Alma Mater Studiorum –Università di Bologna

DOTTORATO DI RICERCA IN

Scienze e Tecnologie Agrarie Ambientali e Alimentari

Ciclo XXXI

Settore Concorsuale: 07/G1 SCIENZE E TECNOLOGIE ANIMALI
Settore Scientifico Disciplinare: AGR/17 ZOOTECNICA GENERALE E MIGLIORAMENTO GENETICO

GENOMIC CHARACTERIZATION OF ITALIAN AND EUROPEAN PIG POPULATIONS

Presentata da: CLAUDIA GERACI

Coordinatore Dottorato | **Supervisore**

Prof. Massimiliano Petracci | **Prof. Luca Fontanesi**

Esame finale anno 2019

Abstract

Over the past ten years, our knowledge about the pig genome has rapidly evolved leading to complete whole genome sequencing, essential for the dissection of molecular basis of pig evolution and various phenotypic traits. As a result of this genomic revolution, we can now take advantage of molecular and bioinformatic tools that allow to study genetic differences or similarities among modern commercial pig breeds, wild boars and local pig populations. This is extremely important in the context of conservation genetics for autochthonous and endangered pig breeds, seen that in many European regions the genetic architecture of local pig populations is still uncharacterized resulting in an unexploited breeding potential. This project aimed to better investigate genomic features of local autochthonous pig breeds focusing analyses on candidate gene markers associated to disease resistance, coat colour and vertebral number and genes potentially involved in feeding preferences.

Considering the economic impact of infectious diseases on the pig production, we used a genotyping approach to define the distribution of disease resistance marker alleles in Italian local pig populations, indirectly confirming, with our results, the robustness of local pig breeds. We also performed an association study between investigated disease resistance markers and production traits, with first results suggesting that it could be possible to introduce disease resistance traits in pig breeding programs without affecting productivity.

Referring to the relationship between local pig populations and wild boars, in the context of the domestication process, we carried out an analysis monitoring the allelic distribution at two evolutionary important loci, involved in coat colour and vertebral number determination. Results of this study suggested that *Sus scrofa* genome is currently experiencing bidirectional introgression of wild and domestic alleles, with autochthonous breeds experiencing a sort of “de-domestication” process and wild resources challenged by a “domestication” drift.

The last part of this project was dedicated to the study of genetic variability of taste receptor genomic regions across different European pig populations. We performed a SNP discovery study to find out similarities and differences in taste sensing system among local breeds. Considering that taste perception is strongly connected to the diet and the environment, comparing different pig breeds and detecting differences in taste receptor genetic sequences, could be informative about the history of breeds and about the impact of ecology in their biodiversity.

Altogether these results can be considered a basis for the use of genetic variability within and among local pig populations and for further studies regarding their full characterization.

Table of contents

General introduction p. 3

Chapter 1: ***The domestic pig*** p. 4

1.1 The origins of pig and the domestication process p. 4

1.2 The molecular perspective in pig domestication p. 6

1.3 Pig genome sequencing p. 9

Chapter 2: ***Pig biodiversity*** p. 11

2.1 Breed differentiation and biodiversity challenge p. 11

2.2 Commercial pig lines p. 14

2.3 Autochthonous pig populations p. 15

Chapter 3: ***The role of genomics in pig breeding*** p. 23

Aim p. 26

Chapter 4: ***Genomics and disease resistance in pigs*** p. 27

4.1 Genetic markers associated with resistance to infectious diseases have no effects on production traits and haematological parameters in Italian Large White pigs p. 29

Chapter 5: ***Phenotypic changes in the domestication process*** p. 51

5.1 Signatures of de-domestication in autochthonous pig breeds and of domestication in wild boar populations from *MC1R* and *NR6A1* allele distribution p. 53

Chapter 6: ***Genomics and nutrient sensing*** p. 75

6.1 DNA pool-seq approach for the identification of single nucleotide polymorphisms in taste receptor genes of European autochthonous pig breeds p. 76

General conclusions p. 100

General introduction

The domestic pig (*Sus scrofa domesticus*) is a member of the Suidae family and it is one of the first animal species in the world to be domesticated. Pig has become over time an important species in animal breeding as it is considered one of the major human nutritional sources of animal-derived proteins (Wang et al., 2017). Pig is also considered an interesting example of domesticated animal whose wild conspecifics have survived and are still living. This is particularly challenging for evolutionary studies together with the great phenotypic varieties represented by existing pig breeds. This variability is the outcome of millions of years of evolution influenced by natural and artificial selections that have impacted morphological, physiological and behavioural traits (Chen et al., 2007).

Domestic pig is not only interesting from an evolutionary point of view, but also plays a central role in biomedical studies and translational medicine. Several researches have been, and are still performed, to investigate on pig as animal model for human diseases. This is possible because pigs and humans are very similar in their physiology and, much more importantly, they share the same disease-causing mutations in genes responsible for severe diseases like Parkinson or Alzheimer or multifactorial traits such as obesity and diabetes (Lunney, 2007; Groenen et al., 2012). The possibility to obtain a huge quantity of molecular data on pig genome thanks to the recent development of Next Generation Sequencing and genotyping technologies, has allowed to discover more about the history of domestic pig and to characterize in time the events marking the separation between early domestication and breeding era (Ramirez et al., 2015).

At the beginning, breeding was intuition-driven, without scientific basis and structured breeding programs. With their early farming activities, humans have strongly modified ancestor wild boars, with selective events that progressively led to the differentiation of modern breeds (Ibáñez-Escriche et al., 2014).

The so called ‘genetic revolution’ in pig history started around 1950, when it became feasible to apply first quantitative genetics approaches and systematic crossbreeding, together with the utilization of estimated breeding values as support to identify animals with the best reproductive qualities for pig industry (Hazel, 1943; Dickerson, 1974; Henderson, 1984; Ibáñez-Escriche et al., 2014). Later, the possibility to employ molecular markers and high-throughput SNP genotyping platforms, led to the discovery of Quantitative Trait Loci (QTLs) and genomic regions, with causal polymorphisms associated with economically important traits for pig industry. In this way, the approach to pig breeding radically changed, opening the path to Genomic Selection (GS) and Marked Assisted Selection (MAS) that are nowadays important tools for pig breeding companies that have implemented breeding schemes and selection strategies (Ibáñez-Escriche et al., 2014).

The dissection of complex traits of economic importance to the pig industry through Genome-Wide Association Analysis (GWAS) is currently a hot topic for researchers. For this reason, studies aiming to identify molecular markers (as causal mutations or polymorphisms in linkage with a QTL) are requested for challenging issues like genetic resistance to diseases, animal performances and productivity (Boddicker et al., 2012; Sanchez et al., 2014; Dunkelberger et al., 2017).

Together with these topics, a lot of studies are presently aimed to define the genetic differences or similarities among modern commercial pig breeds, wild boars and local pig populations. More in details, genomic tools are widely used for conservation genetics in livestock species and for the identification of peculiar traits characterizing local autochthonous pig breeds for defending biodiversity and enhance local pig breeding and production (Čandek-Potokar et al., 2017)

Chapter 1

The domestic pig

1.1 The origins of pig and the domestication process

It is generally accepted that among the six species belonging to the *Sus* genus, *Sus scrofa* is the only one that underwent a fully domestication process from a wild *Sus scrofa* ancestor (Larson et al., 2010). Pig is considered a perfect example of animal species whose wild ancestors are still living, so it represents an incredible opportunity for investigating the history of mammalian evolution and for identifying the signatures of selection occurred during domestication process and natural selection (Chen et al., 2007).

Several studies investigated on the origins of modern pig domestication process and they all agree that the ancestor *Sus scrofa* arose from South East Asia around 3.0-3.5 million years ago (Mya) slowly colonizing and spreading in other regions of Asia and then in European area and North Africa (Groenen et al., 2012; Groenen, 2016). This migration of wild boars to new territories, was probably followed by a long period of geographic isolation corresponding to Calabrian stage, that caused the establishment of two differentiated *Sus scrofa* gene pools, one represented by eastern population (Asia) and the second one gathering western pig groups (Europe, Near East and North Africa) (Ramos-Onsins et al., 2014).

Recent advances in evolutionary genetics proved that by comparing phylogenomic data coming from wild boars and domestic pigs, it is possible to identify two distinct Asian and European lineages that probably diverged during the mid-Pleistocene, around 0.8-1 Mya ago (Giuffra et al., 2000; Fang and

Andersson, 2006; Groenen et al., 2012). During that period, in fact, the geographical distribution of wild boars was remodelled because of significative climate changes that led to the extinction of some *Sus scrofa* populations and to the migration and isolation of some others across Eurasia.

Even if the European clade population seemed to increase after the colonization of these new territories, a consistent drop in the population size was recorded, partially involving the Asian wild boars population, during the Last Glacial Maximum (LGM, around 20.000 years ago) and this is considered the main explanation of the low genetic variability found among modern European wild boars (Groenen et al., 2012; Wang et al., 2017).

The two different clades, the Asiatic and the European one, separated long before the advent of domestication (roughly around 500.000 years ago according to Giuffra et al., 2000) and there are genetic evidences confirming that the two populations of wild boars were domesticated independently, with Asian domestic pigs more recently (around 18th century) hybridized with European pig breeds for their improvement (Giuffra et al., 2000; Larson et al., 2010).

According to population geneticists, two main theories have been suggested for the explanation of pig domestication process, and the two models basically differ in the definition of number of putative domestication sites. These latter are restricted to Near East and China, according to the first proposed model, and extended to Neolithic Europe and Jomon period in Japan according to the second theory (Larson et al., 2010). Additionally, this second model, proposed several and independent events of domestication in the mentioned areas, and it is considered the most reliable according to some scientific evidences that report the absence of Near Eastern wild boar haplotypes in modern domesticated European pigs (Larson et al., 2005).

According to the most ancient archaeological evidence of pig found in Anatolia, the process of domestication from *Sus scrofa* wild boars started approximately 9.000 years ago with two main domestication events occurring independently in multiple regions of the world (Giuffra et al., 2000; Larson et al., 2005).

Regarding the pig domestication process in Asia, China, is considered the main domestication centre with multiple native breeds starting their domestication process around 8.000 years ago. On the counterpart, European pigs were originally domesticated in the region Near East, spreading later to Europe because of human migration (Wang et al., 2017).

From a demographical point of view, Asiatic and European lineages, even if geographically separated, were both influenced by human activity starting from late Quaternary until Neolithic Demographic Transition. This was due to first hunting and agricultural activities and human

colonization of new areas (Wang et al., 2017). Demographic analysis showed that the strongest distinction between the two clades is dated back to 4,500–7,000 years ago with the two populations reporting different population size, furthermore confirming the primary and independent domestication in Europe and China, respectively.

Domestication process can be explained as a mutualistic long-term relationship between humans and plant or animal species that leads to advantages for both parties (Zeder et al., 2006). The impact of this evolutionary process implied changes from genetic, morphophysiological, and behavioural point of view in animals, for the satisfaction of human needs (Ramos-Onsins et al., 2014). Domestication of animals, indeed, was a progressive and complex process that shaped the most important socioeconomic transitions in human history: domestic pigs have probably played a role in the transition of hunter-gatherer nomadic communities to an agricultural and sedentary lifestyle (Larson et al., 2005; Ramos-Onsins et al., 2014).

The process of domestication has been described as a progressive and cumulative interdependent relationship between human societies and plants or animal populations. Most advantageous phenotypes useful for human activities were probably identified by human communities that acted through artificial selection on animal populations defining homogeneous groups of individuals with heritable characteristics different from the wild species of origin (Diamond, 2002). In the case of pigs, this strong selection pressure on wild boar species under domestication, resulted in behavioural change, modification of morphological architecture and changes in physical types. In other words, pig domestication was a genetic adaptation to better satisfy and meet the human needs involving allelic frequency changes (Wang et al., 2017).

Identifying the genes and the allelic variants that led to this transformation of wild boars through domestication and artificial selection, is considered crucial to track evolutionary and genetic history of modern domestic pig.

1.2 The molecular perspective in pig domestication

A multidisciplinary approach is requested to identify and define markers that can trace the process of domestication and identify primary domestication sites and migratory routes, with contributions coming mainly from archaeology, population genetics and biology (Zeder et al., 2006; Ramos-Onsins et al., 2014).

In general, the comparison between wild and domesticated populations is the first step to define differences and to identify the putative domestication centres, keeping in mind that there is no always direct correspondence between true ancestor wild populations and the contemporary wild one (MacHugh et al., 2017).

In the case of pigs, specifically, different events of local introgression from wild boars at the beginning of domestication process, caused a gene flow between European wild boars and pigs at the time of human migration, partially masking the signatures of domestication in their genomes (Larson et al., 2005).

The first evidences about pig domestication process were largely derived by analysis of microsatellites and mitochondrial DNA of wild and domestic pigs from Asia and Europe (Giuffra et al., 2000). mtDNA is considered powerful source for genealogical and evolutionary studies of animal populations, and it was used for screening variation in samples representing different populations, for defining phylogenetic structure of mitochondrial sequences in wild and domestic pigs. The use of mitochondrial DNA for the definition of haplotypes revealed, for example, the presence of two core pig haplotypes that are detectable only in European-derived pig populations, that suggest two independent events of domestication of at least two European wild boar lineages (Zhang and Hewitt, 2003; Larson et al., 2005).

Sequencing the hypervariable region of the mtDNA is a source of important evolutionary information since its evolution rate is more rapid than that of nuclear DNA and does not recombine (Wolstenholme, 1992). However, mitochondrial markers are poor predictors of whole-genome variation: first, because of the presence of mitochondrial pseudogenes in the nuclear genome, secondly, in mtDNA there is a loss of information regarding male-mediated gene flows, and finally mitochondrial markers are susceptible to genetic drift (Zhang and Hewitt, 2003).

For this reason, phylogenetic analysis needed to be also supported by nuclear genetic information, specifically targeting microsatellites and single nucleotide polymorphisms (SNPs).

Microsatellites (known also as simple sequence repeats, SSRs or short tandem repeats, STRs) have been widely used as molecular markers for the definition of variation in pigs (Zhang and Hewitt, 2003). In phylogenetic analysis, they have been investigated for their role in regulation of genes (as they have been proposed to modify genes they are associated to) and for measuring parentage and relatedness within breeds (Laval et al., 2000; SanCristobal et al., 2006). SSRs are frequent in the genome and they present a high level of heterozygosity. Because of the complicated relationship between microsatellite alleles, high mutation rates and questionable neutrality as markers, their

reliability in phylogenetic studies has been discussed a lot (Zhang and Hewitt, 2003; Chen et al., 2007).

Single nucleotide polymorphisms (SNPs) as biallelic markers are less informative than microsatellites, but they have been used largely in genotyping SNP panels that allow, for a single-reaction assay, to detect simultaneously thousands to millions of SNPs. These high-throughput SNP panels are cost effective and they have radically changed the approach to genome research for most livestock species, including pigs (Ramos et al., 2009). High density SNP chips have been developed for investigating on association, admixture and identity by descent mapping, for defining phenotype/genotype associations and for population genetic studies aimed to analyse population structure (Albrechtsen et al., 2010).

The development of such powerful tools for high-density SNP genotyping is consequential to the availability of significant number of SNPs, which is, in turn, linked to the recent development of Next Generation Sequencing (NGS) technology.

The advent of NGS has truly revolutionized genetics: the possibility to cheaply produce a huge amount of data, has totally changed the approach to both, basic and applied research (Metzker, 2010). Most of the applications of next generation sequencing methods to pig genomics are linked to SNP discovery, as it is a cost-effective approach compared to traditional Sanger sequencing (Amaral et al., 2011). Resequencing of divergent populations, as for example wild boars and domestic pigs, has allowed to identify, in the wild boar, regions with low or high variability that are potential targets of selection, like genes that affect reproduction, coat colour, body size, growth and metabolism (Rubin et al., 2012). Genome scans results have also proved that series of loci were under strong selection, during domestication process leading to the phenotypic evolution of European domestic pigs.

The use of NGS and genomic data can reveal the microevolutionary aspects that underlie the phenotypic changes of domestic pigs thanks to the powerful contribution of Paleogenomics, that is the brunch of genomics focusing on the study of ancient DNA (aDNA) from archaeological remains (Ramirez et al., 2015). Artificial selection, caused by breeding processes and activities, has dramatically sculpted pig genome diversity in a very short-time frame. For this reason, performing sequencing of aDNA extracted from fossils, can give an insight on natural selection and proofs of admixture between early domestic animal populations and their wild congeners (MacHugh et al., 2015).

The use of molecular data to study the domestication process is particularly useful to distinguish primary domestication events from more recent ones, tracking domesticated animals and the degree

of variability and phylogenetic discontinuity, especially among different pig breeds (Zeder et al., 2006).

1.3 Pig genome sequencing

Sequencing of pig genome started with the establishment of the Swine Genome Sequencing Consortium (SGSC) in September 2003, involving academic and industry representatives and leading to the assembly and publication of the first draft of reference genome of *Sus scrofa* in 2012 (Schook et al., 2005; Groenen at al., 2012; Groenen, 2016). The published paper reported the analysis of the reference genome of domestic Duroc pig breed, in a comparative study with several wild boars and domesticated pigs coming from Europe and Asia. The results revealed a deep phylogenetic separation between the European and Asian wild boars, dated 1 Mya reporting signatures of selection in genes involved in RNA processing and regulation, and in genomic regions associated with immunity and olfactory sensing. The sequencing output showed that most of the genetic variability between European and Asian samples, is detectable in high recombination regions. In addition, analysis revealed a long-time exchange of genetic material between the domesticated pigs and the wild boars and between the two lineages, with European breeds presenting 35% of Asian fraction (Groenen et al., 2012).

Taking advantage of NGS technologies and molecular data, it was possible to identify some crucial points in the history of pig domestication that marked the remodelling of pig genome. For example, during the late medieval and early modern era, some important changes occurred in the health and size status of domestic pigs probably due to farming strategies that from extensive and uncontrolled, became intensive and stable, limiting the possibility of crossing domesticated pigs with wild animals (Ramirez et al., 2015).

The advent of high throughput sequencing platforms has also provided the path to dissect phenotypic evolution in domestic pigs' traits like behaviour, body composition, reproduction, and coat colour.

Rubin et al. (2012) performed a selective sweep analyses to investigate on three loci harbouring quantitative trait loci that are involved in one of the most important phenotypic change of domestic pig, which is the elongation of the back and the increased number of vertebrae. They looked for non-synonymous mutations that became fixed in three genes (*NR6A1*, *PLAG1*, and *LCORL*) that are known to be associated with stature and body length in other domestic animals. The analysis was based on genetic comparison of samples coming from different European pig populations, European wild boars and Asian domestic pigs. Results showed that most of the commercial European breeds

are homozygous for the same haplotype at these loci with the exception of Iberico breed, remarking the genetic separation between Asiatic and European lineages and the absence of strong modifications in Iberico genome that results in a close relationship with ancestor genome (Rubin et al., 2012; Ramirez et al., 2015).

Another phenotypic trait that has changed during domestication process is coat colour. Coat colour is a monogenic trait and different studies have been performed to identify causative mutations for pigmentation in different species, including pigs (Andersson and Georges, 2004). Coat colour feature differs a lot between domestic pig breeds and the one observed in wild boars, that present brown pigmentation for helping them to camouflage from their predators (Ramos-Onsins et al., 2014). Two most important genes have been investigated for their role in pigmentation, *MC1R* and *KIT*. Fang et al. (2009) explored the genetic variability at *MC1R* locus among wild and domestic pigs from both Europe and Asia, pointing out that even if the same numbers of mutations were reported, they acted differently in wild and domestic, with this latter presenting mutations affecting protein sequence and final coat colour. *KIT* gene and its regulatory elements are known to be responsible for different coat colour phenotypes in pigs, like dominant white, patch, and belt (Groenen et al., 2016; Fontanesi et al., 2016). The complexity of this locus has been investigated by Rubin et al. (2012) that demonstrated that the number of duplications at this site is breed specific and that two of them are exclusively present in white and white spotted pigs.

Many other polygenic traits, related to production, have been intensively studied in pig genome, like *IGF2* gene, which is associated with differences between wild boars and domestic pigs in muscle development, back fat and heart size. QTL analyses proved that *IGF2* locus is a unique adaptive example from pig production point of view because it has no effect on birth weight and it supports muscle growth after birth (Andersson and Georges, 2004). *IGF2* is considered a gene recently affected by selection pressure and it carries a specific allele in most European domestic pigs under intense selection scheme for muscle growth and reduced fat depositions (Rubin et al., 2012).

Information gathered from all these studies represent an example of the output of the current use of genome sequencing technology that is accelerating the discovery of causative genes and polymorphisms specifically involved in the genetic control of economically important traits.

Chapter 2

Pig biodiversity

2.1 Breed differentiation and biodiversity challenge

The independent domestication events of wild boars in Europe and Far East regions, brought to the definition of different gene pools that later in time colonized new areas like Africa and South America, together with human economic expansion and new commercial routes to these two continents (Ramirez et al., 2009).

This spread of genetic resources put the basis for the diversification of livestock breeds and the definition of local types, different from ancestor wild boars and modern ones.

As a matter of fact, the first data collected and available on pig history have proved that modern European pig breeds and Near Eastern wild boars did not share mitochondrial haplotypes, thus suggesting that they probably descended from wild boars that had been domesticated locally (Larson et al., 2005; Manunza et al., 2013). However, the use of whole genome data has proved that the absence of shared features in mitochondrial DNA does not necessarily mean that there is no connection between Asian and European pigs, and that is possible to detect markers in Y and autosomal chromosomes, proving Asian introgressed haplotypes in European pig genes responsible for meat quality, development and fertility (Bosse et al., 2014).

This is the reason why European commercial pig breeds are proved to carry Asian haplotypes in their genomes. This is true for most European populations but not for the local ones (especially for Iberico breed), that present a different conservation status and for which genomic analyses report a lower variation of Asian-derived haplotypes, maybe for not recurrent introgression events or for their mixed European origins (Herrero-Medrano et al., 2013; Bosse et al., 2015).

During 19th century the first gene flow of modern time was recorded between the two pig lineages, with the introduction of Chinese pigs in European population aimed to improve productivity of the European local breeds. This happened because Chinese pigs (Meishan breed in particular) were known to be better in reproduction, meat quality and resistance to diseases. They were then crossed with local breeds, like Large White pigs (also known as Yorskshire pigs) producing hybrids with improved phenotypes (Bosse et al., 2014). It is well documented that this crossbreeding took place mainly in England where pioneers breeders started programs for the improvement of pig breeds in response to the growing social request for meat. This process led to the differentiation of new distinct pig lines reporting new desired phenotypes (Laval et al., 2000; Wilkinson et al., 2013).

Molecular tools available today allow to study these phenotypes, to track pig history and to define the genomic regions that have contributed to the phenotypic diversity of modern breeds in terms of shape, colour, size, production and reproduction performances and that make them extremely different from the wild population (Amills, 2011).

Starting from 19th century with further selection processes during 20th century, pig breeds have been genetically improved for meat quality, muscularity and reproductive attitude. This strong artificial selection led to the institution of official breed standard traits, both in terms of physical features (coat colour, ear morphology) and production traits (intramuscular fat content, backfat depth, carcass weight and length) (BPA, 2002).

The introduced genetic features led to the possibility to improve meat quality as main important output. The first breeding companies emerged around middle of 20th century, with the purpose of genetically improve breeds for meat production creating synthetic or commercial pig lines remodelled for lean meat content, muscularity and enhanced reproduction in response to consumers market (Cesar et al., 2010). This breeding strategy however has caused a loss in genetic diversity because of the utilization of a limited number of sires for reproduction (based on their estimated breeding value, EBV) and a progressive fixation of some valuable alleles over time. This is particularly evident in some international pig lines like Landrace and Large White that have been pushed for increased growth and meat production performance (Ojeda et al., 2008; Ramirez et al., 2009).

Swine industries pursued productivity as a major goal, establishing intensive breeding schemes focused on increase production rates and reduce production costs, with a focus on quantitative and qualitative carcass traits during growth and finishing phase, on daily weight gain feature and feed conversion ratio. In today's swine production the standardization of genetic resources is linked to specific breeds employed for their satisfactory performance and carcass characteristics with Duroc, Large White, Pietrain, Hampshire and Belgian Landrace breeds as main representatives (Carvalho de O. et al., 2016).

Commercial pig lines are standardized for specific traits and for this reason, they differ from European autochthonous pig breeds. This relationship has been investigated by several authors (Laval et al., 2000; Ollivier et al., 2005; SanCristobal et al., 2006; Ollivier, 2009; Nidup and Moran 2011). These analyses were mainly focused on European pigs and they were based on microsatellites detection reporting that the contributions of individual breed to European between-breed diversity, was recorded in a range from 0.04% to 3.94% and that local breeds instead, accounted for 56% of the total, with lower percentages coming also from commercial lines and international breeds.

Conversely, higher contribution to within-breed diversity, was given by commercial breeds followed by international ones (Nidup and Moran, 2011).

These two components of diversity and their respective importance are linked to priorities in pig breeding, whether it is aimed to genetic improvement by selection or adaptability to production systems. Therefore, local breeds should be prioritized for between-breeds diversity, while synthetic lines are important for maintaining genetic variability in within-breeds diversity (Ollivier et al., 2005).

Pig breed diversity is an important issue that deserves to be monitored, supported and pursued. At present, there are official European and FAO databases recording number of individuals per pig breed. In 2011, according to a study from Nidup and Moran, more than 730 breeds or pig lines were estimated, the majority of which were found in China and Europe. However, since 2006, a high percentage (nearly 30%) of breeds are considered in danger for extinction because of breed formation and artificial selection (FAO, 2006).

The dynamics of this genetic erosion is explained by globalization process that affects the decline of pig breeds through changes in food demand, regulations and importation policies and marketing activities of worldwide breeding companies. This results in a lack of economic profitability of local pig breeds when compared to other breeds or their crosses, with a dilution in local pig breeds innate adaptive skills and a huge loss of genetic variability and cultural values represented by autochthonous breeds (Biodiversity Brief 10, EC 2001; Gandini and Oldenbroek, 2007).

Conservation of agrobiodiversity through the management of indigenous pig breeds and their wild relatives, is extremely important not only for their role as food, but also because they represent genetic reservoirs (Yang et al., 2017).

Precise and standardized tools are needed to quantify genetic diversity in European livestock and for the conservation of a rich pool of genetic resources. Microsatellites and SNPs markers have shown to be successful in the evaluation of variability among breeds, but they need to be further developed for establishing conservation programs in all European countries (Laval et al., 2000). High-throughput genomic technologies have been only partially addressed to population studies for dissecting phenotypes among European pig breeds, so they would be advantageous in the evaluation of their genetic characteristics and in the definition of the basis of their resilience and adaptive evolution to the environment.

Pig biodiversity conservation programs strongly require the support of genomic tools for variability discovery and for defining priorities for conservation and management politics both in a short-term commercial context and in a longer-term perspective of maintaining the species biodiversity.

2.2 Commercial pig lines

The role of pigs in human society has changed through history, transforming them from being the beneficiary of the 'leftovers' of human food, to become the most popular source of animal protein in human nutrition (Buchanan and Stalder, 2011). This is linked to the increasing demand for pork raised during recent years in developing countries that has impacted pig production worldwide.

Over time, pig production system has been shaped by human intervention, with pigs changing completely rearing environmental conditions, from forest to pasture-based systems to modern industrialized systems where pigs are kept indoors in dedicated buildings (McGlone J.J., 2013)

Today, pork production can be divided in two main production systems, the traditional one, connected to the breeding and utilization of small local pig populations (most of the times also endangered breeds) and the industrial farming system, which is intensive and aimed to maximize the productivity and the efficiency.

Most pigs used in intensive systems are based on the crossing of two breeds, Large White and Landrace, with a third breed like Duroc or Hampshire used as a terminal sire. These commercial breeds have been selected for maximizing reproductive traits, like the number of piglets they can annually produce, and production traits, like rapid growth rates and high food conversion ratios.

The genetic merit is a key factor for pig breeding companies, therefore commercial pigs are usually the results of crossbred lines that show better production performances when they are compared to original purebred parents (Dekkers, 2007). Using a crossbreeding strategy, pig industry (like industries working on other livestock species) takes advantage of heterosis effects and synergy between parental lines, selecting pigs mainly for growth, carcass and meat quality in the case of meat production. Indeed, farmed pigs are hybrid animals derived from genetic lines, very specialized and selected for their production and market value. Pig lines are developed by companies and they are maintained in terms of integrity, uniformity, productivity and genetic merit (Buchanan and Stalder, 2011). The established breeding schemes are based on the genetic merit evaluation of candidate sire and dam lines derived by breeding values. The possibility offered by new genomic tools is extremely important for pig industry because they allow to get more information about reproducers and support breeders in the identification of the molecular mechanisms affecting economically important traits in pigs (Ernst and Steibel, 2013; Knol et al., 2016).

The idea of improving performances of pigs by managing genetic information gathered from this technology is profitable for pig industry but it is not devoid of side-effects.

The current strategy in pig breeding, that is based on a limited number of breeds for meat production, has raised, indeed, a lot of issues concerning animal welfare, environmental impact, sustainability and biodiversity conservation. A strong contrast is detectable between breeds like the Large White

that still represents almost one third of the total meat produced and consumed in Europe and local pig breeds that are almost at risk for the extinction but potentially efficient from a production point of view (Ollivier et al., 2005).

2.3 Autochthonous pig populations

Local pig breeds embody the identity and the agricultural system of the geographical area they belong to. They represent a cultural value because of their historical role in the definition and maintenance of landscape and gastronomy, and they are deeply connected to folklore and local traditions. For these reasons the uniqueness of traditional breeds must be considered a heritage to be protected, defended and supported (Gandini and Villa, 2003; Gandini and Oldenbroek, 2007). Until today, the merit for the survival of these breeds, (that are mostly considered at the risk for extinction) is to be recognized to local pig farmers that still see value in local resources and try to conserve them.

In addition to their cultural merit, local pig breeds offer valuable, but untapped, economic potential with their production traits, that need to be further investigated for developing better management and breeding programs that could be beneficial, in turn, for their protection and their conservation.

The investigation of genetic structure and background of local and traditional European pig breeds is considered challenging and interesting by the pig sector, seen that they are generally well adapted to specific local agro-climatic environments and production systems. However, only few cases of successful local breed chains exist in Europe and the current state of pig genetic resources in European area is represented by a mix of many local breeds, mostly rare, whose level of genetic diversity remains not properly exploited (Ollivier et al., 2005).

In general, there are two problems linked to the assessment of pig diversity: the first is the collection of specimens of a reasonable number of individuals per breed (they are very small populations), the second, that is also the most limiting factor is cost of genetic characterization (Megens et al., 2008).

To address this problem, different European collaborative projects have been funded for the establishment of networking activities and research collaborations aiming to describe distinctiveness of pigs from different parts of Europe. Among these projects, TREASURE project (https://treasure.kis.si; https://cordis.europa.eu/project/rcn/193290_en.html) was funded to explore the diversity of local pig breeds and their production systems to develop sustainable pork chains that promote high quality traditional products derived from these pigs.

By providing reliable data (e.g. biological specimens, phenotypic characterization, production traits records, demographic information) necessary to study the genetic potential of endangered pig breeds,

this kind of projects aim to describe local pigs at phenotypic, genomic and functional level for a better management of these populations in their ecological niche.

Genomic data resulting from these studies could be exploited within a short time for improving management practices and marketing strategies and for promoting inherent value of regional pig breeds in terms of biodiversity resources and improved breeding programs. This could also result in highlighting the market potential and value of products that are traditionally linked to a specific territory, fulfilling social issues like animal welfare, environmental impact of production and sustainable exploitation of locally available resources.

The main peculiarity of studying autochthonous pig breeds is that they show adaptability to specific local conditions and environments, high potential for fat deposition and high meat quality in terms of intramuscular and intermuscular fat content. For this reason, identifying most relevant genes and mutations associated with adaptive traits like pig morphology and productivity, or reproductive and disease resistance traits, could be powerful for optimizing the breeding strategies.

In this context, the genomic information gathered on local pig breeds, the knowledge of potentially negative alleles and genetic basis responsible for specific adaptive traits (resilience, robustness), the development of marker panels (reasonably priced tools) useful to solve traceability problems would represent innovation potential to serve, set up or improve breeding programs, in-situ management of genetic resources, and contribute to a better utilization of local pig breeds. Genetic and genomic tools that are presently not available could be developed and used by the farmers to better exploit local pig breeds. New breeding programs and management of local pig populations need to rely on genomic and molecular information.

References

Albrechtsen, A., Nielsen, F.C., Nielsen, R., 2010. Ascertainment biases in SNP chips affect measures of population divergence. Mol. Biol. Evol. 27, 2534–47. doi:10.1093/molbev/msq148

Amaral, A.J., Ferretti, L., Megens, H.-J., Crooijmans, R.P.M.A., Nie, H., Ramos-Onsins, S.E., Perez-Enciso, M., Schook, L.B., Groenen, M.A.M., 2011. Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. PLoS One 6, e14782. doi:10.1371/journal.pone.0014782

Amills, M., 2011. Biodiversity and Origin of Pig Breeds. Bulletin UASVM Animal Science and Biotechnologies, 68, 1-2

Andersson, L., Georges, M., 2004. Domestic-animal genomics: deciphering the genetics of complex traits 5. doi:10.1038/nrg1294

Biodiversity Brief 10, 2001. Biodiversity in Development- Livestock and biodiversity. European Commision

Boddicker, N., Waide, E.H., Rowland, R.R.R., Lunney, J.K., Garrick, D.J., Reecy, J.M., Dekkers, J.C.M., 2012. Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. J. Anim. Sci. 90, 1733–1746. doi:10.2527/jas2011-4464

Bosse, M., Megens, H.-J., Frantz, L.A.F., Madsen, O., Larson, G., Paudel, Y., Duijvesteijn, N., Harlizius, B., Hagemeijer, Y., Crooijmans, R.P.M.A., Groenen, M.A.M., 2014. Genomic analysis reveals selection for Asian genes in European pigs following human-mediated introgression. Nat. Commun. 5, 4392. doi:10.1038/ncomms5392

Bosse, M., Madsen, O., Megens, H.J., Frantz, L.A.F., Paudel, Y., Crooijmans, R.P.M.A., Groenen, M.A.M., 2015. Hybrid origin of European commercial pigs examined by an in-depth haplotype analysis on chromosome 1. Frontiers in Genetics. 5, 442-451. doi: 10.3389/fgene.2014.00442

BPA, 2002. British Pig Breeds. Trumpington, Cambridge: British Pig Association

Buchanan, D.S. , Stalder, K., 2011. Breeds of Pigs. In: Rothschild, M. and Ruvinsky, A. (Eds.) The genetics of the Pig. 2nd edition, CABI, 2011

Čandek-Potokar, M., Nieto, R., Pugliese, C., Araujo, J.P., Charneca, R., Garcia Casco, J.M., González Sánchez, E., Hernandez-Garcia, F.I., Izquierdo, M., Karolyi, D., Kušec, G., Lebret, B., Mercat, M.J., Petig, M., Radović, C., Savić, R., 2017. Local Pig Breeds: Nutritional Requirements, Innovative Practices and Local Feeding Resources as Challenges in Project TREASURE. https://zenodo.org/record/1136173

Carvalho de O., P.L., Soares de V., E., Ansolin, R.D., Lorscheiter, K.A., da Silva Chambo da S., P.C., Carvalho, S.T. , Almeida Santana, A.L., Costa Oliveira da C., A., Luiz Genova, J.L., 2016. Performance and carcass traits of different commercial pig lines. Ciênc. Agrár. 37, 3317-3326. doi: 10.5433/1679-0359.2016v37n5p3317

Cesar, A.S.M. Silveira, A.C.P., Freitas, P.F.A. Guimarães, E.C., Batista, D.F.A., Torido, L.C.,

Meirelles, F.V., Antunes, R.C., 2010 Influence of Chinese breeds on pork quality of commercial pig lines. Genet. Mol. Res. 9, 2, 727-733. doi: 10.4238/vol9-2gmr733

Chen, K., Baxter, T., Muir, W.M., Groenen, M.A., Schook, L.B., 2007. Genetic resources, genome mapping and evolutionary genomics of the pig (Sus scrofa). Int. J. Biol. Sci. 3, 153–65. doi:10.7150/ijbs.3.153

Dekkers, J.C.M., 2007. Marker-assisted selection for commercial crossbred performance1. J. Anim. Sci. 85, 2104–2114. doi:10.2527/jas.2006-683

Diamond, J., 2002. Evolution, consequences and future of plant and animal domestication. Nature 418, 700–707. doi:10.1038/nature01019

Dickerson, G.E., 1974. Evaluation and utilization of breed differences. In: Proceedings of the workingsymposium on breed evaluation and crossing with farm animals. Zeist, September 1974.

Dunkelberger, J.R., Serão, N.V.L., Weng, Z., Waide, E.H., Niederwerder, M.C., Kerrigan, M.A., Lunney, J.K., Rowland, R.R.R., Dekkers, J.C.M., 2017. Genomic regions associated with host response to porcine reproductive and respiratory syndrome vaccination and co-infection in nursery pigs. BMC Genomics 18, 1–19. doi:10.1186/s12864-017-4182-8

Ernst, C.W., Steibel, J.P., 2013. Molecular advances in QTL discovery and application in pig breeding. Trends Genet. 29, 215–224. doi:10.1016/J.TIG.2013.02.002

Fang, M., Andersson, L., 2006. Mitochondrial diversity in European and Chinese pigs is consistent with population expansions that occurred prior to domestication. Proc. R. Soc. B Biol. Sci. 273, 1803–1810. doi:10.1098/rspb.2006.3514

Fang, M., Larson, G., Soares Ribeiro, H., Li, N., Andersson, L., Barsh, G.S., 2009. Contrasting Mode of Evolution at a Coat Colour Locus in Wild and Domestic Pigs. doi:10.1371/journal.pgen.1000341

FAO, 2006. The State of World's Animal Genetic Resources for Food and Agriculture. FAO: Rome

Fontanesi, L., Scotti, E., Gallo, M., Nanni Costa, L., Dall’Olio, S., 2016. Authentication of “mono-breed” pork products: Identification of a coat colour gene marker in Cinta Senese pigs useful to this purpose. Livest. Sci. 184, 71–77. doi:10.1016/J.LIVSCI.2015.12.007

Gandini, G. and Oldenbroek, K. in Oldenbroek, K. (Ed.), 2007. Utilisation and conservation of farm animal genetic resources. Wageningen Academic Publishers, The Netherlands. doi:10.3920/978-90-8686-592-5

Gandini, G.C., Villa, E., 2003. Analysis of the cultural value of local livestock breeds: a methodology. J. Anim. Breed. Genet. 120, 1–11. doi:10.1046/j.1439-0388.2003.00365.x

Giuffra, E., Kijas, J.M., Amarger, V., Carlborg, O., Jeon, J.T., Andersson, L., 2000. The origin of the domestic pig: independent domestication and subsequent introgression. Genetics 154, 1785–91.

Groenen, M.A.M., Archibald, A.L., Uenishi, H., Tuggle, C.K., Takeuchi, Y., Rothschild, M.F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H.-J., Li, S., Larkin, D.M., Kim, H., Frantz, L.A.F., Caccamo, M., Ahn, H., Aken, B.L., Anselmo, A., Anthon, C., Auvil, L., Badaoui, B.,

Beattie, C.W., Bendixen, C., Berman, D., Blecha, F., Blomberg, J., Bolund, L., Bosse, M., Botti, S., Bujie, Z., Bystrom, M., Capitanu, B., Carvalho-Silva, D., Chardon, P., Chen, C., Cheng, R., Choi, S.-H., Chow, W., Clark, R.C., Clee, C., Crooijmans, R.P.M.A., Dawson, H.D., Dehais, P., De Sapio, F., Dibbits, B., Drou, N., Du, Z.-Q., Eversole, K., Fadista, J., Fairley, S., Faraut, T., Faulkner, G.J., Fowler, K.E., Fredholm, M., Fritz, E., Gilbert, J.G.R., Giuffra, E., Gorodkin, J., Griffin, D.K., Harrow, J.L., Hayward, A., Howe, K., Hu, Z.-L., Humphray, S.J., Hunt, T., Hornshøj, H., Jeon, J.-T., Jern, P., Jones, M., Jurka, J., Kanamori, H., Kapetanovic, R., Kim, J., Kim, J.-H., Kim, K.-W., Kim, T.-H., Larson, G., Lee, K., Lee, K.-T., Leggett, R., Lewin, H.A., Li, Y., Liu, W., Loveland, J.E., Lu, Y., Lunney, J.K., Ma, J., Madsen, O., Mann, K., Matthews, L., McLaren, S., Morozumi, T., Murtaugh, M.P., Narayan, J., Truong Nguyen, D., Ni, P., Oh, S.-J., Onteru, S., Panitz, F., Park, E.-W., Park, H.-S., Pascal, G., Paudel, Y., Perez-Enciso, M., Ramirez-Gonzalez, R., Reecy, J.M., Rodriguez-Zas, S., Rohrer, G.A., Rund, L., Sang, Y., Schachtschneider, K., Schraiber, J.G., Schwartz, J., Scobie, L., Scott, C., Searle, S., Servin, B., Southey, B.R., Sperber, G., Stadler, P., Sweedler, J. V., Tafer, H., Thomsen, B., Wali, R., Wang, J., Wang, J., White, S., Xu, X., Yerle, M., Zhang, G., Zhang, J., Zhang, J., Zhao, S., Rogers, J., Churcher, C., Schook, L.B., 2012. Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491, 393–398. doi:10.1038/nature11622

Groenen, M.A.M., 2016. A decade of pig genome sequencing: a window on pig domestication and evolution. Genet. Sel. Evol. 48, 23. doi:10.1186/s12711-016-0204-2

Hazel, L.N., 1943. The Genetic Basis for Constructing Selection Indexes. Genetics 28, 476–90.

Henderson, C.R., 1984. Applications of Linear Models in Animal Breeding. University of Guelph, Guelph, Ont.

Herrero-Medrano, J.M., Megens, H.-J., Groenen, M.A.M., Ramis, G., Bosse, M., Pérez-Enciso, M., Crooijmans, R.P.M.A., 2013. Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. BMC Genet. 14, 106. doi:10.1186/1471-2156-14-106

Ibáñez-Escriche, N., Forni, S., Noguera, J.L., Varona, L., 2014. Genomic information in pig breeding: Science meets industry needs. Livest. Sci. 166, 94–100. doi:10.1016/j. livsci.2014.05.020

Knol, E.F., Nielsen, B., Knap, P.W., 2016. Genomic selection in commercial pig breeding. Anim. Front. 6, 15–22. doi:10.2527/af.2016-0003

Larson, G., Dobney, K., Albarella, U., Fang, M., Matisoo-Smith, E., Robins, J., Lowden, S., Finlayson, H., Brand, T., Willerslev, E., Rowley-Conwy, P., Andersson, L., Cooper, A., 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science 307, 1618–21. doi:10.1126/science.1106927

Larson, G., Liu, R., Zhao, X., Yuan, J., Fuller, D., Barton, L., Dobney, K., Fan, Q., Gu, Z., Liu, X.-H., Luo, Y., Lv, P., Andersson, L., Li, N., 2010. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. Proc. Natl. Acad. Sci. U. S. A. 107, 7686–91. doi:10.1073/pnas.0912264107

Laval, G., Iannuccelli, N., Legault, C., Milan, D., Groenen, M.A., Giuffra, E., Andersson, L., Nissen, P.H., Jørgensen, C.B., Beeckmann, P., Geldermann, H., Foulley, J.-L., Chevalet, C., Ollivier, L., 2000. Genetic diversity of eleven European pig breeds. Genet. Sel. Evol. 32, 187.

doi:10.1186/1297-9686-32-2-187

Lunney, J.K., 2007. Advances in swine biomedical model genomics. Int. J. Biol. Sci. 3, 179–84.

MacHugh, D.E., Larson, G., Orlando, L., 2017. Taming the Past: Ancient DNA and the Study of Animal Domestication. Annu. Rev. Anim. Biosci. 5, 329–351. doi:10.1146/annurev-animal-022516-022747

Manunza, A., Zidi, A., Yeghoyan, S., Balteanu, V.A., Carsai, T.C., Scherbakov, O., Ramírez, O., Eghbalsaied, S., Castelló, A., Mercadé, A., Amills, M., 2013. A high throughput genotyping approach reveals distinctive autosomal genetic signatures for European and Near Eastern wild boar. PLoS One 8, e55891. doi:10.1371/journal.pone.0055891

McGlone, J.J., 2013. The Future of Pork Production in the World: Towards Sustainable, Welfare-Positive Systems. Anim. an open access J. from MDPI 3, 401–15. doi:10.3390/ani3020401

Megens, H.-J., Crooijmans, R.P.M.A., San Cristobal, M., Hui, X., Li, N., Groenen, M.A.M., 2008. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. Genet. Sel. Evol 40, 103–128. doi:10.1051/gse:2007039

Metzker, M.L., 2010. Sequencing technologies — the next generation. Nat. Rev. Genet. 11, 31–46. doi:10.1038/nrg2626

Nidup, K., Moran, C., 2011. Genetic diversity of domestic pigs as revealed by microsatellites: a mini review. Genom. and Quantit. Genet. 2, 5-18

Ojeda, A., Huang, L.S., Ren, J., Angiolillo, A., Cho, I.C., Soto, H., Lemús-Flores, C., Makuza, S.M., Folch, J.M., Pérez-Enciso, M., 2008. Selection in the making: a worldwide survey of haplotypic diversity around a causative mutation in porcine IGF2. Genetics. 178, 3, 1639-52. doi: 10.1534/genetics.107.084269

Ollivier, L., Alderson, L., Gandini, G.C., Foulley, J.L., Haley, C.S., Joosten, R., Rattink, A.P., Harlizius, B., Groenen, M.A.M., Amigues, Y., Boscher M.Y., Russell, G., Law, A., Davoli, R., Russo, V., Matassino, D., Désautés, C., Fimland, E., Bagga, M., Delgado, J.V., Vega-Pla, J.L., Martinez, A.M., Ramos, A.M., Glodek, P., Meyer, J.N., Plastow, G.S., Siggens, K.W., Archibald, A.L., Milan, D., San Cristobal, M., Laval, G., Hammond, K., Cardellino, R., Chevalet, C., 2005. An assessment of European pig diversity using molecular markers: Partitioning of diversity among breeds. Conserv. Genet. 6, 729-741. doi: 10.1007/s10592-005-9032-6

Ollivier, L., 2009. European pig genetic diversity: a minireview. Animal. 3, 915-924. doi:10.1017/S1751731109004297

Ramirez, O., Ojeda, A., Tomas, A., Gallardo, D., Huang, L.S., Folch, J.M., Clop, A., Sanchez, A., Badaoui, B., Hanotte, O., Galman-Omitogun, O., Makuza, S.M., Soto, H., Cadillo, J., Kelly, L., Cho, I.C., Yeghoyan, S., Perez-Enciso, M., Amills, M., 2009. Integrating Y-Chromosome, Mitochondrial, and Autosomal Data to Analyze the Origin of Pig Breeds. Mol. Biol. Evol. 26, 2061–2072. doi:10.1093/molbev/msp118

Ramírez, O., Burgos-Paz, W., Casas, E., Ballester, M., Bianco, E., Olalde, I., Santpere, G., Novella, V., Gut, M., Lalueza-Fox, C., Saña, M., Pérez-Enciso, M., 2015. Genome data from a sixteenth century pig illuminate modern breed relationships. Heredity (Edinb). 114, 175–184. doi:10.1038/hdy.2014.81

Ramos-Onsins, S.E., Burgos-Paz, W., Manunza, A., Amills, M., 2014. Mining the pig genome to investigate the domestication process. Heredity (Edinb). 113, 471–484. doi:10.1038/hdy.2014.68

Ramos, A.M., Crooijmans, R.P.M.A., Affara, N.A., Amaral, A.J., Archibald, A.L., Beever, J.E., Bendixen, C., Churcher, C., Clark, R., Dehais, P., Hansen, M.S., Hedegaard, J., Hu, Z.-L., Kerstens, H.H., Law, A.S., Megens, H.-J., Milan, D., Nonneman, D.J., Rohrer, G.A., Rothschild, M.F., Smith, T.P.L., Schnabel, R.D., Van Tassell, C.P., Taylor, J.F., Wiedmann, R.T., Schook, L.B., Groenen, M.A.M., 2009. Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology. PLoS One 4, e6524. doi:10.1371/journal.pone.0006524

Rubin, C.J., Megens, H.-J., Martinez Barrio, A., Maqbool, K., Sayyab, S., Schwochow, D., Wang, C., Carlborg, Ö., Jern, P., Jørgensen, C.B., Archibald, A.L., Fredholm, M., Groenen, M.A.M., Andersson, L., 2012. Strong signatures of selection in the domestic pig genome. doi:10.1073/pnas.1217149109

San Cristobal, M., Chevalet, C., Haley, C.S., Joosten, R., Rattink, A.P., Harlizius, B., Groenen, M.A., Amigues, Y., Boscher, M.Y., Russell, G., Law, A., Davoli, R., Russo, V., Désautés, C., Alderson, L., Fimland, E., Bagga, M., Delgado, J.V., Vega-Pla, J.L., Martinez, A.M., Ramos, M., Glodek, P., Meyer, J.N., Gandini, G.C., Matassino, D., Plastow, G.S., Siggens, K.W., Laval, G., Archibald, A.L., Milan, D., Hammond, K., Cardellino, R., 2006. Genetic diversity within and between European pig breeds using microsatellite markers. Anim Genet. 37, 189-98. doi: 10.1111/j.1365-2052.2005.01385.x

Sanchez, M.-P., Tribout, T., Iannuccelli, N., Bouffaud, M., Servin, B., Tenghe, A., Dehais, P., Muller, N., Del Schneider, M., Mercat, M.-J., Rogel-Gaillard, C., Milan, D., Bidanel, J.-P., Gilbert, H., 2014. A genome-wide association study of production traits in a commercial population of Large White pigs: evidence of haplotypes affecting meat quality. Genet. Sel. Evol. 46, 12. doi:10.1186/1297-9686-46-12

Schook, L.B., Beever, J.E., Rogers, J., Humphray, S., Archibald, A., Chardon, P., Milan, D., Rohrer, G., Eversole, K., 2005. Swine Genome Sequencing Consortium (SGSC): a strategic roadmap for sequencing the pig genome. Comp. Funct. Genomics 6, 251–5. doi:10.1002/cfg.479

Wang, J., Zou, H., Chen, L., Long, X., Lan, J., Liu, W., Ma, L., Wang, C., Xu, X., Ren, L., Gu, Z., Li, N., Hu, X., Zhao, Y., Zhao, Y., 2017. Convergent and divergent genetic changes in the genome of Chinese and European pigs. Sci. Rep. 7, 8662. doi:10.1038/s41598-017-09061-z

Wilkinson, S., Lu, Z.H., Megens, H.-J., Archibald, A.L., Haley, C., Jackson, I.J., Groenen, M.A.M., Crooijmans, R.P.M.A., Ogden, R., Wiener, P., 2013. Signatures of Diversifying Selection in European Pig Breeds. PLoS Genet. 9, e1003453. doi:10.1371/journal.pgen.1003453

Wolstenholme, D.R., 1992. Animal mitochondrial DNA: structure and evolution. Int. Rev. Cytol.

141, 173–216.

Yang, B., Cui, L., Perez-Enciso, M., Traspov, A., Crooijmans, R.P.M.A., Zinovieva, N., Schook, L.B., Archibald, A., Gatphayak, K., Knorr, C., Triantafyllidis, A., Alexandri, P., Semiadi, G., Hanotte, O., Dias, D., Dovč, P., Uimari, P., Iacolina, L., Scandura, M., Groenen, M.A.M., Huang, L., Megens, H.-J., 2017. Genome-wide SNP data unveils the globalization of domesticated pigs. Genet. Sel. Evol. 49, 71. doi:10.1186/s12711-017-0345-y

Zeder, M.A., Emshwiller, E., Smith, B.D., Bradley, D.G., 2006. Documenting domestication: the intersection of genetics and archaeology. Trends Genet. 22, 139–155. doi:10.1016/J.TIG.2006.01.007

Zhang, D.X., Hewitt, G.M., 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Mol. Ecol. 12, 563–84. doi:10.1046/j.1365-294X.2003.01773.x

Chapter 3

The role of genomics in pig breeding

The increasing discovery of thousands of molecular markers and the availability of genome maps have radically transformed livestock breeding by enabling the use of molecular genetics approaches in the identification of genes and polymorphisms impacting relevant phenotypes in breeding programs (Ernst and Steibel, 2013). High-throughput genotyping technologies allow to acquire a huge amount of data explaining biological diversity and phenotypes that can be exploited, especially by industries, for the improvement of breeding strategies.

Genomic selection (GS) was proposed for the first time by Meuwissen et al. (2001) and it was described as a new approach, based on the use of genetic markers covering the whole genome for the prediction of genetic merit of animals.

More specifically, genomic selection is an enhanced form of marker-assisted selection (MAS) that owes its powerful development to three main key points: a) the availability of sequenced genome information and polymorphisms detection for major livestock species, b) the possibility to cost-effectively genotype thousands of SNPs contemporarily by using SNP genotyping chips, c) the advancement in statistic methods for the evaluation of the effects of markers (Samorè and Fontanesi, 2016).

By estimating the effects of all these SNPs simultaneously without testing, genomic selection differs from marker assisted selection, that uses instead a restricted number of markers, considering the rest as having no effect (Meuwissen et al., 2016).

Genomic selection is considered complementary to standard strategies, like phenotypic parameters recording and pedigree information collection. Before genomic revolution, in fact, breeding development and genetic improvement of animals were based on the exploitation of information obtained on relatives through selection indices, but also taking advantage of estimated breeding values (EBVs) as random effects and relying on statistical methods to estimate genetic relationships (Hickey et al., 2017).

The most impacting difference between traditional methods and genomic prediction is that this latter by using genomic information allows to evaluate individuals outside from the nucleus of production (Garrick, 2017). This, results in the possibility to separate the individuals in the population eligible for candidate's selection from the limited number of those with truly recorded phenotypes.

The use of genomic selection results in a more detailed identification and explanation of genetic variance, in a higher accuracy in the estimation of the effect of QTL alleles and in a time-saving approach with the reduction of the generation interval. On the other side, the practical application of the approach itself, is considered extremely challenging because it requires the development of tools and cost-effective methods for the translation of genomic data in feasible implemented breeding programs (Goddard and Hayes, 2007; Ernst and Steibel, 2013).

In pig breeding sector, the genomic selection is oriented towards the implementation of different phenotypes and there are some works assessing the application of genomic selection with references to different selection schemes and production systems. Traits mainly investigated are production traits (like growth, average daily gain and carcass weight), meat quality traits (like intramuscular fat content), maternal performance traits (like piglets' number, mortality and survival), disease resistance (to viral or bacterial pathogens) and all those traits with low heritability or linked to measurement in late life.

Some preliminary available data derived from commercial production systems report a comparison between the new approach and the traditional one, and they show a substantial improvement in the calculation of breeding values and in the increased accuracy in selection of candidates when genomic selection is applied. This most reliable selection is estimated to results in a 35% increase in the rate of genetic gain with a faster genetic improvement estimated around 25% in pig production chain worldwide (Gjerlaug-Enger et al., 2014; Hickey et al., 2017).

The overall results from scientific literature is that results are extremely variable, and they show different results depending on traits studied and considered selection programs. For this reason, the implementation and the application of genomic selection in pig sector need to be further investigated to define costs and benefits for pig industry.

References

Ernst, C.W., Steibel, J.P., 2013. Molecular advances in QTL discovery and application in pig breeding. Trends Genet. 29, 215–224. doi:10.1016/j.tig.2013.02.002

Hickey, J.M., Chiurugwi, T., Mackay, I., Powell, W., Hickey, J.M., Chiurugwi, T., Mackay, I., Powell, W., Eggen, A., Kilian, A., Jones, C., Canales, C., Grattapaglia, D., Bassi, F., Atlin, G., Gorjanc, G., Dawson, I., Rabbi, I., Ribaut, J.-M., Rutkoski, J., Benzie, J., Lightner, J., Mwacharo, J., Parmentier, J., Robbins, K., Skot, L., Wolfe, M., Rouard, M., Clark, M., Amer, P., Gardiner, P., Hendre, P., Mrode, R., Sivasankar, S., Rasmussen, S., Groh, S., Jackson, V., Thomas, W., Beyene, Y., Beyene, Y., 2017. Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. Nat. Genet. 49, 1297–1303. doi:10.1038/ng.3920

Garrick, D.J., 2017. The role of genomics in pig improvement. Anim. Prod. Sci. 57, 2360. doi:10.1071/AN17277

Gjerlaug-Enger, E., Nordbø, Ø., Grindflek, E., 2014. Proceedings, 10 th World Congress of Genetics Applied to Livestock Production Genomic selection in pig breeding for improved meat quality.

Goddard, M.E., Hayes, B.J., 2007. Genomic selection. J. Anim. Breed. Genet. 124, 323–330. doi:10.1111/j.1439-0388.2007.00702.x

Meuwissen, T.H., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–29.

Meuwissen, T., Hayes, B., Goddard, M., 2016. Genomic selection: A paradigm shift in animal breeding. Anim. Front. 6, 6–14. doi:10.2527/af.2016-000

Samorè, A.B., Fontanesi, L., 2016. Italian Journal of Animal Science Genomic selection in pigs: state of the art and perspectives Genomic selection in pigs: state of the art and perspectives. doi:10.1080/1828051X.2016.1172034

Aim

The aim of this thesis was to describe the singularity of European local and commercial pig breeds at the genomic level to provide a first insight into the genetic architecture and population structure of morphological and economically relevant traits. In the next three chapters of this thesis we investigated candidate gene markers associated to disease resistance, coat colour and vertebral number and genes potentially involved in feeding preferences. Research activities took advantage from the availability of the latest version of the *Sus scrofa* genome and applied different genomic approaches combined with specifically designed experimental works, including next generation sequencing, bioinformatic analyses, targeted SNP genotyping, association analyses, population structure analyses coupled with traditional and novel phenotyping strategies. The study also investigated several wild boar populations in a comparative analysis with local pig breeds. The obtained result opened new opportunities to design breeding plans in commercial populations and define new management strategies for autochthonous pig breeds.

Chapter 4

Genomics and disease resistance in pigs

Infectious diseases have deep economic impacts on the pig breeding industry worldwide. It is estimated that the total cost of diseases corresponds to 20% of turnover in developed countries and up to 35–50% of turnover in the developing ones (Bishop and Woolliams, 2014). Pig production is affected by endemic, epizootic and stress-induced diseases and this is true especially for intensive production systems, where animals tend to be more susceptible to infections. A wide range of diseases have been reported in pork chain during the three phases of production (pre-weaning, growing-finishing, breeding) and, according to the responsible pathogen, they affect pigs causing disorders of digestive tract (i.e. birth diarrhea, gastroenteritis, swine dysentery), infections of the respiratory tract (i.e. influenza, pneumonia, porcine respiratory and reproductive syndrome), inflammation of specific areas (i.e. mastitis, dermatitis) (FAO, 2009).

Even if some protection measures (i.e. improved hygienic conditions of breeding stations, animal health status monitoring and pharmaceutical intervention) have been adopted in pig industry, the problem of infectious diseases still affects pig production chain. Most common control measures adopted for managing diseases are mainly antibiotics administration and vaccination programs.

The widespread use of high levels of antibiotics led to the development of antibiotic resistance with the results that some pathogens became endemic, making it more difficult to treat or eradicate efficiently diseases.

The use of vaccines, on the other hand, reduces clinical signs caused by the infection, decreases the process of shedding of pathogens in vaccinated animals with a lower transmission of infective agents, but it is documented that the majority of vaccination programs, available for livestock species, only guarantee a partial protection from infection, with the consequential demand for additional alternative methods to face the problem (Rose and Andraud, 2017).

Considering the lack of efficient solutions to control diseases in pig farming, a promising alternative can be provided using genomic information for the investigation of genetic traits conferring disease resistance in pigs. Disease resistance can be described as the ability of the host to naturally resist to a pathogen infection thanks to a genetic predisposition in facing exposure to pathogens. Best et al. (2008) defined disease resistance as the mechanism of reduction of the parasite burden, by contrasting infection and growth rate of pathogen itself.

Disease resistance traits and animal robustness are thus considered a crucial target in genomic studies: identifying genomic regions responsible for disease control and immunity response is particularly interesting for improving the selective breeding schemes.

The possibility to identify individuals more resistant or, at least, less susceptible to disease, is expected in having as a major outcome the improvement of pig farming programs that aim contemporarily to breed for resistance and improve production traits. This balance is important, because the inclusion in programs of genetic traits conferring resistance must be concordant with improved productivity, which is always the main objective for pig industry (Knap, 2005; Guy et al., 2012).

Taking advantage of genomic information, genetic maps and sequencing-derived data, some researches have been recently performed for the discovery of genetic markers and genes that can be exploited for the improvement of pig production. The result is that a few genetic markers associated to disease resistance have been identified and used in marker-assisted selection (MAS) in a few pig populations, as part of disease control programs (Meijerink et al., 1997; Coddens et al., 2008; Boddicker et al., 2012).

Different approaches can be used for the identification of molecular mechanisms connected to disease resistance and these different methods need to be merged to effectively identify involved genes and immunological pathways (Zhao et al., 2012).

The candidate gene approach is based on the identification of genes whose biological role is well known to have, directly or indirectly, a role in the etiology of the disease under investigation. So, the candidate gene is believed to contribute to a complex phenotype just considering the previous knowledge collected on the gene itself, regarding its biochemical function and phenotypes caused by mutations in the gene sequence. The identified genetic variants that disrupt protein function or are in linkage disequilibrium with functional genes, are then tested by genotyping in the population to statistically define the association between polymorphisms and examined phenotype (Tabor et al., 2002).

Another strategy used for targeting genes involved in disease resistance is the genome-wide scan approach. Genomic scans are used to systematically search for chromosomal regions, genes and specific SNPs responsible for quantitative traits (quantitative trait loci, QTL) that have economic impact (Zhao et al., 2012). With this strategy, a consistent number of QTLs has been precisely mapped and located on almost all pig chromosomes. Traditionally, research on QTLs was mainly focused on traits impacting growth, reproduction, carcass composition and meat quality, while a relative smaller

number of QTLs has been identified for disease resistance and immunity-linked traits (Rothschild et al., 2007).

The increasing attention towards the identification of QTLs associated with disease resistance is linked to the availability of pig whole-genomic screens and dense marker maps available today, that facilitate the identification of markers close to these QTLs or in linkage disequilibrium with them.

Before implementing breeding for genetic resistance, it is important to collect data regarding the genetic correlation between disease resistance and production traits, in terms of possible antagonistic effects that could make it difficult to contemporarily follow conventional selection programs and to breed for new traits.

Since these interactions or antagonism between growth, innate immunity and disease resistance traits could exist, we decided to investigate the association between four disease resistance markers already reported by other studies in a few genes with seven production traits and 15 haematological parameters in an Italian commercial pig breed.

4.1 Genetic markers associated with resistance to infectious diseases have no effects on production traits and haematological parameters in Italian Large White pigs

Introduction

Infectious diseases are the most important causes of economic losses for the pork industry worldwide. Selection and breeding for disease resistance is considered one possible approach to mitigate this problem (Mellencamp et al., 2008). Based on their relevance and potential responsiveness to genetic selection, diarrhea caused by *Escherichia coli* and Porcine Reproductive and Respiratory Syndrome (PRRS) are classified among the infectious diseases with the highest priorities in pig breeding strategies (Davies et al., 2009).

Different Enterotoxigenic *E. coli* (ETEC) strains are among the prevalent etiological agents responsible for the incidence of diarrhea in piglets during the neonatal or post-weaning periods (Chan et al., 2013). Genetic resistance to diarrhea caused by ETEC is determined by the presence or the absence of host intestinal cell surface receptors for bacterial fimbriae which are the most important factors in conferring virulence to *E. coli* strains (Schroyen et al., 2012). Using invasive *in vitro* adhesion assays, resistant pigs can be identified by a much lower number of bacteria attached to the intestinal cell surface compared to susceptible ones (Sellwood et al., 1975). ETEC strains presenting F4 fimbriae with different antigenic variant combinations have been shown to be the most prevalent

strains causing Neonatal Diarrhea (ND) in pigs (Moon et al., 1999). A single nucleotide polymorphism (SNP) in intron 7 (rs338992994, NC_010455.5:g.134226654C>G) of the mucin 4 (*MUC4*) gene (which encodes a large membrane-bound mucin) has been demonstrated to be in linkage disequilibrium with the receptor for ETEC strain exhibiting F4 fimbriae with both antigenic variants *ab* or *ac*. Allele G of this marker is associated with the presence of the receptor that confers susceptibility to ND and is presumed to be dominant over the resistance allele C, associated with the absence of the receptor (Jorgensen et al., 2004). This marker, not completely concordant with the adhesion assay results (Rasschaert et al., 2007), still remains the most reliable polymorphism used in marker assisted selection (MAS) programs to reduce ND incidence in piglets (Goetstouwers et al., 2014).

Post-Weaning Diarrhea (PWD) is determined by *E. coli* strains exposing F18 fimbriae, that seem more prevalent than ETEC F4 (Moon et al., 1999). A nucleotide substitution in the alpha (1,2)-fucosyltransferase 1 (*FUT1*) gene (rs335979375, NC_010448.4:g.54079560A>G) has been found to be in high linkage disequilibrium with the gene encoding for the receptor for ETEC strain exhibiting F18 fimbriae with *ab* antigenic variant (ETEC F18ab). The recessive A allele is associated with the absence of ETEC F18ab receptor, that in turn, determines resistance to PWD (Meijerink et al., 1997). *FUT1* gene encodes an enzyme that catalyzes an intermediary step in the formation of antigens pertaining to the porcine AO blood group system. The indicated polymorphism in this gene is considered thus the most reliable discovered marker for the identification of PWD resistant pigs and for this reason has been included in MAS programs (Meijerink et al., 1997; Coddens et al., 2007, 2008).

PRRS is responsible for gross production losses due to by late-term reproductive failure in sows and respiratory illness in growing pigs. The etiological agent is an Arteriviridae positive-sense RNA virus with two main different strains showing only about 60% of sequence identity (Kappes and Faaberg, 2015; Neumann et al., 2005). Genetic variation in the host response to this disease was first observed in the late 90s (Halbur et al., 1998) and was subsequently investigated by a series of challenging studies conducted by the PRRSV Host Genetic Consortium (PHGC) (Lunney et al., 2011) leading to the discovery of an important QTL on porcine chromosome (SSC) 4 responsible for 15% of the genetic variation of the host's immune response to this infection (Boddicker et al., 2012). This QTL region includes five genes encoding guanylate binding protein (GBP) family members, that are mediators of the proinflammatory immune response post viral or bacterial infections (Koltes et al., 2015, Pilla et al., 2014). Two SNPs have been found to be in linkage disequilibrium with this SSC4 QTL. One SNP (rs340943904, NC_010446.5:g.127301202G>T), located in intron 9-10 of the *GBP1* gene (encoding a negative regulator of T-cell responses), has been identified as the putative causative

mutation responsible of this QTL (Koltes et al., 2015). The presence at this locus of nucleotide G introduces a splice acceptor site that leads to the addition of 5 nucleotides into intron 9-10. This causes, in turn, the shift of the reading frame and the formation of a premature stop codon in *GBP1* sequence with the transcription of non-functional products. However, only pigs with GG genotype produce entirely non-functional transcripts and are regarded as susceptible, suggesting a dominant effect of the resistance allele T (Koltes et al., 2015). The second SNP, WUR10000125 (rs80800372, NC_010446.5:g.127441677A>G), is a polymorphism located next to a putative polyadenylation site in the 3'-untranslated region of the *GBP1* gene which can potentially affect transcript stability, with consequences on protein synthesis and expression (Gol et al., 2015). A few studies have been carried out in different pig populations to investigate the role of WUR10000125 in PRRSV resistance and its possible impact on performance traits (Boddicker et al., 2014; Niu et al., 2015, 2016; Abella et al., 2016; Dulkenberger et al., 2017a; Waide et al., 2017). This polymorphism is currently included in commercial porcine SNP genotyping chips. Pigs with unfavorable A allele at this marker showed higher viremia and lower weight gain following infection whereas pigs carrying the G variant, which might be the dominant allele, reported a lower viremia and showed higher growth rate under PRRSV infection (Boddicker et al., 2012). The same allele has been also associated with lower porcine circovirus type 2b (PCV2b) virus load and might be useful to improve host response on coinfection of PRRSV with PCV2b (Dunkelberger et al., 2017b).

Since previous studies have suggested a negative correlation between growth traits and disease resistance and a few of these markers have been suggested to have antagonistic pleiotropic effects on these economically relevant aspects (Doeschl-Wilson et al., 2009; Fontanesi et al., 2012; Huang et al., 2008), investigating the effects of these markers on growth, carcass and meat quality and production traits is essential for their proper use in MAS programs.

In this study, we genotyped *MUC4*, *FUT1* and *GBP1* gene markers i) to compare allele frequencies distribution in Italian local pig breeds and Large White pigs, ii) to retrospectively investigate allele frequency changes of these markers over two decades of selection among Italian Large White boars and iii) to evaluate the effect of these polymorphisms on performance, carcass and meat quality traits and haematological parameters (some of which related to the immune response) in Italian Large White pigs.

Materials and methods

Animals and traits

All animals used in this study were not raised for any experimental purposes. They were kept according to Italian and European legislation for pig production and all procedures described were in compliance with national and European Union regulations for animal care and slaughtering.
Three groups of pigs were included in this study. One group was made by a total of 117 unrelated animals of four Italian local pig breeds (Apulo-Calabrese, Casertana, Cinta Senese and Nero Siciliano) that were used for allele frequency analysis and comparison across breeds.
Two groups of Italian Large White pigs were analyzed in this study. The first group consisted of 189 Italian Large White boars born and raised from 1992 to 2012 under the national heavy pig selection scheme of this breed (Fontanesi et al., 2015; Schiavo et al., 2016). These boars were used for the evaluation of allele frequency changes over years. The animals were selected among all Italian Large White boars born from 1992 to 2012 (n. = 5983) and approved for reproduction based on their genetic merit after evaluation by the National Pig Breeders Association (ANAS). The boars were ranked according to the reliability of their estimated breeding values (EBVs; calculated in 2012) and those with the highest reliability (n. = 189) were selected and used to constitute eight groups (n.19-28 boars per group) based on the years in which they were born (groups were constituted using 2-4 years windows; 2011-2012 did not have boars with high EBV reliability, thus this window was not included in the two-decades timeframe). More information about these boars are reported in Fontanesi et al. (2015).
The second group of Italian Large White pigs consisted of 557 performance tested animals (382 gilts and 175 castrated males). These animals were used in the association studies between gene markers and phenotypic traits (see below). These pigs were part of the sib-testing program of the Italian Large White population for which triplets of pigs from the same litter (2 females and 1 castrated male) were individually performance tested at the ANAS Central Test Station for the genetic evaluation of a boar from the same litter (sib-testing). The test period started when piglets were 30-45 days old and ended when they reached 155 ± 5 kg live weight at approximately nine months of age. The nutritive level was *quasi ad libitum*, meaning that about 60% of the pigs were able to ingest the entire supplied ratio (Fontanesi et al., 2008, 2012). During the test period feed intake was recorded daily, body weight was measured bimonthly, and then average daily gain (ADG, in g) and feed:gain ratio (FGR) were calculated. At the end of the testing period, animals were slaughtered in a commercial abattoir after electrical stunning and following standard procedures. These pigs were slaughtered in 17 different days in the years 2012-2013. After slaughtering, weight of lean cuts (LC, the weight of neck and loin in kg), weight of the hams (HW, in kg) and backfat thickness (BFT, determined at the level of

Musculus gluteus medius and expressed in mm) were measured on the carcasses. Visible intermuscular fat (VIF) was scored on the exposed muscles of the legs as previously described (Fontanesi et al. 2017b; Bertolini et al., 2018). Ham weight loss at first salting (HWLFS, in g) was calculated during the first week of seasoning (Fontanesi et al., 2017a; Bertolini et al., 2018). Haematological parameters were determined on blood samples, collected just after jugulation and exsanguination, in EDTA added tubes (Vacutest Kima s.r.l.). A total of 15 haematological parameters (erythrocyte traits: red blood cell count, haemoglobin, haematocrit, mean corpuscular volume, mean corpuscular haemoglobin, mean corpuscular haemoglobin concentration and red cell distribution width; leukocyte traits: white blood cell count, lymphocyte count, neutrophil count, eosinophil count, basophil count and monocyte count; platelet traits: platelet count and mean platelet volume) were measured on an Olympus AU 400 (Beckman Coulter) automated analyzer at the Veterinary Haematological Laboratory of the University of Bologna under standard procedures.

SNP genotyping

Genomic DNA was extracted from blood using the NucleoSpin®Tissue commercial kit (Macherey-Nagel, Düren, Germany). Genotyping was obtained for four SNPs in three genes already reported to be associated with resistance to porcine bacterial or viral determined diseases: rs338992994 (g.134226654C>G) in the *MUC4* gene; rs335979375 (g.54079560A>G) in the *FUT1* gene; rs340943904 (g.127301202G>T) and rs80800372 (WUR10000125 or g.127441677A>G) in the *GBP1* gene. PCR primers and genotyping protocols are reported in Supplementary Table 1. PCR-RFLP was used for the genotyping of three SNPs (rs338992994, rs335979375 and rs340943904) whereas results for the WUR10000125 marker were obtained from the genotyping of the PorcineSNP60 BeadChip (Illumina Co., St. Diego, CA, USA) SNP chip, version 1 (as reported by Fontanesi et al. 2017a,b). PCR was carried on a 2720 Thermal Cycler (Applied Biosystems) in a total volume of 20 µL that included 10 pmol of each primer, 2.5 mM each dNTP, 2 µL of KAPA Taq Buffer 10x (containing 1.5 mM of $MgCl_2$ at 1X), 0.5 U of KAPA Taq DNA Polymerase (KAPA Biosystems, Boston, MA, USA). Digested PCR products were visualized by electrophoresis on 2.5%-3.0% agarose gels in TBE (Tris-Borate-EDTA) 1X with the addition of GelRed Nucleic Acid Gel Stain (Biotium Inc., Hayward, CA, USA).

Statistical analyses

Allele and genotype frequencies were obtained by counting. Hardy Weinberg equilibrium was evaluated using the HWE software program (Linkage Utility Programs, Rockefeller University, New York, NY). For each breed, PLINK software v1.07 (Purcell et al., 2007) was used to calculate the

linkage disequilibrium (LD, here reported by using the r^2 measure) and PHASE software v2.1.1 (Stephens et al., 2001) was used to identify haplotypes between the two SNPs (rs340943904 and WUR10000125) genotyped in the *GBP1* gene (1000 iterations were set as parameter for computing haplotypes).

Allele frequency changes over years of the genotyped SNPs in the first group of Italian Large White pigs were evaluated using logistic regression models. To calculate the P value associated with the change in allele frequency, the 'glm' function in R software (R Development Core Team, 2008) was used by specifying the binomial family parameter with the model that included the time of each group. Time as a covariate was expressed as an integer ranging from 1 to 8 (groups 1 to 8, depending on the years, as shown in Fig.1). Each animal was codified with a vector in which each marker was, in turn, codified with 0, 1 or 2, depending on the number of copies of the minor allele.

The second group of Italian Large White performance tested pigs was used in the association analysis between the recorded phenotypic traits and the genotyped SNPs. Association with production traits was obtained using Random Residuals (RRs) as defined by Fontanesi et al. (2010) on traits (ADG, BFT, LC, FGR, HW, VIF and HWLFS) measured by ANAS, as described above. As previously reported, the use of RRs reduces the type I error and the overestimation of the effects that would be obtained by using EBVs (Fontanesi et al., 2013). Haematological parameters were normalized in R (packages "MASS" and "CAR") by applying the Box–Cox transformation as previously described by Bovo et al. (2016).

Association analysis was carried out in GEMMA (Zhou and Stephens, 2012) by fitting linear mixed models that included a pedigree relationship matrix ***K***. The assessment of the association between each SNP and trait was obtained by testing the null hypothesis $H_0:\beta = 0$. Additive genetic model, assuming a trend per copy of the minor allele, was used to specify the dependency of trait on genotype categories. The model for the (normalized) haematological traits included as covariates sex, date of slaughtering and carcass weight. Because of the estimation of RR for each trait already corrects for possible variables, these covariates were not added while fitting the linear model. *GBP1* haplotypes were not used in the association analysis as recombinant haplotypes were carried by few animals, preventing a correct estimation of their effect. SNPs were defined to be significantly associated if their P nominal values were below 5.68×10^{-04} (Bonferroni corrected $P < 0.05$). This threshold was determined by multiplying the four SNPs by the 22 traits under investigation (seven production traits and 15 haematological traits). Suggestively significant results were considered if *P* nominal values were below 0.05.

Results

Allele and genotype frequencies for the investigated SNPs are reported in Table 1. Only the *FUT1* polymorphism in the Italian Large White population (second group of pigs) significantly deviated (P < 0.05) from HWE (Supplementary Table 2). All other investigated polymorphisms in the analyzed breeds were in HWE (Supplementary Table 2).

Frequency of the *MUC4* C allele (conferring resistance to ND) was 0.57 in the Italian Large White population, the lowest frequency in the analyzed breeds. This allele was always the major allele in all local breeds in which its frequency ranged from 0.72 (Nero Siciliano) to 1.00 (Cinta Senese).

Allele A of the *FUT1* marker (the allele associated with resistance to PWD) was the minor allele in all breeds. Its frequency in Italian Large White pigs was similar to that of most local breeds apart Casertana which showed the highest frequency of this allele among all investigated breeds (0.554).

The frequency of allele T of the *GBP1* rs340943904 marker (the dominant allele associated with resistance to PRRS) was low (0.087) in the Italian Large White population. Its frequency was higher in the Italian local breeds, except in Apulo-Calabrese where it was 0.125. Cinta Senese was again the breed that showed the highest frequency of the resistance allele (0.437).

Allele G of WUR10000125 (*GBP1* rs80800372) SNP (associated with resistance to PRRS) was low in all breeds and ranged from 0.1 (Casertana) to 0.168 (Cinta Senese). It was 0.07 in the commercial Italian Large White breed.

The two analyzed *GBP1* SNPs were in complete LD only in Apulo-Calabrese pigs (r^2 = 1.00). The value of r^2 was 0.133 in Cinta Senese, 0.698 in Nero Siciliano and 0.722 in Casertana. In the Italian Large White breed r^2 was 0.696. Table 2 reports the *GBP1* haplotype frequencies observed in the analyzed breeds. The haplotype derived by the combination of the two susceptibility alleles (i.e. GA, respectively for rs340943904 and WUR10000125) was, on average, the most frequent overall investigated pigs (0.80). The haplotype derived by the two resistance associate alleles (i.e. TG at the same positions) showed the lowest frequency in the Italian Large White population.

Figure 1 shows the frequencies of the resistance associated alleles for the four SNPs under investigation in this study in the Italian Large White boars over two decades. The plot describes a retrospective analysis of frequency changes in this timeframe in the whole Italian Large White population (Fontanesi et al., 2015). This overview might give a general impression of slow reduction of the frequency of the resistance alleles for all analysed markers (Fig. 1). However, the logistic regression model indicated that only for *FUT1* SNP there was a significant change of allele frequencies over the investigated 20 years (P = 0.027) with a gradual reduction of the resistance associated allele (0.25 in 1996-1997 and 0.06 in 2008-2010). *MUC4*, *FUT1* and *GBP1* rs80800372 and rs340943904 allele frequency distribution did not significantly change in the considered

timeframe (Table 3). For these investigated markers, there was a peak for the resistance alleles in the 1996-1997 years that was not subsequently confirmed by the trends of the following years (Fig. 1). Overall allele frequencies of all four investigated SNPs in the 2008-2010 window was a little bit lower than the frequency observed in the second groups of Italian Large White pigs used in the association studies, which were born in the years 2011-2012 (as they were slaughtered in 2012-2013; Supplementary Table 3). This result may indicate, indirectly, that allele frequencies in the sib-tested population tend to fluctuate, a few years later, around values already observed in the boar population in the considered timeframe.
Results of association analysis between the analyzed SNPs and ADG, BFT, FGR, LC, HW, VIF and HWLFS are reported in Table 4. None of the four SNPs was associated with any production traits both considering a Bonferroni corrected threshold (significant association) and a P nominal value of 0.05 (suggestive association).
Association study for the haematological traits showed only four suggestive associations (P nominal value < 0.05; Supplementary Table 4). *MUC4* was suggestively associated with mean platelet volume and lymphocyte count, *FUT1* was suggestively associated with red blood cell count and *GBP1* rs80800372 was suggestively associated with monocyte count.

Discussion

A few gene markers have been already shown to be associated with resistance to some of the most important diseases that have a negative economic impact on the pig production chain worldwide (Vögeli et al., 1997; Huang et al., 2008; Boddicker et al., 2012, Schroyen et al., 2016). Based on these works that prompted the use of polymorphisms in the *MUC4*, *FUT1* and *GBP1* genes in MAS to increase disease resistance in several pig populations, it is important to first evaluate allele frequency distribution in other populations that were not part of these investigations and to verify if these markers are associated with other production traits. This is essential to evaluate if antagonist effects of these genes might reduce the genetic progress on economic traits that are currently used in the breeding program of commercial populations.
This study investigated the allele frequency distribution of *MUC4*, *FUT1* and *GBP1* polymorphisms in an important heavy pig breed (Italian Large White) and in several Italian local breeds (Apulo-Calabrese, Casertana, Cinta Senese and Nero Siciliano) which complemented the results on the commercial breed. The Italian Large White breed was also investigated to evaluate if allele frequencies of the analyzed markers changed over a period of 20 years of selection for animals with carcass and meat quality traits more adapted to the dry-cured ham production chain (Bosi and Russo, 2004). None of the four investigated SNPs were associated with any traits included in the national

selection program of the Italian Large White breed. In addition, none of the haematological traits (some of which might be important to define indirectly the potential immune response of the animals) was significantly associated with the same markers, suggesting that other biological mechanisms might be involved in explaining the reported effect on disease resistance.

The results of the association studies with production traits are in line with the general stable allele frequencies of these markers over two decades of selection in the Italian Large White breed that significantly increased average daily gain and improved all meat and carcass traits needed for the production of dry-cured hams (Fontanesi et al., 2015). Had the analyzed gene markers strongly affected the traits already used in the selection programs of this breed, a strong allele frequency modification over time should have been expected. This was already found for other gene markers already shown to be associated with important production traits (i.e. *IGF2*, *FTO, MC4R* and *VRTN*; Fontanesi et al. 2015). Only the *GBP1* rs340943904 showed a significant allele frequency change in the investigated timeframe, probably due to genetic drift or a small effect (not detected in the association analysis) on traits under direct or indirect selection in this population. In this context, despite the results were not statistically significant, it is worth to note that there was a general, even small, decrease, of allele frequency of the resistance associate alleles at all loci. Considering that there would be no detrimental effect on any production traits (according to the results obtained in the association study we carried out), it should be important to re-consider the selection program of this breed to reverse this trend by adopting a targeted MAS program that aims to increase the frequency of the favorable alleles (associated to resistance to ND, PWD and PRRS) at these markers and, indirectly, increase disease resistance in this population. This could be important considering that the favorable allele at these markers was in almost all cases the minor allele in the Italian Large White population (except for the *MUC4* SNP). In contrast, in most local breeds, the favorable allele was usually the most frequent one or showed a higher frequency than in the Italian Large White population. This aspect might confirm the higher rusticity (partly derived by a lower susceptibility to infection diseases) that is usually attributed to local pig breeds compared to commercial populations. In these autochthonous breeds, the effect of natural selection, and thus adaptation to harsh environmental conditions, could have increased the frequency of disease resistance alleles.

In a previous study that was based on extremely divergent Italian Large White pigs for estimated breeding values for a few production traits, the same *MUC*4 SNP showed a significant association between ADG and the resistance allele that was more frequent in slow-growing pigs (Fontanesi et al., 2012). The design of that experiment might have captured more subtle effects that were not evident in the current study which is based on animals coming from a more homogeneous sib-tested population (born in the years 2011-2012). The general direction of the effect seems however, even

not significant, the same as that previously observed (Fontanesi et al., 2012), i.e. genotype CC grew slower than the other genotypes (data not shown). In another study, Yan et al. (2009) reported faster growth rates during the fattening period for animals presenting ETEC F4ab and F4ac receptors. Yan et al. (2009) also indicated that susceptible pigs with F4abR demonstrated higher values of ADG during the suckling period. These results might confirm, to some extent, a small or population dependent effect of this marker on production traits.

The *FUT1* marker that, in this study, did not affect any traits in the Italian Large White population was, instead, associated with growth rate (with a positive effect of allele A) in a Swiss pig population (Kadarmideen, 2008). Similar results were also obtained by Bao et al. (2012) in a study on Sutai pigs (hybrids of Duroc and Meishan pigs) where animals with the AA genotype had the highest ADG value before weaning.

Allele frequencies of the WUR10000125 (*GBP1* rs80800372) polymorphism obtained in this study were similar to those reported by other works which showed that the unfavorable A allele is the most frequent one in commercial lines selected for increased growth rate (ranging from 0.7 to 0.9; Boddicker et al., 2012; Abella et al., 2016). In a recent study on nonchallenged animals, Dunkelberger et al. (2017a) estimated the effect of WUR10000125 on traits under selection in four different purebred pig lines. Even if the effect of this marker acted differently among commercial lines, nonsignificant effect was reported for most of the traits and for overall selection index value, suggesting that selecting for the resistance allele should result in pig resistant to PRRS without consequences on productive traits in nonchallenging conditions.

The linkage between the two *GBP1* markers associated to PRRS resistance (i.e. rs340943904 and rs80800372) was previously evaluated only by Koltes et al. (2015) in a commercial crossbred population who reported that the two markers were in perfect LD. Our results contrast with what was reported by Koltes et al. (2015). LD varied among breeds with extreme values in local pig populations (r^2 = 1.00 in Apulo-Calabrese and r^2 = 0.133 in Cinta Senese). Therefore, more than two haplotypes were identified in Italian Large White, Casertana, Cinta Senese and Nero Siciliano breeds suggesting that previous association studies based on WUR10000125 should be revised including the genotyping of the other associated *GBP1* SNP in populations in which these two markers are not in complete LD.

Conclusions

Disease resistance to infective agents is considered a complex trait with low heritability. However, a few gene markers have been shown to explain a relevant fraction of the genetic variability of important diseases. This study showed that implementing a MAS program in the Italian Large White pig breed, aimed to indirectly increase disease resistance, might not have any relevant adverse effect

on all other performance, carcass and meat quality traits included in the genetic merit index used in the national selection scheme. This conclusion is based on the absence of negative effects on the investigated traits for the alleles in the *MUC4*, *FUT1* and *GBP1* already shown by other studies to be associated to resistance to bacterial and viral diseases. In our case, before implementing MAS based on these genes, further studies should be carried out to investigate their effect on reproduction traits in this heavy pig breeds that is usually used as maternal line in the heavy pig breeding programs. Other studies are also needed to evaluate the level of LD between *GBP1* SNPs and the effect of their haplotypes on production traits for the subsequent use of these markers in MAS to reduce susceptibility to PRRS.

Remarks

The information reported in this chapter were submitted as manuscript to Livestock Science Journal in October 2018. The submitted material is currently under revision.

Acknowledgements

Different partners contributed to the realization of the study, therefore we would like to thank the Italian Association of Pig Breeders (Associazione Nazionale Allevatori Suini – ANAS), the Department of Statistical Sciences of the University of Bologna, the Department of Animal Genetic Improvement at INIA, Madrid, and the Italian ‘Consiglio per la Ricerca in agricoltura e l’analisi dell’Economia Agraria’(CREA).

Table 1: Allele and genotype frequencies of the analyzed polymorphisms in the Italian Large White population and in the autochthonous pig breeds.

	***MUC4*[1]** rs338992994						**_FUT1_** rs335979375						**_GBP1_** rs340943904						**_GBP1_** _rs80800372_					
Breeds	No. of Pigs	Allele Frequency		Genotype Frequency			No. of pigs	Allele Frequency		Genotype Frequency			No. of pigs	Allele Frequency		Genotype Frequency			No. of pigs	Allele Frequency		Genotype Frequency		
		C	G	CC	CG	GG		A	G	AA	AG	GG		T	G	TT	TG	GG		G	A	GG	AG	AA
Italian Large White[2]	500	0.575	0.425	0.328	0.494	0.178	510	0.115	0.885	0.026	0.178	0.796	557	0.087	0.913	0.013	0.149	0.838	557	0.07	0.93	0.009	0.12	0.871
Apulo Calabrese[3]	15	0.83	0.17	0	0.333	0.667	25	0.06	0.94	0	0.12	0.88	28	0.125	0.875	0	0.25	0.75	28	0.125	0.875	0	0.25	0.75
Casertana[3]	27	0.91	0.09	0.815	0.185	0	28	0.554	0.446	0.393	0.321	0.286	30	0.134	0.866	0	0.267	0.733	30	0.1	0.9	0	0.2	0.8
Cinta Senese[3]	22	1.000	0	1.000	0	0	29	0.069	0.931	0	0.862	0.138	16	0.437	0.563	0.188	0.5	0.312	16	0.168	0.832	0	0.188	0.812
Nero Siciliano[3]	30	0.72	0.28	0.534	0.366	0.1	23	0.108	0.892	0	0.217	0.783	26	0.23	0.77	0.077	0.308	0.615	26	0.16	0.84	0.077	0.192	0.731

[1]Frequencies reported from Fontanesi et al. (2012) except for Italian Large White pigs, which are from this study.
[2]Total individuals were 557 but for *MUC4* and *FUT1* genotyping was missing for some individuals. Allele and genotype frequencies were reported here only for the second group of Italian Large White pigs, which included performance tested pigs born in the years 2011-2012.
[3]Genotyping results in the local breeds were obtained for 15-30 pigs each.

Table 2: Haplotype frequencies in the Italian Large White and Italian local pig breeds for the two markers in the *GBP1* gene.

Breed[1]	***GBP1* haplotypes**[2]	**Haplotype frequency**
Italian Large White (557)	TG	0.066
	TA	0.024
	GG	0.005
	GA	0.905
Apulo Calabrese (28)	TG	0.125
	GA	0.875
Casertana (30)	TG	0.100
	TA	0.033
	GA	0.867
Cinta Senese (16)	TG	0.094
	TA	0.343
	GA	0.563
Nero Siciliano (26)	TG	0.173
	TA	0.058
	GA	0.769

[1] The number of pigs is included in parenthesis.
[2] The first and the second nucleotides are for the rs340943904 and rs80800372 SNPs, respectively.

Table 3: Effect of time on resistance allele frequency changes in candidate boars over 20 years.

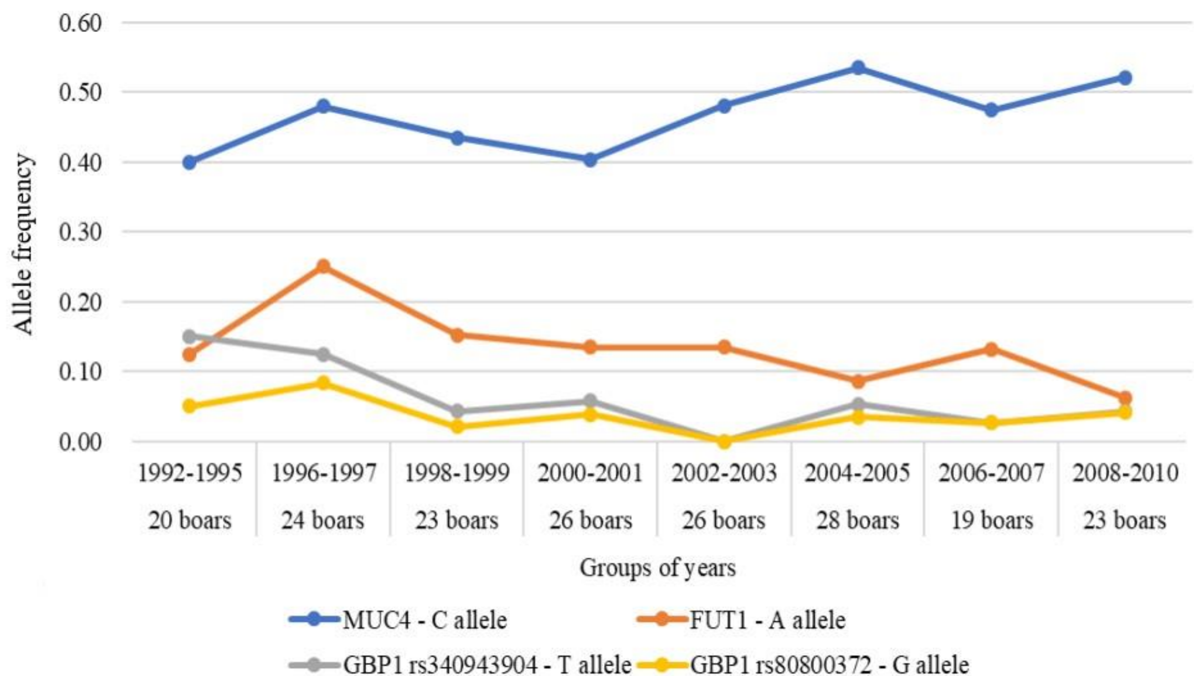
Gene/SNPs	**Allele**	**Chi square value**	**P value**
MUC4 - rs338992994	C	0.600	0.423
FUT1 - rs335979375	A	5.000	0.027
GBP1 - rs340943904	T	1.260	0.263
GBP1 - rs80800372	G	0.650	0.422

Table 4: Results of the association study between the analyzed markers and production, carcass and meat quality traits in the second group of Italian Large White pigs. The table reports the calculated P nominal value.

Traits[1]	*MUC4* rs338992994	*FUT1* rs335979375	*GBP1* rs340943904	*GBP1* rs80800372
ADG	0.420	0.337	0.291	0.101
BFT	0.068	0.730	0.118	0.236
LC	0.133	0.534	0.768	0.633
FGR	0.577	0.308	0.738	0.605
HW	0.910	0.878	0.639	0.976
VIF	0.678	0.143	0.731	0.953
HWLFS	0.388	0.491	0.160	0.261

[1]ADG: average daily gain, BFT: back fat thickness; LC: lean meat cuts; FGR: feed gain ratio; HW: ham weight; VIF: visible intermuscular fat; HWLFS: ham weight loss at first salting.

Figure 1: Allele frequency trends of four single nucleotide polymorphisms associated to disease resistance in Italian Large White boars under the national selection program over about two decades. *GBP1* haplotype frequency changes were not evaluated due to the low frequencies of three haplotypes out of four reported in this population (see Table 2 for the haplotype frequencies in this breed).



References

Abella, G., Pena, R.N., Nogareda, C., Armengol, R., Vidal, A., Moradell, L., Tarancon, V., Novell, E., Estany, J., Fraile, L., 2016. A WUR SNP is associated with European Porcine Reproductive and Respiratory Virus Syndrome resistance and growth performance in pigs. Res. Vet. Sci. 104, 117–122. doi:10.1016/j.rvsc.2015.12.014

Bao, W. B., Ye, L., Zhu, J., Pan, Z.Y., Zhu, G.Q., Huang, X.G., Wu, S.L., 2012. Evaluation of M307 of FUT1 gene as a genetic marker for disease resistance breeding of Sutai pigs. Mol. Biol. Rep. 39, 4223–4228. doi:10.1007/s11033-011-1208-1

Bertolini, F., Schiavo, G., Galimberti, G., Bovo, S., D'Andrea, M., Gallo, M., Buttazzoni, L., Rothschild, M.F., Fontanesi, L., 2018. Genome-wide association studies for seven production traits highlight genomic regions useful to dissect dry-cured ham quality and production traits in Duroc heavy pigs. Animal 1–8. doi:10.1017/S1751731118000757

Best, A., White, A., Boots, M., 2008. Maintenance of host variation in tolerance to pathogens and parasites. Proc. Natl. Acad. Sci. U. S. A. 105, 20786–91. doi:10.1073/pnas.0809558105

Bishop, S.C., Woolliams, J.A., 2014. Genomics and disease resistance studies in livestock. Livest. Sci. 166, 190–198. doi:10.1016/j.livsci.2014.04.034

Boddicker, N., Waide, E.H., Rowland, R.R.R., Lunney, J.K., Garrick, D.J., Reecy, J.M., Dekkers, J.C.M., 2012. Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. J. Anim. Sci. 90, 1733–1746. doi:10.2527/jas2011-4464

Boddicker, N.J., Garrick, D.J., Rowland, R.R.R., Lunney, J.K., Reecy, J.M., Dekkers, J.C.M., 2014. Validation and further characterization of a major quantitative trait locus associated with host response to experimental infection with porcine reproductive and respiratory syndrome virus. Anim. Genet. 45, 48–58. doi:10.1111/age.12079

Bosi, P., Russo, V., 2004. The production of the heavy pig for high quality processed products. Ital. J. Anim. Sci. 3, 309–321. doi:10.4081/ijas.2004.309

Bovo, S., Schiavo, G., Mazzoni, G., Dall'Olio, S., Galimberti, G., Calò, D.G., Scotti, E., Bertolini, F., Buttazzoni, L., Samorè, A.B., Fontanesi, L., 2016. Genome-wide association study for the level of serum electrolytes in Italian Large White pigs. Anim. Genet. 47, 597–602. doi:10.1111/age.12459

Chan, G., Farzan, A., DeLay, J., McEwen, B., Prescott, J.F., Friendship, R.M., 2013. A retrospective study on the etiological diagnoses of diarrhea in neonatal piglets in Ontario, Canada, between 2001 and 2010. Can. J. Vet. Res. 77, 254–60.

Coddens, A., Verdonck, F., Tiels, P., Rasschaert, K., Goddeeris, B.M., Cox, E., 2007. The age-dependent expression of the F18+ E. coli receptor on porcine gut epithelial cells is positively correlated with the presence of histo-blood group antigens. Vet. Microbiol. 122, 332–341. doi:10.1016/j.vetmic.2007.02.007

Coddens, A., Verdonck, F., Mulinge, M., Goyvaerts, E., Miry, C., Goddeeris, B., Duchateau, L., Cox, E., 2008. The possibility of positive selection for both F18+ Escherichia coli and stress resistant pigs opens new perspectives for pig breeding. Vet. Microbiol. 126, 210–215. doi:10.1016/j.vetmic.2007.06.021

Davies, G., Genini, S., Bishop, S.C., Giuffra, E., 2009. An assessment of opportunities to dissect host genetic variation in resistance to infectious diseases in livestock. Anim. Anim. Consort. 3, 415–436. doi:10.1017/S1751731108003522

Doeschl-Wilson, A.B., Kyriazakis, I., Vincent, A., Rothschild, M.F., Thacker, E., Galina-Pantoja, L., 2009. Clinical and pathological responses of pigs from two genetically diverse commercial lines to porcine reproductive and respiratory syndrome virus infection. J. Anim. Sci. 87, 1638–1647. doi:10.2527/jas.2008-1447

Dunkelberger, J.R., Mathur, P.K., Lopes, M.S., Knol, E.F., Dekkers, J.C.M., 2017a. A major gene for host response to porcine reproductive and respiratory syndrome is not unfavorably associated with overall performance under nonchallenging conditions in commercial pig lines. J. Anim. Sci. 95, 2838–2847. doi:10.2527/jas.2017.1524

Dunkelberger, J.R., Serão, N.V.L., Niederwerder, M.C., Kerrigan, M.A., Lunney, J.K., Rowland, R.R.R., Dekkers, J.C.M., 2017b. Effect of a major quantitative trait locus for porcine reproductive and respiratory syndrome (PRRS) resistance on response to coinfection with PRRS virus and porcine circovirus type 2b (PCV2b) in commercial pigs, with or without prior vaccination for PRRS1. J. Anim. Sci. 95, 584–598. doi:10.2527/jas.2016.1071

FAO, 2009. Farmer's Hand Book on Pig Production (For the small holders at village level) Food and Agriculture Organization of the United Nations European Comission.

Fontanesi, L., Davoli, R., Nanni Costa, L., Beretti, F., Scotti, E., Tazzoli, M., Tassone, F., Colombo, M., Buttazzoni, L., Russo, V., 2008. Investigation of candidate genes for glycolytic potential of porcine skeletal muscle: Association with meat quality and production traits in Italian Large White pigs. Meat Sci. 80, 780–787. doi:10.1016/j.meatsci.2008.03.022

Fontanesi, L., Speroni, C., Buttazzoni, L., Scotti, E., Dall'Olio, S., Nanni Costa, L., Davoli, R., Russo, V., 2010. The insulin-like growth factor 2 (IGF2) gene intron3-g.3072G>A polymorphism is not the only Sus scrofa chromosome 2p mutation affecting meat production and carcass traits in pigs: Evidence from the effects of a cathepsin D (CTSD) gene polymorphism. J. Anim. Sci. 88,

2235–2245. doi:10.2527/jas.2009-2560

Fontanesi, L., Bertolini, F., Dall'Olio, S., Buttazzoni, L., Gallo, M., Russo, V., 2012. Analysis of association between the MUC4 g.8227C>G polymorphism and production traits in Italian heavy pigs using a selective genotyping approach. Anim. Biotechnol. 23, 147–155. doi:10.1080/10495398.2011.653462

Fontanesi, L., Buttazzoni, L., Galimberti, G., Calò, D.G., Scotti, E., Russo, V., 2013. Association between melanocortin 4 receptor (MC4R) gene haplotypes and carcass and production traits in Italian Large White pigs evaluated with a selective genotyping approach. Livest. Sci. 157, 48–56. doi:10.1016/j.livsci.2013.07.006

Fontanesi, L., Schiavo, G., Scotti, E., Galimberti, G., Calò, D.G., Samorè, A.B., Gallo, M., Russo, V., Buttazzoni, L., 2015. A retrospective analysis of allele frequency changes of major genes during 20 years of selection in the Italian Large White pig breed. J. Anim. Breed. Genet. 132, 239–246. doi:10.1111/jbg.12127

Fontanesi, L., Schiavo, G., Gallo, M., Baiocco, C., Galimberti, G., Bovo, S., Russo, V., Buttazzoni, L., 2017a. Genome-wide association study for ham weight loss at first salting in Italian Large White pigs: towards the genetic dissection of a key trait for dry-cured ham production. Anim. Genet. 48, 103–107. doi:10.1111/age.12491

Fontanesi, L., Schiavo, G., Galimberti, G., Bovo, S., Russo, V., Gallo, M., Buttazzoni, L., 2017b. A genome-wide association study for a proxy of intermuscular fat level in the Italian Large White breed identifies genomic regions affecting an important quality parameter for dry-cured hams. Anim. Genet. 48, 459–465. doi:10.1111/age.12542

Goetstouwers, T., Van Poucke, M., Coppieters, W., Nguyen, V.U., Melkebeek, V., Coddens, A., Van Steendam, K., Deforce, D., Cox, E., Peelman, L.J., 2014. Refined Candidate Region for F4ab/ac Enterotoxigenic Escherichia coli Susceptibility Situated Proximal to MUC13 in Pigs. PLoS One 9, e105013. doi:10.1371/journal.pone.0105013

Gol, S., Estany, J., Fraile, L.J., Pena, R.N., 2015. Expression profiling of the GBP1 gene as a candidate gene for porcine reproductive and respiratory syndrome resistance. Anim. Genet. 46, 599–606. doi:10.1111/age.12347

Guy, S.Z.Y., Thomson, P.C., Hermesch, S., 2012. Selection of pigs for improved coping with health and environmental challenges: breeding for resistance or tolerance? Front. Genet. 3, 281. doi:10.3389/fgene.2012.00281

Halbur, P.G., Rothschild, M.F., Thacker, B.J., Meng, X.-J., Paul, P.S., Bruna, J.D., 1998. Differences in susceptibility of Duroc, Hampshire, and Meishan pigs to infection with a high virulence strain (VR2385) of porcine reproductive and respiratory syndrome virus (PRRSV). J. Anim. Breed.

Genet. 115, 181–189. doi:10.1111/j.1439-0388.1998.tb00341.x

Huang, S.Y., Chung, M.T., Tsou, H.L., Li, H.L., 2008. Association of polymorphism in the alpha (1,2) fucosyltransferase gene with growth performance in two Western pig breeds in Taiwan. Livest. Sci. 114, 336–340. doi:10.1016/j.livsci.2007.08.012

Jorgensen, C.B., Cirera, S., Archibald, A., Andersson, L., Fredholm, M., Edfors-Lilja, I., 2004. Porcine polymorphisms and methods for detecting them. US Patent 10536184 November 25, 2003

Kadarmideen, H.N., 2008. Biochemical, ECF18R, and RYR1 gene polymorphisms and their associations with osteochondral diseases and production traits in pigs. Biochem. Genet. 46, 41–53. doi:10.1007/s10528-007-9127-5

Kappes, M.A., Faaberg, K.S., 2015. PRRSV structure, replication and recombination: Origin of phenotype and genotype diversity. Virology 479-480, 475-486. doi:10.1016/j.virol.2015.02.012

Koltes, J.E., Fritz-Waters, E., Eisley, C.J., Choi, I., Bao, H., Kommadath, A., Serão, N.V.L., Boddicker, N.J., Abrams, S.M., Schroyen, M., Loyd, H., Tuggle, C.K., Plastow, G.S., Guan, L., Stothard, P., Lunney, J.K., Liu, P., Carpenter, S., Rowland, R.R.R., Dekkers, J.C.M., Reecy, J.M., 2015. Identification of a putative quantitative trait nucleotide in guanylate binding protein 5 for host response to PRRS virus infection. BMC Genomics 16, 412. doi:10.1186/s12864-015-1635-9

Knap, P.W., 2005. Breeding robust pigs. Aust. J. Exp. Agric. 45, 763. doi:10.1071/EA05041

Lunney, J.K., Steibel, J.P., Reecy, J.M., Fritz, E., Rothschild, M.F., Kerrigan, M., Trible, B., Rowland, R.R., 2011. Probing genetic control of swine responses to PRRSV infection: current progress of the PRRS host genetics consortium. BMC Proc. 5 Suppl 4, S30. doi:10.1186/1753-6561-5-S4-S30

Meijerink, E., Fries, R., Vögeli, P., Masabanda, J., Wigger, G., Stricker, C., Neuenschwander, S., Bertschinger, H.U., Stranzinger, G., 1997. Two alpha (1,2) fucosyltransferase genes on porcine chromosome 6q11 are closely linked to the blood group inhibitor (S) and Escherichia coli F18 receptor (ECF18R) loci. Mamm. Genome 8, 736–741

Mellencamp, M.A., Galina-Pantoja, L., Gladney, C.D., Torremorell, M., 2008. Improving pig health through genomics: a view from the industry. Dev. Biol. 132, 35–41. doi:10.1159/000317142

Moon, H.W., Hoffman, L.J., Cornick, N.A., Booker, S.L., Bosworth, B.T., 1999. Prevalences of Some Virulence Genes among *Escherichia Coli* Isolates from Swine Presented to a Diagnostic Laboratory in Iowa. J. Vet. Diagnostic Investig. 11, 557–560. doi:10.1177/104063879901100617

Neumann, E.J., Kliebenstein, J.B., Johnson, C.D., Mabry, J.W., Bush, E.J., Seitzinger, A.H., Green,

A.L., Zimmerman, J.J., 2005. Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States. J. Am. Vet. Med. Assoc. 227, 385–392. doi:10.2460/javma.2005.227.385

Niu, P., Kim, S.W., Kim, W.I., Kim, K.-S., 2015. Association analyses of DNA polymorphisms in immune-related candidate genes GBP1, GBP2, CD163, and CD169 with porcine growth and meat quality traits. J. Biomed. Res. 16, 40–46. doi:10.12729/jbr.2015.16.2.040

Niu, P., Shabir, N., Khatun, A., Seo, B.-J., Gu, S., Lee, S.M., Lim, S.K., Kim, K.S., Kim, W.I., 2016. Effect of polymorphisms in the GBP1, Mx1 and CD163 genes on host responses to PRRSV infection in pigs. Vet. Microbiol. 182, 187–195. doi:10.1016/j.vetmic.2015.11.010

Pilla, D.M., Hagar, J.A., Haldar, A.K., Mason, A.K., Degrandi, D., Pfeffer, K., Ernst, R.K., Yamamoto, M., Miao, E.A., Coers, J., 2014. Guanylate binding proteins promote caspase-11-dependent pyroptosis in response to cytoplasmic LPS. Proc. Natl. Acad. Sci. U. S. A. 111, 6046–6051. doi:10.1073/pnas.1321700111

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575. doi:10.1086/519795

Rasschaert, K., Verdonck, F., Goddeeris, B.M., Duchateau, L., Cox, E., 2007. Screening of pigs resistant to F4 enterotoxigenic Escherichia coli (ETEC) infection. Vet. Microbiol. 123, 249–253. doi:10.1016/j.vetmic.2007.02.017

Rose, N., Andraud, M., 2017. The use of vaccines to control pathogen spread in pig populations. doi:10.1186/s40813-017-0053-6

Rothschild, M.F., Hu, Z.-L., Jiang, Z., 2007. Advances in QTL Mapping in Pigs. Int. J. Biol. Sci 3 (3):192-197

Schiavo, G., Galimberti, G., Calò, D.G., Samorè, A.B., Bertolini, F., Russo, V., Gallo, M., Buttazzoni, L., Fontanesi, L., 2016. Twenty years of artificial directional selection have shaped the genome of the Italian Large White pig breed. Anim. Genet. 47, 181–191. doi:10.1111/age.12392

Schroyen, M., Stinckens, A., Verhelst, R., Niewold, T., Buys, N., 2012. The search for the gene mutations underlying enterotoxigenic Escherichia coli F4ab/ac susceptibility in pigs: a review. Vet. Res. 43, 70-78. doi:10.1186/1297-9716-43-70

Schroyen, M., Eisley, C., Koltes, J.E., Fritz-Waters, E., Choi, I., Plastow, G.S., Guan, L., Stothard, P., Bao, H., Kommadath, A., Reecy, J.M., Lunney, J.K., Rowland, R.R.R., Dekkers, J.C.M., Tuggle, C.K., 2016. Bioinformatic analyses in early host response to Porcine Reproductive and

Respiratory Syndrome virus (PRRSV) reveals pathway differences between pigs with alternate genotypes for a major host response QTL. BMC Genomics 17, 1–16. doi:10.1186/s12864-016-2547-z

Sellwood, R., Gibbons, R.A., Jones, G.W., Rutter, J.M., 1975. Adhesion of enteropathogenic Escherichia coli to pig intestinal brush borders: the existence of two pig phenotypes. J. Med. Microbiol. 8, 405–411. doi:10.1099/00222615-8-3-405

Stephens, M., Smith, N., and Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet., 68, 978-989.

Tabor, H.K., Risch, N.J., Myers, R.M., 2002. Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat. Rev. Genet. 3, 391–397. doi:10.1038/nrg796

Vögeli, P., Meijerink, E., Fries, R., Neuenschwander, S., Vorländer, N., Stranzinger, G., Bertschinger, H.U., 1997. A molecular test for the detection of E. coli F18 receptors: a breakthrough in the struggle against edema disease and post-weaning diarrhea in swine. Schweiz. Arch. Tierheilkd. 139, 479–484.

Waide, E.H., Tuggle, C.K., Serão, N.V.L., Schroyen, M., Hess, A., Rowland, R.R.R., Lunney, J.K., Plastow, G., Dekkers, J.C.M., 2017. Genomewide association of piglet responses to infection with one of two porcine reproductive and respiratory syndrome virus isolates1. J. Anim. Sci. 95, 16–38. doi:10.2527/jas.2016.0874

Yan, X.M., Ren, J., Huang, X., Zhang, Z.Y., Ouyang, J., Zeng, W.H., Zou, Z.Z., Yang, S.J., Yang, B., Huang, L.S., 2009. Comparison of production traits between pigs with and without the Escherichia coli F4 receptors in a White Duroc x Erhualian intercross F2 population. J. Anim. Sci. 87, 334–339. doi:10.2527/jas.2008-1107

Zhao, S., Zhu, M., Chen, H., 2012. Immunogenomics for identification of disease resistance genes in pigs: a review focusing on Gram-negative bacilli. doi:10.1186/2049-1891-3-34

Zhou, X., Stephens, M., 2012. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 44, 821–824. doi:10.1038/ng.2310

Supplementary material

Supplementary Table 1: PCR conditions and genotyping protocols.

Genes/markers	**PCR primers**	**Temp.**[1]	**Product size**[2]	**RFLP**[3]
FUT1	F:5'-CTTCCTGAACGTCTATCAAGACC-3' R:5'-CTTCAGCCAGGGCTCCTTTAAG-3'	62.5	421	*Hha*I; Alelle A: 328+93; Allele G: 241+93+87
MUC4	F:5'-GTGCCTTGGGTGAGAGGTTA-3' R:5'-CACTCTGCCGTTCTCTTTCC-3'	57.0	367	*Xba*I; Allele C: 367; Allele G: 216+151
GBP1 rs340943904[4]	F:5'-TCCTGATAGAATCTTTGCCGC-3' R:5'-GGAGGGAGAAGGATGGGTAC-3'	60.0	218	*Fsp*BI; Allele T: 113+29+23; Allele G: 136+29

[1] Annealing temperature in °C.
[2] Amplified fragment size (in bp).
[3] Restriction enzymes used in the RFLP methods are indicated together with the resulted fragments of the two alleles obtained after enzyme digestion (in bp).
[4] The *GBP1* rs80800372 marker was analyzed using the Illumina PorcineSNP60 BeadChip.

Supplementary Table 2: Hardy-Weinberg Equilibrium (HWE) test for all single nucleotide polymorphisms in the analyzed pig breeds.

SNP		**Italian Large White**	**Apulo-Calabrese**	**Casertana**	**Cinta Senese**	**Nero Siciliano**
MUC4	GG CG CC	89 247 164	10 5 0	22 5 0	0 0 22	16 11 3
	***P* value**	0.810	0.315	0.474	-	0.599
FUT1	GG AG AA	406 91 13	22 3 0	8 9 11	25 4 0	18 5 0
	***P* value**	0.012	0.661	0.062	0.586	0.434
GBP1 **rs340943904**	GG TG TT	467 83 7	21 7 0	22 8 0	5 8 3	16 8 2
	***P* value**	0.168	0.316	0.266	0.949	0.508
GBP1 **rs80800372**	AA AG GG	485 67 5	21 7 0	24 6 0	13 3 0	19 5 2
	***P* value**	0.162	0.316	0.413	0.577	0.127

Supplementary Table 3: Resistance allele frequency values of four SNPs associated to disease resistance in Italian Large White boars under the national selection program over about two decades.

Groups of years	*MUC4* – C rs338992994	*FUT1* – A rs335979375	*GBP1* - T rs340943904	*GBP1* – G rs80800372
1992-1995	0.40	0.13	0.15	0.05
1996-1997	0.48	0.25	0.13	0.08
1998-1999	0.43	0.15	0.04	0.02
2000-2001	0.40	0.13	0.06	0.04
2002-2003	0.48	0.13	0.00	0.00
2004-2005	0.53	0.09	0.05	0.03
2006-2007	0.47	0.13	0.03	0.03
2008-2010	0.52	0.06	0.04	0.04

Supplementary Table 4: Results of the association study between the analyzed markers and haematological traits in the second group of Italian Large White pigs. The table reports the calculated P nominal value.

Haematological traits[1]	*MUC4* rs338992994	*FUT1* rs335979375	*GBP1* rs340943904	*GBP1* rs80800372
HGB	0.927	0.205	0.876	0.408
HCT	0.942	0.119	0.869	0.498
RBC	0.823	0.030	0.881	0.414
MCV	0.917	0.662	0.764	0.821
MCHC	0.971	0.485	0.939	0.660
MCH	0.898	0.647	0.701	0.912
RDW	0.322	0.416	0.108	0.537
MPV	0.012	0.703	0.584	0.463
PLT	0.222	0.977	0.343	0.454
WBC	0.577	0.803	0.992	0.154
LINFO	0.012	0.821	0.142	0.216
EOSINO	0.069	0.677	0.803	0.447
NEUTRO	0.750	0.931	0.466	0.558
BASO	0.436	0.697	0.329	0.112
MONO	0.622	0.252	0.309	0.017

[1] HGB: Hemoglobin; HCT: Hematocrit; RBC: Red blood cell count; MCV: Mean corpuscular volume; MCHC: Mean corpuscular hemoglobin concentration; MCH: Mean corpuscular hemoglobin; RDW: Red cell distribution width; MPV: Mean platelet volume; PLT: Platelet count;WBC: White blood cell count; LINFO: Lymphocyte count; EOSINO: Eosinophil count; NEUTRO: Neutrophil count; BASO: Basophil count; MONO: Monocyte count.

Chapter 5

Phenotypic changes in the domestication process

As previously reported, genomic innovation in the field of livestock production has resulted in the opportunity to study the process of pig domestication by identifying regions of the genome that provide information about selective sweeps and breed differentiation. More specifically both, selective and non-selective evolutionary processes were investigated for defining milestones in the history of pig domestication. One of the most representative examples of selective process is undoubtedly the one regarding coat colour phenotype (Wiener and Wilkinson, 2011).

Coat colour is considered an impacting morphological change in the domestication process and the first genetic studies, about its variation in pigs, started around 1900 (Andersson and Plastow, 2011). Genetic variation of coat colour has been deeply investigated, and molecular studies (Johansson et al., 1992; Mariani et al., 1996; Kim et al., 2000) reported three major coat colour loci in the pig genome: *Dominant white* locus (I), *Extension* locus (E) and *Agouti* locus (A).

In pigs, the two major genes, *KIT* (v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog, also known as *Dominant white* locus) and *MC1R* (melanocortin 1 receptor, also known as *Extension* locus) are responsible for most of the coat colour variation among the European pig breeds with some other minor genes participating in defining pattern of pigmentation. (Fontanesi and Russo, 2013).

Dominant white phenotype is the most recurrent coat colour in European domestic pigs and in commercial pig breeds (Landrace and Large White). This white skin phenotype is linked to mutations occurring in the *KIT* gene, that encodes for the mast cell growth factor receptor, and has (additionally to other biological functions, like erythropoiesis or T-cell differentiation) an important role in the migration and survival of melanocyte precursor cells. Allele series variability, copy number variations (CNV) and a splice site mutation (Marklund et al., 1998; Pielberg et al., 2002; Johansson et al., 2005; Fontanesi et al., 2010) were identified as responsible for different pigmentation in pigs going from dominant white colour in commercial breeds to belted (i.e. Cinta Senese) or spotted (i.e. Pietrain) phenotypes.

Coat colour pigmentation is also defined by the *Extension* locus in pigs, as well as in other domesticated animal species like horses, goats and cattle breeds (Crepaldi et al., 2005). *MC1R* gene is considered an extremely reliable marker for coat colour definition and skin colour variation (Dun et al., 2007). This gene encodes for the melanocyte-stimulating hormone receptor that belongs to the family of G-protein coupled receptors and by signaling, it regulates the relative amount of eumelanin

(black/brown) and phaeomelanin (red/yellow) pigments synthetized by melanocytes of hair follicles (Kijas et al., 1998).

The allelic series at *Extension locus* in pigs was sequenced and characterized for its role in impacting pigmentation. Mutations occurring in the sequence lead to different coat colours, spanning from red to black phenotypes (Wiener and Wilkinson, 2011). More in detail, the wild-type (E^{+}) allele is responsible for the production of both pigments turning in a grey/brown coat with variable shades, that is characteristic of European wild boar. Two evolutionary independent mutations are known to be associated with the dominant black colour, and they are the black alleles E^{D1} and E^{D2}. Two additional alleles can be found at this locus, and they are the recessive red allele *e* (loss of function mutation with recessive coat colour, i.e. Duroc breed) and E^{P} which leads to a black spotting on a red or white background (Fang et al., 2009).

European pig breeds show a large number of differences for this phenotype and there are some breeds in which specific *MC1R* haplotypes are considered unique and characteristic, with the implication of using *MC1R* alleles as markers for the differentiation of meat and meat derived products, obtained from wild boar or other domestic pig breeds (D’Alessandro et al., 2007; Fontanesi et al., 2014).

In addition to coat colour phenotype, another important characteristic that is used to differentiate between wild boar and domestic pig breeds, is the number of vertebral bodies (Klomtong et al., 2015). During the domestication process, the intensive selective breeding for an enlarged body size has progressively led to changes in western commercial pig breeds, presenting a higher number of vertebrae (n= 22/23) when compared to wild boar (n=19). This, of course, turned to be impacting in terms of body size, carcass length and meat production (Yang et al., 2009). Two major QTLs affecting the number of vertebrae with a combined effect have been fine mapped and localized in two different pig chromosomes by Mikawa et al. (2007). Two candidate genes underlying these two QTLs were identified in *VRTN* and *NR6A1* genes.

VRTN is also called “vertebrae development associated” gene and its role in the definition of vertebrae number has been investigated in different studies (Hirose et al., 2013; Fan et al., 2013). An insertion in the promoter sequence of this gene is responsible for its effect on thoracic vertebrae number, and the frequency of this genetic variation is variable in western pig breeds. However, there are still contrasting results regarding studies of association between VRTN gene and meat and carcass traits in both commercial and local pig breeds (Burgos et al., 2015).

The other candidate gene in linkage disequilibrium with QTL on chromosome 1, is *NR6A1*. The causative mutation for this QTL is a missense substitution found in the sequence of the nuclear

receptor subfamily 6, group A, member 1 (*NR6A1*) gene (Yang et al., 2009). The amino acid substitution is in the hinge region of the *NR6A1* gene and the presence of this polymorphism alters the binding of the encoded nuclear receptor to its coregulators, thus enabling the biological role of the protein (Mikawa et al., 2007).

The mutation is associated to an increased number of vertebrae and it results to be fixed in European commercial pig breeds, with the wild type genotype fixed instead in wild boars (Fontanesi et al., 2014; Burgos et al., 2015; Klomtong et al., 2015).

Information about the distribution of these genetic mutations in coat colour and vertebrae number related genes are not always available for autochthonous pig breeds. Considering that wild boars and domestic pigs belong to the same species (*Sus scrofa*) and that, especially for local pig populations, they can also share the same environment, assessing the possible interactions between wild and domestic animals is extremely important for their genetic characterization.

Furthermore, the study of the genetic diversity in these regions (with specific reference to *MC1R* and *NR6A1*) has an important implication to understand the real impact of the selective pressure in domestic animals when compared to wild relatives. The study of the introgression of wild boar populations in the domesticated genomes is a key point in understanding the history and evolution of divergence between commercial pigs and wild boars and to fully understand variability among local pig breeds (Fang et al., 2009).

Starting from the awareness that autochthonous pig breeds are usually reared in extensive or semi-extensive production systems that might facilitate the contact with wild boars and, thus, reciprocal genetic exchanges, we decided to investigate variants in *MC1R* and *NR6A1* genes in pigs of twelve local pig breeds raised in Italy and South-East countries. We compared the data with the genetic variability at these loci investigated in wild boars from populations spread in the same macro-geographic areas.

5.1 Signatures of de-domestication in autochthonous pig breeds and of domestication in wild boar populations from *MC1R* and *NR6A1* allele distribution

Introduction

The domestication process in all livestock species has been determined by a complex series of spatial-temporal events causing continuous genetic changes derived by population admixture and isolation that shaped the animal genome from the corresponding ancestral wild genetic pools (Larson and

Burger, 2013). The reconstruction of the domestication history of the pig (*Sus scrofa*), from the earliest events (that might date back some 9,000-10,000 years ago) till the constitution of the modern breeds largely relied on local wild boar populations that were the sources of the domestic pools (e.g. Larson et al., 2007, 2010; Ramos-Oisins et al., 2014; Iacolina et al., 2016; Yang et al., 2017). It also seems clear that the domestication of the pig was not based on few fixed events but occurred over many millennia and included repeated admixture of domestic populations with local wild boars, that, in most cases, shared the same environments (Larson and Burger, 2013). The results of this continuous process determined several morphological, behavioral and physiological modifications of the pigs that satisfied the primary farmers' needs and led to the fixation of a few traits (e.g. coat colour and shape of the animals), regarded among the first domestication-derived phenotypes (Clutton-Brock, 1999).

Coat colour in pigs is largely affected by variability at the *Extension* locus. This locus is characterized by different alleles at the melanocortin 1 receptor (*MC1R*) gene (Kijas et al., 1998, 2001). Among the described *MC1R* alleles, the wild type allele (E^{+}, indicated also as allele *0101*; Fang et al., 2009) is the typical form in European wild boars (determining the classical grey/brown coat colour), whereas several other alleles are considered as domestic variants. These include i) alleles E^{D1} (indicated as alleles *0201*, *0202* and *0203*; of Asian origin) and E^{D2} (or allele *0301;* of European origin) that cause the dominant black coat colour, ii) allele E^{P} (identified also as alleles *0501*, *0502* and *0503*) which is common in spotted and completely white pigs and iii) allele *e* which is the recessive allele producing a reddish coat colour of the Duroc breed. Most of the commercial pig breeds are fixed for one of the domestic alleles (Kijas et al., 1998, 2001; Fontanesi et al., 2010).

Another mutation affecting a domestication-selected trait has been reported in the nuclear receptor subfamily 6 group A member 1 (*NR6A1*) gene. The mutated allele (derived by a missense mutation: p.P192L) is associated with an increased number of thoracic and dorsal vertebrae (21-23 *vs* 19 vertebrae) as compared to the wild type allele (Mikawa et al., 2007). The positive effect of the mutated allele on the number of vertebrae that is, in turn, associated with increased length of the animals, more meat, higher number of teats (1-3 more teats) and thus increased reproduction potentials of the sows (Borchers et al., 2004; Mikawa et al., 2007). These effects indirectly determined its fixation in commercial pig breeds and lines through a directional selection pressure on these traits (e.g. Rubin et al., 2012; Fontanesi et al., 2014).

Domestic alleles at both *MC1R* and *NR6A1* genes have been also described in wild boar populations as the result of introgression from domestic populations, probably derived by accidental crossbreeding with free-ranging domestic pigs or by deliberate crosses in captive farming systems

and subsequent release of crossbred animals to improve performances of wild boar populations (e.g. Koutsogiannouli et al., 2010; Frantz et al., 2012, 2013; Fontanesi et al., 2014).

Local pig breeds are usually constituted by small populations that have recently (or at some time during their developmental history) experienced bottlenecks, genetic drifts and, in some cases, introgression from other populations that contributed to increasing variability and reducing inbreeding (Porter, 1993). Local pig breeds are usually reared in extensive or semi-extensive production systems and are less productive than commercial breeds (considering both performance and reproductive traits). As the populations of many local breeds are usually too small to run effective selection programmes, only conservation programmes can be considered. Polymorphisms in coat colour genes have been proposed as useful markers for the authentication of mono-breed pork products and might be also considered as targets for their characterization (Kijas et al., 1998; D'Alessandro et al., 2007; Fontanesi, 2009; Fontanesi et al., 2016).

In this study, we analysed *MC1R* and *NR6A1* gene variants in twelve local pig breeds raised in Italy and South-East Europe (Slovenia, Croatia, Serbia and Bulgaria) and compared the data with the genetic variability at these loci investigated in wild boar populations spread in same three separated macro-geographic areas: one represented by Sardinia (isolated because of a geographical barrier recently strengthened by the ban of exchange of pigs and wild boars caused by the presence of African Swine Fever since 1978; Jurado et al., 2018), another constituted by the Italian peninsula which traditionally has a continuous genetic flow with Sicily and the third represented by the Balkan countries.

Materials and methods

A total of 712 pigs belonging to six Italian [Apulo-Calabrese, n. 73; Casertana, n. 114; Cinta Senese (Siena Belted), n. 80; Mora Romagnola, n. 74; Nero Siciliano (Sicilian Black), n. 70; and Sarda, n. 58], one Slovenian (Krškopolje, n. 31), two Croatian (Black Slavonian, n. 27; Turopolje, n. 47), two Serbian (Mangalitsa, n. 47; Moravka, n. 47) and one Bulgarian (East Balkan Swine, n. 44) autochthonous breeds were investigated (detailed information on the geographical distribution, standard coat colour, vertebrae and teat range numbers of these breeds are reported in Table S1). In addition to autochthonous pig populations, a total of 229 wild boars were sampled. Collection was opportunistic, derived by hunting or via inspection of carcasses/hunted animals by forest policemen. Of these wild boars, 139 were sampled in two isolated Italian areas (113 in the Appennini mountains in North of Italy; and 26 in Sardinia island). The remaining 90 wild boars were from western and central Balkan countries (16 from Bosnia and Herzegovina, 10 from Croatia, 16 from Montenegro, 16 from the North of Macedonia, 17 from Serbia and 15 from Slovenia) that altogether were thereafter

named as South-East (SE) European population. Information of the investigated wild boars are reported in Table S2. Blood, hair roots or meat were sampled from these animals. DNA was extracted using the Wizard (R) Genomic DNA Purification kit (Promega Corporation, Madison, WI, USA) or by standard phenol–chloroform–isoamyl alcohol extraction (Sambrook et al., 1989).

Five autosomal polymorphisms were genotyped: three single nucleotide polymorphisms (SNPs) and one insertion/deletion (indel) in the *MC1R* gene that, on the whole, can distinguish all major alleles at the *Extension* locus (i.e. E^{+}, E^{D1}, E^{D2}, E^{P} and *e*) described by Kijas et al. (1998, 2001); the missense mutation in the *NR6A1* gene (rs326780270 C>T or p.P192L) that is considered the causative mutation of the QTL for number of vertebrae reported on porcine chromosome 1 (Mikawa et al., 2007). Genotyping protocols were based on PCR-RFLP, fragment analysis of PCR amplicons and on an OpenArray™ genotyping platform (details are reported in Table S3).

Results and discussion

Genotyping results obtained in the autochthonous pig breeds and in the wild boar populations are reported in Table S3 and Table S5, respectively. Allele frequencies at the *MC1R* and *NR6A1*genes are summarized in Table 1. Figure 1 reports a Principal Component Analysis (PCA) plot based on the first two principal components (PC1 and PC2) obtained analyzing allele frequency data with the R “prcomp” function and a cluster representation of the analysed breeds and populations generated with the “dist” and “hclust” functions of R using allele frequency information to calculate the Euclidean distance among groups (R Core Team 2018).

None of the pig breeds and none of the wild boar populations were fixed for one allele at both loci. The East Balkan Swine breed showed all *MC1R* and *NR6A1* alleles. A few breeds showed only one allele at one of the investigated genes. In particular: Turopolje was fixed for the *MC1R* E^{P} allele; Mangalitsa was fixed for the wild type E^{+} allele, as already reported by Fang et al. (2009), which confirms the inference determined by its phenotypic description (Porter, 1993). The fixation of the wild type allele in Mangalitsa could facilitate the expression of the agouti signalling protein (*ASIP*) gene, determining the classical black-and-tan phenotype (according to the epistatic interaction between the *MC1R* and *ASIP* genes) of this breed (Drögemüller et al., 2006). The wild type allele was also observed in all Italian (Sarda was the breed with the highest frequency: 30%) and in two East European breeds (Moravka, 3%; and East Balkan Swine, 28%). Among the Italian breeds, allele E^{D2} at the *MC1R* gene was the most frequent one (ranging from 57% in Nero Siciliano to 86% in Apulo-Calabrese) except in the Mora Romagnola breed, which showed a quite unique genetic structure with only two alleles: the highest frequency of the recessive *e* allele (82%) and the presence of the E^{+} allele (18%). All other Italian breeds showed at least four of the five analysed *MC1R* alleles.

Allele E^{D1} was highly frequent in the Black Slavonian (88%) and quite frequent in Moravka (37%), testifying their genetic origin or contamination with Asian derived populations or breeds, including Large Black and Berkshire pigs (which are well known to be derived by crossbreeding with Chinese pigs; Porter, 1993; Megens et al., 2008; Gvozdanović et al., 2018).

Krškopolje showed a high frequency (82%) of the other dominant black allele (E^{D2}). At the *NR6A1* gene, Cinta Senese, Mora Romagnola, Krškopolje and Moravka pigs showed only the domestic allele (i.e. *T*). This allele was the most frequent in almost all other breeds except in the East Balkan Swine in which allele *C* had frequency 81%, suggesting a high level of wild boar gene introgression. Based on the *MC1R* and *NR6A1* data, this breed was the closest one to the wild boar populations (Fig. 1).

Only the Sardinian wild boar population was fixed for the *MC1R* E^{+} allele even if at the *NR6A1* gene the frequency of the domestic allele was 12%. This allele was also identified in the other two wild boar populations, although at lower frequencies (2% in the North Italian wild boars and 7% in the SE European group). Three (E^{D2}, E^{P} and *e*) and two ($E^{D2}$and E^{P}) *MC1R* domestic alleles were identified in the North Italian and SE-European wild boar populations, respectively. The most frequent domestic alleles were E^{D2} (11%) and E^{P} (5%) in the two populations, respectively.

These results clearly indicated that autochthonous breeds have a complex history. These breeds could have recently experienced crossbreeding with other commercial breeds that might have introduced or re-introduced heterogeneity at the *MC1R* gene. Crossbreeding is mentioned in several historical records (available for some of the investigated breeds; e.g. Mora Romagnola crossed with Duroc) and testified by oral transmitted information (e.g. Nero Siciliano crossed with Pietrain). The presence of the *MC1R* E^{+} allele could be derived by planned (the case of Mora Romagnola) or accidental crossbreeding with wild boars that also might be the source of introgression of the wild type allele (i.e. *C*) of the *NR6A1* gene observed in local pig breeds. These genetic fluxes could be important to re-introduce genetic variability in pig populations that usually have quite a high level of inbreeding. On the other hand, this heterogeneity should be managed to create phenotypically uniform populations that might better match the standard traits defined by their herd books, useful to acquire a specific identity of these pig genetic resources and make it possible the application of DNA based systems for the authentication of their products (Fontanesi, 2009; Fontanesi et al., 2016).

Moreover, these results confirm that wild boar genetic integrity has been "polluted" by domestic alleles (as already described: e.g. Goedbloed et al., 2013) which might be putatively derived from commercial pig and/or local breeds (e.g. *MC1R* E^{P} allele, fixed in commercial white breeds but also in Turopolje; Table 1 and Fontanesi et al., 2010) or from autochthonous black pig breeds (i.e. E^{D2}) that are in close contact with wild animals, when raised in free-range systems (see Table 1). The introgression direction of the *MC1R* and *NR6A1* alleles from domestic pig breeds to wild boar

populations was tested with the four-taxon ABBA/BABA test computing the Patterson's *D*-statistic (Dasmahapatra et al., 2012; Martin et al., 2015). Four populations can be considered in the model and five plausible scenarios were tested, defining different wild boar and domestic populations in the ABBA positions (see Table S6 and Fig. S1). In these scenarios, the outgroup population (P4) was a hypothetical ancestral wild boar population with ancestral allele frequency equal to 1 (as requested by the test), P3 was a domestic population (defined by one single breed or grouping different breeds, using the geographic criteria), P2 and P1 were different wild boar populations. Even if the test cannot be formally used to evaluate the opposite introgression (as it would have assumed that a derived pig is the ancestor of a wild boar, contrasting the coalescent model), we also modelled one BABA scenario assuming an ancestral domestic population, with domestic allele frequency equal to 1. In all tested scenarios for both loci, absolute *D* values ranged from 0.184 to 1.000 (Table S6). Results indicated an ABBA excess meaning that there were events of domesticated alleles introgression whereas the tested BABA scenario, even if not formally appropriated, might indicate an inverse introgression flow.

The relatively high frequency of the domestic *NR6A1* allele in wild boar populations might be derived by a reproductive advantage, and in turn, a slightly higher fitness of the carriers of the domestic allele that might tend to increase its frequency in natural environments. An increased number of vertebrae associated to the domestic allele would contribute to increased body size and length of the animals, with subsequent effects on reproduction traits, obtained directly with an increment in litter size (derived by a higher uterus capacity) or indirectly with an increased number of teats as also reported in QTL studies with domestic pigs (Duijvesteijn et al., 2014). This matter should be better investigated to acquire phenotypic evidences associated to genotyping data (Fulgione et al., 2016). Combining information from the two investigated genes, signals of introgression of domestic alleles were observed in 16% (North of Italy) and 34% (SE-Europe) of the investigated wild boars (in Sardinia, 12% carried the *NR6A1* C allele). These regions constitute three contiguous but separated European areas considered among those with the highest number of autochthonous pig breeds (Porter 1993). From a molecular ecology perspective, highly introgressed populations could provide the opportunity to evaluate the effect of natural selection on domestic alleles that reached wild populations through interdemic gene flow (Fulgione et al., 2016).

Conclusions

Altogether these results indirectly demonstrate that bidirectional introgression of wild and domestic alleles is part of the human- and naturally-driven evolutionary forces that are continuously shaping the *Sus scrofa* genome. From one hand, this species is still experiencing a "de-domestication" process

(i.e. the genome of autochthonous breeds, that can be considered at the "domestication border") and, from the other hand, a "domestication" drift (i.e. the wild boar populations). In the case of autochthonous breeds, the domestication might still be considered a work in progress against de-domestication forces, led by the way in which these genetic resources are managed. Both aspects need to be further investigated and evaluated in genetic conservation programs of wildlife and domestic populations.

Remarks

The information reported in this chapter were submitted as short communication to Animal Genetics Journal in October 2018. The submitted material is currently under revision.

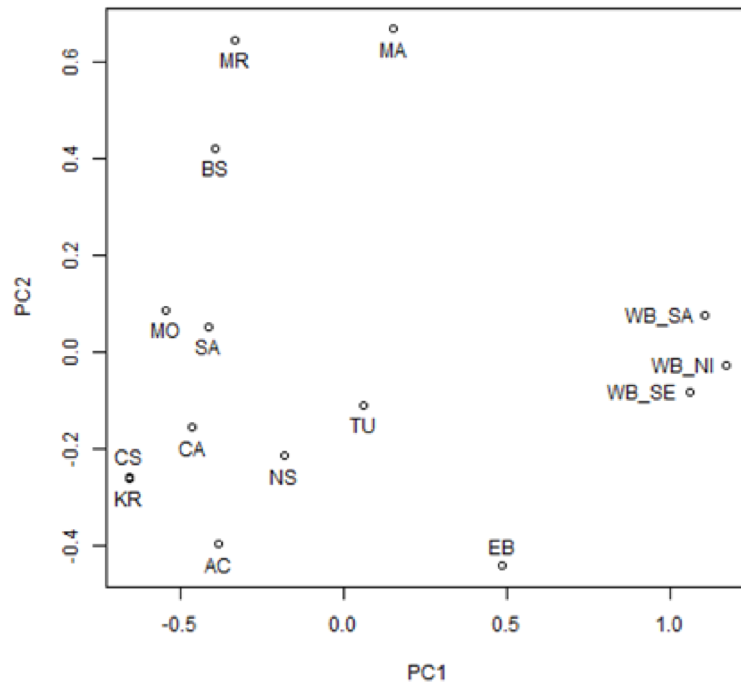
Acknowledgements

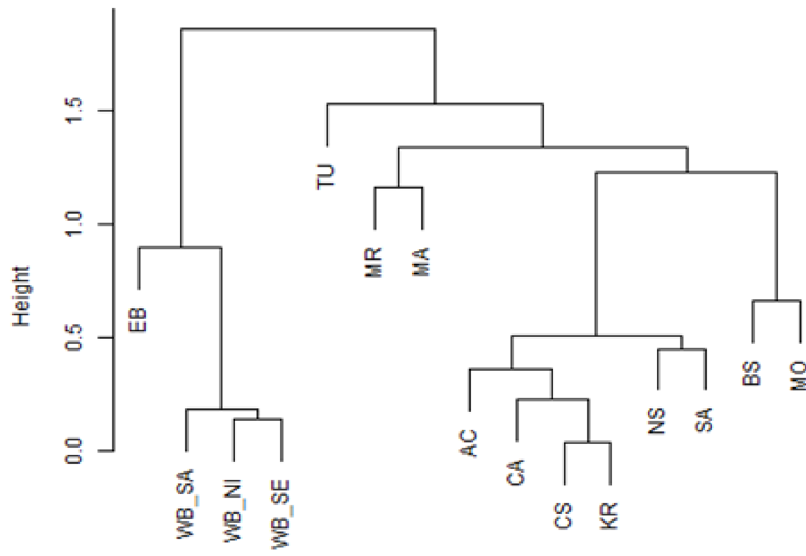
This work was realized thanks to the collaboration of many partners, therefore, we would like to thank for their contribution and efforts: the Italian Pig Breeders Association (Associazione Nazionale Allevatori Suini – ANAS), the Department of Biology and Ecology of the University of Novi Sad (Novi Sad, Serbia), the Association for Breeding and Preserving of the East Balkan Swine (Bulgaria), the Department of Pig Breeding and Genetics, Institute for Animal Husbandry (Belgrade, Serbia) the Faculty of Agriculture, University of Belgrade (Belgrade, Serbia), the Department of Animal Science of the University of Zagreb (Zagreb, Croatia), the Faculty of Agrobiotechnical Sciences, University of Osijek (Osijek, Croatia), the Kmetijski inštitut Slovenije (Ljubljana, Slovenia), the Department of Animal Genetic Improvement at INIA (Madrid, Spain) and the Agris Sardegna agency for scientific investigation (Sassari, Italy).

Table 1: Frequencies of melanocortin 1 receptor (*MC1R*) and nuclear receptor subfamily 6 group A member 1 (*NR6A1*) alleles in domestic pig breeds and wild boar populations.

Breeds/Populations	No. of animals	*MC1R* alleles					*NR6A1* alleles	
		E^+	E^{D1}	E^{D2}	E^P	*e*	*C*	*T*
Domestic pigs								
Apulo-Calabrese	73	0.05	0.00	0.86	0.03	0.06	0.23	0.77
Casertana	114	0.17	0.00	0.71	0.10	0.02	0.06	0.94
Cinta Senese	80	0.01	0.01	0.83	0.13	0.02	0.00	1.00
Mora Romagnola	74	0.18	0.00	0.00	0.00	0.82	0.00	1.00
Nero Siciliano	70	0.17	0.07	0.57	0.19	0.00	0.28	0.72
Sarda	58	0.30	0.06	0.58	0.00	0.06	0.01	0.99
Krškopolje	31	0.00	0.00	0.82	0.16	0.02	0.00	1.00
Black Slavonian	27	0.00	0.88	0.04	0.06	0.02	0.07	0.93
Turopolje	47	0.00	0.00	0.00	1.00	0.00	0.43	0.57
Mangalitsa	47	1.00	0.00	0.00	0.00	0.00	0.02	0.98
Moravka	47	0.03	0.37	0.43	0.17	0.00	0.00	1.00
East Balkan Swine	44	0.28	0.17	0.49	0.05	0.01	0.81	0.19
Wild boars								
North of Italy	113	0.92	0.00	0.01	0.05	0.02	0.98	0.02
Sardinia	26	1.00	0.00	0.00	0.00	0.00	0.88	0.12
South-East Europe	90	0.87	0.00	0.11	0.02	0.00	0.93	0.07

Figure 1: Graphical representations of the analysed breeds and wild boar populations based on *MC1R* and *NR6A1* data. **a)** Principal Component Analysis (PCA) plot. **b)** Cluster dendrogram of the pig breeds and wild populations. AC = Apulo-Calabrese; CA = Casertanta; CS = Cinta Senese; MR = Mora Romagnola; NS = Nero Siciliano; SA = Sarda; KR = Krškopolje; BS = Black Slavonian; TU = Turopolje; MA = Mangalitsa; MO = Moravka; EB = East Balkan Swine; WB_NI = Wild boars from North of Italy; WB_SA = Wild boars from Sardinia; WB_SE = Wild boars from South-East Europe.


a)
PC2
PC1
MR
MA
BS
MO
SA
WB_SA
WB_NI
WB_SE
TU
CS
CA
KR
NS
AC
EB


b)
Height
EB
WB_SA
WB_NI
WB_SE
TU
MR
MA
AC
CA
CS
KR
NS
SA
BS
MO

References

Andersson, L., Plastow, G., 2011. Molecular genetics of coat colour variation. In: Rothschild MF, Ruvinsky A, editors. The genetics of the pig. 2. Wallingford: CABI; 2011. pp. 38–50.

Borchers, N., Reinsch, N., Kalm, E., 2004. The number of ribs and vertebrae in a Pietrain cross: variation, heritability and effects on performance traits. J. Anim. Breed. Genet. 121, 392–403. doi:10.1111/j.1439-0388.2004.00482.x

Burgos, C., Latorre, P., Altarriba, J., Carrodeguas, J.A., Varona, L., López-Buesa, P., 2015. Allelic frequencies of NR6A1 and VRTN, two genes that affect vertebrae number in diverse pig breeds: A study of the effects of the VRTN insertion on phenotypic traits of a Duroc × Landrace–Large White cross. Meat Sci. 100, 150–155. doi:10.1016/ j.meatsci.2014.09.143

Clutton-Brock, J., 1999. A natural history of domesticated mammals. 2nd Edition. Cambridge University Press.

Crepaldi, P., Fornarelli, F., Marilli, M., 2005. MC1R gene: comparison between different farm animal species, Ital.J.Anim.Sci. 4, 43-45. https://doi.org/10.4081/ijas.2005.2s.43

D'Alessandro, E., Fontanesi, L., Liotta, L., Davoli, R., Chiofalo, V., Russo, V., 2007. Analysis of the MC1R Gene in the Nero Siciliano Pig Breed and Usefulness of this Locus for Breed Traceability. Vet. Res. Commun. 31, 389–392. doi:10.1007/s11259-007-0063-y

Dasmahapatra, K.K., Walters, J.R., Briscoe, A.D., Davey, J.W., Whibley, A., Nadeau, N.J., Zimin, A. V., Hughes, D.S.T., Ferguson, L.C., Martin, S.H., Salazar, C., Lewis, J.J., Adler, S., Ahn, S.-J., Baker, D.A., Baxter, S.W., Chamberlain, N.L., Chauhan, R., Counterman, B.A., Dalmay, T., Gilbert, L.E., Gordon, K., Heckel, D.G., Hines, H.M., Hoff, K.J., Holland, P.W.H., Jacquin-Joly, E., Jiggins, F.M., Jones, R.T., Kapan, D.D., Kersey, P., Lamas, G., Lawson, D., Mapleson, D., Maroja, L.S., Martin, A., Moxon, S., Palmer, W.J., Papa, R., Papanicolaou, A., Pauchet, Y., Ray, D.A., Rosser, N., Salzberg, S.L., Supple, M.A., Surridge, A., Tenger-Trolander, A., Vogel, H., Wilkinson, P.A., Wilson, D., Yorke, J.A., Yuan, F., Balmuth, A.L., Eland, C., Gharbi, K., Thomson, M., Gibbs, R.A., Han, Y., Jayaseelan, J.C., Kovar, C., Mathew, T., Muzny, D.M., Ongeri, F., Pu, L.-L., Qu, J., Thornton, R.L., Worley, K.C., Wu, Y.-Q., Linares, M., Blaxter, M.L., ffrench-Constant, R.H., Joron, M., Kronforst, M.R., Mullen, S.P., Reed, R.D., Scherer, S.E., Richards, S., Mallet, J., Owen McMillan, W., Jiggins, C.D., Jiggins, C.D., 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487, 94–98. doi:10.1038/nature11041

Drögemüller, C., Giese, A., Martins-Wess, F., Wiedemann, S., Andersson, L., Brenig, B., Fries, R., Leeb, T., 2006. The mutation causing the black-and-tan pigmentation phenotype of Mangalitza pigs maps to the porcine ASIP locus but does not affect its coding sequence. Mamm. Genome 17, 58–66. doi:10.1007/s00335-005-0104-1

Duijvesteijn, N., Veltmaat, J.M., Knol, E.F., Harlizius, B., 2014. High-resolution association mapping of number of teats in pigs reveals regions controlling vertebral development. BMC Genomics 15, 542. doi:10.1186/1471-2164-15-542

Dun, G., Li, X., Cao, H., Zhou, R., Li, L., 2007. Variations of Melanocortin Receptor 1 (MC1R) Gene in Three Pig Breeds. J. Genet. Genomics 34, 777–782. doi:10.1016/S1673-8527(07)60088-5

Fan, Y., Xing, Y., Zhang, Z., Ai, H., Ouyang, Z., Ouyang, J., Yang, M., Li, P., Chen, Y., Gao, J., Li, L., Huang, L., Ren, J., 2013. A Further Look at Porcine Chromosome 7 Reveals VRTN Variants Associated with Vertebral Number in Chinese and Western Pigs. PLoS One 8, e62534. doi:10.1371/journal.pone.0062534

Fang, M., Larson, G., Soares Ribeiro, H., Li, N., Andersson, L., Barsh, G.S., 2009. Contrasting Mode of Evolution at a Coat Color Locus in Wild and Domestic Pigs. doi:10.1371/journal.pgen.1000341

Fontanesi, L., 2009. Genetic authentication and traceability of food products of animal origin: new developments and perspectives. Ital. J. Anim. Sci. 8, 9–18. doi:10.4081/ijas.2009.s2.9

Fontanesi, L., D'Alessandro, E., Scotti, E., Liotta, L., Crovetti, A., Chiofalo, V., Russo, V., 2010. Genetic heterogeneity and selection signature at the KIT gene in pigs showing different coat colours and patterns. Anim. Genet. 41, 478–492. doi:10.1111/j.1365-2052.2010.02054.x

Fontanesi, L., Ribani, A., Scotti, E., Utzeri, V.J., Veličković, N., Dall'Olio, S., 2014. Differentiation of meat from European wild boars and domestic pigs using polymorphisms in the MC1R and NR6A1 genes. Meat Sci. 98, 781–784. doi:10.1016/j.meatsci.2014.07.026

Fontanesi, L., Russo, V., 2013. Acta argiculturae Slovenica, Supplement 4, 15-20, Ljubljana 2013 Invited lecture COBISS: 1.06 Agris category code: L10 Molecular genetics of coat colour in pigs.

Fontanesi, L., Scotti, E., Gallo, M., Nanni Costa, L., Dall'Olio, S., 2016. Authentication of "mono-breed" pork products: Identification of a coat colour gene marker in Cinta Senese pigs useful to this purpose. Livest. Sci. 184, 71–77. doi:10.1016/J.LivSci.2015.12.007

Frantz, A.C., Massei, G., Burke, T., 2012. Genetic evidence for past hybridisation between domestic pigs and English wild boars. Conserv. Genet. 13, 1355–1364. doi:10.1007/s10592-012-0379-1

Frantz, A.C., Zachos, F.E., Kirschning, J., Cellina, S., Bertouille, S., Mamuris, Z., Koutsogiannouli, E.A., Burke, T., 2013. Genetic evidence for introgression between domestic pigs and wild boars (Sus scrofa) in Belgium and Luxembourg: a comparative approach with multiple marker systems. Biol. J. Linn. Soc. 110, 104–115. doi:10.1111/bij.12111

Fulgione, D., Rippa, D., Buglione, M., Trapanese, M., Petrelli, S., Maselli, V., 2016. Unexpected but welcome. Artificially selected traits may increase fitness in wild boar. Evol. Appl. 9, 769–776. doi:10.1111/eva.12383

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

Gvozdanović, K., Margeta, V., Margeta, P., Djurkin Kušec, I., Galović, D., Dovč, P., Kušec, G., 2018. Genetic diversity of autochthonous pig breeds analyzed by microsatellite markers and mitochondrial DNA D-loop sequence polymorphism. Anim. Biotechnol. 1–10. doi:10.1080/10495398.2018.1478847

Hirose, K., Mikawa, S., Okumura, N., Noguchi, G., Fukawa, K., Kanaya, N., Mikawa, A., Arakawa, A., Ito, T., Hayashi, Y., Tachibana, F., Awata, T., 2013. Association of swine vertnin (*VRTN*) gene with production traits in Duroc pigs improved using a closed nucleus breeding system. J Anim. Sci. 84, 213–221. doi:10.1111/j.1740-0929.2012.01066.x

Iacolina, L., Brajković, V., Canu, A., Šprem, N., Cubric-Curik, V., Fontanesi, L., Saarma, U., Apollonio, M., Scandura, M., 2016. Novel Y-chromosome short tandem repeats in Sus scrofa and their variation in European wild boar and domestic pig populations. Anim. Genet. 47, 682–690. doi:10.1111/age.12483

Johansson, A., Pielberg, G., Andersson, L., Edfors-Lilja, I., 2005. Polymorphism at the porcine Dominant white/KIT locus influence coat colour and peripheral blood cell measures. Anim. Genet. 36, 288–296. doi:10.1111/j.1365-2052.2005.01320.x

Johansson, M., Ellegren, H., Marklund, L., Gustavsson, U., Ringmar-Cederberg, E., Andersson, K., Edfors-Lilja, I., Andersson, L., 1992. The gene for dominant white color in the pig is closely linked to ALB and PDGRFRA on chromosome 8. Genomics 14, 965–9.

Jurado, C., Fernández-Carrión, E., Mur, L., Rolesu, S., Laddomada, A., Sánchez-Vizcaíno, J.M., 2018. Why is African swine fever still present in Sardinia? Transbound. Emerg. Dis. 65, 557–566. doi:10.1111/tbed.12740

Kijas, J.M., Wales, R., Törnsten, A., Chardon, P., Moller, M., Andersson, L., 1998. Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. Genetics 150, 1177–85.

Kijas, J.M., Moller, M., Plastow, G., Andersson, L., 2001. A frameshift mutation in MC1R and a high frequency of somatic reversions cause black spotting in pigs. Genetics 158, 779–85.

Kim, K.S., Mendez, E.A., Marklund, S., Clutter, A.C., Pomp, D., 2000. Rapid communication: Linkage mapping of the porcine Agouti gene. J Anim Sci. 78, 1395-6.

Klomtong, P., Chaweewan, K., Phasuk, Y., Duangjinda, M., 2015. MC1R, KIT, IGF2, and NR6A1 as markers for genetic differentiation in Thai native, wild boars, and Duroc and Chinese Meishan pigs. Genet. Mol. Res. 14, 12723–12732. doi:10.4238/2015.October.19.16

Koutsogiannouli, E.A., Moutou, K.A., Sarafidou, T., Stamatis, C., Mamuris, Z., 2010. Detection of hybrids between wild boars (Sus scrofa scrofa) and domestic pigs (Sus scrofa f. domestica) in Greece, using the PCR-RFLP method on melanocortin-1 receptor (MC1R) mutations. Mamm. Biol. 75, 69–73. doi: 10.1016/j.mambio.2008.08.001

Larson, G., Albarella, U., Dobney, K., Rowley-Conwy, P., Schibler, J., Tresset, A., Vigne, J.-D., Edwards, C.J., Schlumbaum, A., Dinu, A., Balaçsescu, A., Dolman, G., Tagliacozzo, A., Manaseryan, N., Miracle, P., Van Wijngaarden-Bakker, L., Masseti, M., Bradley, D.G., Cooper, A., 2007. Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. Proc. Natl. Acad. Sci. U. S. A. 104, 15276–81. doi:10.1073/pnas.0703411104

Larson, G., Liu, R., Zhao, X., Yuan, J., Fuller, D., Barton, L., Dobney, K., Fan, Q., Gu, Z., Liu, X.-H., Luo, Y., Lv, P., Andersson, L., Li, N., 2010. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. Proc. Natl. Acad. Sci. U. S. A. 107, 7686–91. doi:10.1073/pnas.0912264107

Larson, G., Burger, J., 2013. A population genetics view of animal domestication. Trends Genet. 29, 197–205. doi: 10.1016/j.tig.2013.01.003

Mariani, P., Moller, M.J., Høyheim, B., Marklund, L., Davies, W., Ellegren, H., Andersson, L., 1996 The extension coat color locus and the loci for blood group O and tyrosine aminotransferase are on pig chromosome 6. J. Hered. 87, 272–276.

Marklund, S., Kijas, J., Rodriguez-Martinez, H., Rönnstrand, L., Funa, K., Moller, M., Lange, D., Edfors-Lilja, I., Andersson, L., 1998. Molecular Basis for the Dominant White Phenotype in the Domestic Pig. Genome Res. 8, 826.

Martin, S.H., Davey, J.W., Jiggins, C.D., 2015. Evaluating the use of ABBA-BABA statistics to locate introgressed loci. Mol. Biol. Evol. 32, 244–57. doi:10.1093/molbev/msu269

Megens, H.J., Crooijmans, R.P.M.A., San Cristobal, M., Hui, X., Li, N., Groenen, M.A.M., 2008. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. Genet. Sel. Evol 40, 103–128. doi:10.1051/gse:2007039

Mikawa, S., Morozumi, T., Shimanuki, S.-I., Hayashi, T., Uenishi, H., Domukai, M., Okumura, N., Awata, T., 2007. Fine mapping of a swine quantitative trait locus for number of vertebrae and analysis of an orphan nuclear receptor, germ cell nuclear factor (NR6A1). Genome Res. 17, 586–593. doi:10.1101/gr.6085507

Pielberg, G., Olsson, C., Syvänen, A.C., Andersson, L., 2002. Unexpectedly high allelic diversity at the KIT locus causing dominant white color in the domestic pig. Genetics 160, 305–11.

Porter, V. 1993. Pigs: A Handbooks to the Breeds of the World. Cornell University Press, USA.

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramos-Onsins, S.E., Burgos-Paz, W., Manunza, A., Amills, M., 2014. Mining the pig genome to investigate the domestication process. Heredity (Edinb). 113, 471–484. doi:10.1038/hdy.2014.68

Rubin, C.-J., Megens, H.-J., Barrio, A.M., Maqbool, K., Sayyab, S., Schwochow, D., Wang, C., Carlborg, O., Jern, P., Jorgensen, C.B., Archibald, A.L., Fredholm, M., Groenen, M.A.M., Andersson, L., 2012. Strong signatures of selection in the domestic pig genome. Proc. Natl. Acad. Sci. 109, 19529–19536. doi:10.1073/pnas.1217149109

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA.

Wiener, P., Wilkinson, S., 2011. Deciphering the genetic basis of animal domestication. Proc. R. Soc. B Biol. Sci. 278, 3161–3170. doi:10.1098/rspb.2011.1376

Yang, B., Cui, L., Perez-Enciso, M., Traspov, A., Crooijmans, R.P.M.A., Zinovieva, N., Schook, L.B., Archibald, A., Gatphayak, K., Knorr, C., Triantafyllidis, A., Alexandri, P., Semiadi, G., Hanotte, O., Dias, D., Dovč, P., Uimari, P., Iacolina, L., Scandura, M., Groenen, M.A.M., Huang, L., Megens, H.-J., 2017. Genome-wide SNP data unveils the globalization of domesticated pigs. Genet. Sel. Evol. 49, 71. doi:10.1186/s12711-017-0345-y

Yang, G., Ren, J., Zhang, Z., Huang, L., 2009. Genetic evidence for the introgression of Western *NR6A1* haplotype into Chinese Licha breed associated with increased vertebral number. Anim. Genet. 40, 247–250. doi:10.1111/j.1365-2052.2008.01820.x

Supplementary material

Supplementary Table 1: Information on the sampled animals of the domestic pig breeds.

Breeds	**Geographic region**	**Standard coat colour**	**Average thoracic vertebrae number**	**Average teats number**	**No. of farms**	**No. of animals (males and females)**	**Genotyping methods**[2]	**Years**[3]
Apulo-Calabrese	Central-South of Italy (Lazio, Basilicata and Calabria regions)	Solid black[1]	na[4]	10-16[5]	5	73 (20, 53)	PCR-RFLP, fragment analysis	2010-2015
Casertana	Central-South of Italy (Campania and Molise regions)	Solid black or dark grey[1]	na	10-16[5]	8	114 (30, 84)	PCR-RFLP, fragment analysis	2010-2013
Cinta Senese	Central Italy (Tuscany region)	Black with white belt[1]	14[6]	10-16[5]	10	80 (30, 50)	PCR-RFLP, fragment analysis	2010-2015
Mora Romagnola	North of Italy (Romagna region)	Dark red/black with paler abdomen[1]	na	10-16[5]	8	74 (20, 54)	PCR-RFLP, fragment analysis	2015
Nero Siciliano	Sicily island	Solid black[1] (a few animals could have white spots)	na	10-16[5]	5	70 (15, 55)	PCR-RFLP, fragment analysis	2010
Sarda	Sardinia island	No fixed coat colour[1]	na	8-16[5]	4	58 (25, 33)	PCR-RFLP, fragment analysis	2015
Krškopolje	Slovenia	Black with white belt of varying size and shape	14-15[7]	14-16[7]	6	36 (15, 21)	OpenArray ™	2015
Black Slavonian	East Croatia	Solid black	na	10-12[8]	6	30 (15, 15)	PCR-RFLP, OpenArray ™	2015
Turopolje	West Croatia	Grey/pale red	na	10-12[9]	2	45 (17, 28)	OpenArray ™	2015

Mangalitsa	Serbia	Grey/black	na	10-12[10]	4	46 (21, 25)	OpenArray ™	2015
Moravka	Serbia	Solid black	na	8-12[11]	7	47 (22, 25)	OpenArray ™	2015
East Balkan Swine	Bulgaria	Solid black (but not fixed)	14[12]	10-12[13]	2	44 (20, 24)	PCR-RFLP, fragment analysis	2015

[1] According to the standard described in the breed herd book.

[2] Genotyping protocols are described in Table S3.

[3] Years of collection of the analysed samples.

[4] na = not available.

[5] From: ANAS (2017) Norme tecniche del Libro Genealogico e Registro Anagrafico della specie suina. D.M. 20304, Ministero delle Politiche Agricole, Alimentari e Forestali.

[6] From: Franci, O., Campodoni, G., Bozzi, R., Pugliese, C., Acciaioli, A., Gandini, G., 2003. Productivity of Cinta Senese and Large White x Cinta Senese pigs reared outdoors in woodlands and indoors. 2. Slaughter and carcass traits. Ital. J. Anim. Sci. 2, 59-65. https://doi.org/10.4081/ijas.2003.59

[7] From: H2020 TREASURE project.

[8] From: Uremović, M., 2004. Povećanje gospodarske vrijednosti Crne slavonske pasmine svinja. Poglavlje u knjizi: Uremović Marija: Crna slavonska pasmina svinja: hrvatska izvorna pasmina, Vukovarsko-srijemska županija, Vukovar, 70-78.

[9] From: Robić Z., 2002. Contribution to the renewal of the Turopolje breed of pigs. Agron. Glas. 64, 305-20.

[10] From: Belic J., 1951. Specijalna zootehnika (ovčarstvo i svinjarstvo). Univerzitet u Beogradu, Naučna knjiga, Beograd, 1-376.

[11] From: Petrović, M., Mijatović, M., Radojković, D., Radović, Č., Marinkov, G., Stojanović, Lj., 2007. Genetic resources in pig breeding – Moravka. Biotech Anim Husbandry. 23, 1-11.

[12] From: Hlebarov, G., 1921. The East-Balkan Pig, Cooperative Printing Edison. Sofia (Bg).

[13] From: Marchev, J., Doneva, R.K., Dimitrova, D., 2017. East Balkan swine – autochthonous Bulgarian pig breed. Arch. Zootec. Proceedings IX Simposio Internacional sobre el Cerdo Mediterráneo, 61-65.

Supplementary Table 2: Information on the investigated wild boars.

Wild boar populations	No. of animals	Genotyping methods[4]	Years [5]
North of Italy	113 [1]	PCR-RFLP and fragment analysis	2010-2013
Sardinia	26 [2]	PCR-RFLP and fragment analysis	2014
South-East Europe	90 [3]	PCR-RFLP and fragment analysis	2010-2012

[1] 111 of these animals were investigated by Fontanesi et al. (2014).
[2] Analysed in this study.
[3] Already investigated by Fontanesi et al. (2014).
[4] Genotyping results reported by Fontanesi et al. (2014) did not distinguish alleles E^{D2} and E^{P} of the *MC1R* gene. All wild boars were re-genotyped in the current study to obtain information from all *MC1R* alleles (see Table S3 for details).
[5] Years of collection of the analysed samples. Information on the sex was not recorded for all animals and therefore was not reported in the table.

Supplementary Table 3: PCR-RFLP and fragment analysis protocols used for the genotyping of markers at the *MC1R* and *NR6A1* genes.

Primer pair names/genes	**Genotyping method**	**Primer sequences (5'-3'): Forward and *Reverse* primers**	**Amplified region (bp)**	**PCR conditions**[1]	**Genotyping protocols/system**
MC1R_1	PCR-RFLP	CTGCACTCGCCCATGTACTA *AGCAGAGGCTGGACACCAT*	196	61/3.0	Amplicons digested with *BspH*I (c.367G = 196 bp in E^{+}, E^{D1} and *e*; c.367A = 154 + 42 bp in E^{D2} and E^{P}) [2]
MC1R_2	PCR-RFLP	GCGGGTACTGTACGTCCACAT *CCCAGCAGAGGAGGAAGAC*	154	61/3.0	Amplicons digested with *Hha*I (c.727G = 108 + 46 bp in E^{+}, E^{D1}, E^{D2} and E^{P}; c.727A = 154 bp in *e*); Amplicons digested with *Bst*UI (c.729G = 109 + 47 bp in E^{+}, E^{D2} and E^{P}; c.729A = 154 bp in E^{D1} and *e*) [2]
MC1R_indel	Fragment analysis	CACCTCTGGGAGCCATGA *GTCTGGTTGGTCTGGTTG*	168/170	55/2.5	Amplicons analysed in a capillary sequencer (ABI3100 Avant, ABI Prism)
MC1R_OA	OpenArray™ Genotyping platform	-	-	-	QuantStudioTM 12 K Flex Real-Time PCR System (Thermo Fisher Scientific) [3]
NR6A1	PCR-RFLP	GGTATCCTGAGCACCCAGTC *ACCTGGAGGACAGTGTGGAG*	203	55/2.5	Amplicons digested with *Msp*I (g.299084751C = 180 + 23 bp; g.299084751T = 203 bp) [2]
NR6A1_OA	OpenArray™ Genotyping platform	-	-	-	QuantStudioTM 12 K Flex Real-Time PCR System (Thermo Fisher Scientific) [3]

[1] Annealing temperature (°C) / [$MgCl_2$] mM.
[2] Genotyping protocols were based on PCR-RFLP. Restriction enzymes used to genotype the amplified fragments are indicated together with the size of the fragments obtained after digestions and extension alleles that have the indicated PCR-RFLP patterns (Fontanesi et al., 2010, 2014).
[3] Genotyping details are reported in Muñoz et al. (submitted).

Supplementary Table 4: *MC1R* and *NR6A1* genotyping data in the autochthonous pig breeds.

Breeds	No. of pigs	***MC1R* genotypes[1]**																***NR6A1* genotypes[1]**			
		E^+/E^+	E^+/E^{D1}	E^{D1}/E^{D1}	E^+/E^{D2}	E^{D1}/E^{D2}	E^{D2}/E^{D2}	E^+/E^P	E^{D1}/E^P	E^{D2}/E^P	E^P/E^P	E^+/e	E^{D1}/e	E^{D2}/e	E^P/e	e/e	HWE[2]	C/C	C/T	T/T	HWE[2]
Apulo-Calabrese	73	1	0	0	5	0	53	1	0	3	0	0	0	9	0	0	0.82	2	29	42	0.22
Casertana	114	0	0	0	38	0	51	2	0	19	0	0	0	3	1	0	0.08	5	4	104	0
Cinta Senese	80	0	0	1	1	0	54	0	0	20	0	0	0	4	0	0	0.12	0	0	80	-
Mora Romagnola	74	2	0	0	0	0	0	0	0	0	0	22	0	0	0	50	1	0	0	74	-
Nero Siciliano	70	4	0	1	10	6	23	6	2	18	0	0	0	0	0	0	0.24	9	20	41	0.02
Sarda	58	0	0	2	35	3	14	0	0	0	0	0	0	1	0	3	0	0	1	57	0.93
Krškopolje	31	0	0	0	0	0	20	0	0	10	0	0	0	1	0	0	0.99	0	0	31	-
Black Slavonian	27	0	0	22	0	1	0	0	2	0	0	0	1	1	0	0	0.91	0	4	23	0.57
Turopolje	47	0	0	0	0	0	0	0	0	0	47	0	0	0	0	0	-	6	27	13	0.17
Mangalitsa	47	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	0	2	45	0.83
Moravka	47	0	2	5	1	14	9	0	9	7	0	0	0	0	0	0	0.79	0	0	47	-
East Balkan Swine	44	4	4	3	12	5	11	0	0	4	0	1	0	0	0	0	0.39	2	13	29	0.73

[1] The number of animals for each genotype is reported
[2] P value of the Hardy Weinberg equilibrium calculated as reported in Supplementary Table 5.

Supplementary Table 5: *MC1R* and *NR6A1* genotyping data in the wild boar populations. South-Europe (SE) wild boars are also considered separately according to the country of origin.

Populations	No. of animals	*MC1R* genotypes[1]																*NR6A1* genotypes[1]			
		E^+/E^+	E^+/E^{D1}	E^{D1}/E^{D1}	E^+/E^{D2}	E^{D1}/E^{D2}	E^{D2}/E^{D2}	E^+/E^P	E^{D1}/E^P	E^{D2}/E^P	E^P/E^P	E^+/e	E^{D1}/e	E^{D2}/e	E^P/e	e/e	HWE[2]	C/C	C/T	T/T	HWE[2]
North of Italy	113	96	0	0	2	0	0	10	0	0	0	3	0	0	2	0	0.78	108	5	0	0.74
Sardinia	26	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	21	4	1	0.28
Bosnia and Herzegovina	16	15	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	15	1	0	0.86
Croatia	10	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	7	3	0	0.47
North of Macedonia	16	9	0	0	6	0	1	0	0	0	0	0	0	0	0	0	1	14	2	0	0.71
Montenegro	16	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	15	1	0	0.86
Serbia	17	9	0	0	7	0	0	1	0	0	0	0	0	0	0	0	0.99	15	1	1	0.51
Slovenia	15	9	0	0	3	0	0	3	0	0	0	0	0	0	0	0	1	13	2	0	0.71
SE Europe	90	68	0	0	17	0	1	4	0	0	0	0	0	0	0	0	1	79	10	1	0.38

[1] The number of animals for each genotype is reported.
[2] P value of the Hardy Weinberg equilibrium calculated with the HWE software program (Linkage Utility Programs, Rockefeller University, New York, NY).

Supplementary Table 6: *D* statistics of the ABBA-BABA test (Dasmahapatra *et al.* 2012; Martin *et al.* 2014) for *MC1R* and *NR6A1* loci.

Scenario	Positions of the populations in the scenario (P1-P4)[1,2]	*D* value for *MC1R*[3,4]	*D* value for *NR6A1*[4]
ABBA 1	P1: North Italian + Sardinian wild boars P2: SE European wild boars P3: All pig breeds P4: Outgroup (ancestral wild boar)	0.333	0.268
ABBA 2	P1: North Italian wild boars P2: Sardinian wild boars P3: Italian + Sardinian pig breeds P4: outgroup (ancestral wild boar)	1.000	0.715
ABBA 3	P1: Sardinian wild boars P2: North Italian wild boars P3: Italian pig breeds P4: Outgroup (ancestral wild boar)	1.000	-0.715
ABBA 4	P1: North Italian wild boars P2: Sardinian wild boars P3: Italian pig breeds P4: Outgroup (ancestral wild boar)	-1.000	0.715
ABBA 5	P1: North Italian + Sardinian wild boars P2: SE European wild boars P3: SE European pig breeds P4: Outgroup (ancestral wild boar)	0.333	0.268
BABA 1	P1: Italian + Sarda pig breeds P2: SE European pig breeds P3: All wild boar populations P4: outgroup (ancestral domestic pig)	0.335	0.184

[1] P4 is the ancestral population. The other positions (P1-P3) are those reported in the trees of Fig. S1 and indicate the direction of the introgression.

[2] Populations were grouped as follows: South-East (SE) European wild boars include wild boars sampled in Bosnia and Herzegovina, Croatia, North of Macedonia, Montenegro, Serbia and Slovenia; Italian pig breeds include Apulo-Calabrese, Casertana, Cinta Senese, Mora Romagnola and Nero Siciliano pig breeds; SE European pig breeds include Krškopolje, Black Slavonian, Turopolje, Mangalitsa, Moravka and East Balkan Swine.
[3] All domestic *MC1R* alleles were considered together in the formula for the calculation of D.
[4] Negative values indicate inverted gene flow between P2 and P1 (derived by allele frequencies in the Sarda pig breed in which the wild type alleles are absent or almost absent in the two loci).

Chapter 6

Genomics and nutrient sensing

Recent analyses of pig genomes have revealed important events in the history of swine demography and evolution. In a study examining the evolution and the mutation rate of pig genes, Groenen et al. (2012) identified pig-specific evolutionary breakpoint regions (EBRs) where a set of genes involved in taste perception lies, indicating that pig domestic taste phenotypes may have been affected by genomic rearrangements during pig evolution. Evidences of positive selection in taste perception genes (especially in bitter sensing) were detected through genome-wide scans and sequences comparison, across different mammalian species, in evolutionary analyses (Shi and Zhang, 2006; Kosiol et al., 2008; Li and Zhang, 2014). Rearrangements at these loci are considered responsible for pig adaptation to sensory perception of taste, like in the case of bitter compounds, for which pigs have developed tolerance, even at higher concentrations when compared to human sensing.

The evolution of chemosensing system in animals is not surprising considering that it played a determinant role in defining food preferences, helping animals in the discrimination of different foods and acting as a mean of defense from toxic compounds. This is the reason why animals are extremely averse to bitter-tasting compounds and attracted by sweet and umami ones (Zhao et al., 2003; Li and Zhang, 2014).

The interest towards pig taste sensory systems has started from the late 19th century, with the publication of works investigating the anatomy of taste receptors and simultaneously testing physiological responses and feed intake according to the nature of proposed foods (Roura and Fu, 2017). In this way, it was possible to gradually gather information about pig preferences and appetite behaviour and the connection between hormones and hunger-satiety cycle.

However, the possibility to use genomic tools in modern era, facilitated the study of molecular basis of these processes, progressively identifying genetic mutations in taste receptor genes and their association with pig growth, fat deposition and obesity (Clop et al., 2016; Ribani et al., 2016; Cirera et al., 2018).

This is particularly interesting because pigs are considered a valuable model for human studies: as they share anatomical and physiological features (Clouard et al., 2012) and some taste receptors and hormones, pigs are currently used for the understanding of the impact of nutritional interventions in human microbiota and in the mechanisms involved in appetite/satiety regulation (Roura et al., 2016).

As for other mammals, pig taste sensing is based on perception of five basic tastes: sweet, salty, sour, umami and bitter. Most of the studies on pig taste sensing are focused on two different families of receptors, called TAS1Rs and TAS2Rs, involved in sweet/umami and bitter taste sensing, respectively (Kiuchi et al., 2006; da Silva et al., 2014; Avau and Depoortere, 2015).

TAS2R is the largest family of taste receptors, and the number of bitter genes greatly varies in size among species (Li and Zhang, 2014), with pigs having the lowest number (14 identified, according to the last genome release) when compared to species like human or mouse. This confirms that bitter taste repertoire is closely related to the dietary habits of different species and adaptation to dietary change (Dong et al., 2009).

Considering that taste perception is strongly connected to the diet and the environment, comparing different pig breeds and detecting differences in taste receptor genetic sequences, could be informative in terms of how ecology impacts biodiversity and evolution of these genes (Shi and Zhang, 2006).

For this reason, we decided to investigate on SNPs distribution in both families of taste receptor genes across 19 European autochthonous pig breeds, comparing them with 3 commercial pig lines. We used a Pool-Seq approach to perform a SNP discovery study, with the final aim of identifying breed peculiarities that could be relevant in explaining the adaptation to their own environment, thus confirming selection pressure and the potential role of geographical origins in feeding preferences.

6.1 DNA pool-seq approach for the identification of single nucleotide polymorphisms in taste receptor genes of European autochthonous pig breeds.

Introduction

Taste sensing system consists of a network of nutrient sensory cells that are responsible for the recognition of chemical compounds according to 5 perceptual qualities: salty, sour, sweet, umami and bitter (Lindemann, 1996; Drayna, 2005, Chaudari and Roper, 2010). These sensory cells, located in different tracts and tissues of gastrointestinal system, play a central role in signaling, to the central nervous system, the 5 food sensing classes that are necessary for growth and survival of animals (Kiuchi et al., 2006). Taste receptors and their genes, expressed in sensory cells mainly located in buds of the tongue, are known to recognize the presence of dietary compounds in the oral cavity and

they are connected to physiological mechanisms controlling feed intake and eating behaviour in animals (Roura and Fu, 2017).

From a structural point of view, there are two families of taste receptors: those linked to salty and sour perception, that are related to ligand gated trans-membrane ion channels and those belonging to the family of G-protein coupled receptors (GPCRs) that are mainly involved in sweet, umami and bitter sensing. GPCRs group, indeed, can be in turn divided in two subfamilies, with subfamily 1 (known as TAS1Rs) responsible for sensing simple sugars and some L-amino acids coming from energy-rich diet, and subfamily 2 (known as TAS2Rs) recognizing selectively bitter and toxic compounds. G-protein coupled receptors are also involved in nutrient sensing of amino acids, peptones and fatty acids.

TAS1Rs group consists of 3 members that act by assembling in two heteromeric GPCR complexes, transducing tastants-induced stimulus via G-protein signaling cascades (Bachmanov and Beauchamp, 2007). More specifically, the heterodimeric complex derived by the combination of TAS1R2 and TAS1R3 protein products initiate the response towards sweet compounds, while the one made by the combination of *TAS1R1* and *TAS1R3* works as umami taste receptor.

TAS2Rs family instead, comprises approximately 10-40 functional genes in mammalian genomes (Ribani et al., 2016). According to the latest release of *S. scrofa* genome 11.1 on NCBI database, 14 genes belong to TAS2Rs family and they all act as bitter sensors to unpleasant compounds.

Considering that taste sensing and peripheral chemosensing can influence dietary choice and can impact on food/feed intake in pigs (Chaudhari and Roper, 2010) an increased interest has been recently recorded for understanding this relationship. During the last ten years, thanks to the advent of the genomic technologies and the increasing availability of new genomic tools for exploitation of biological data, some scientific works have been published on nutrient chemosensing in pigs (da Silva et al., 2014; Roura and Fu, 2017). Most of the studies have focused on investigating the variability of variants and polymorphisms in both families of taste receptor genes (Ribani et al., 2016; Cirera et al., 2018) A comprehensive understanding of how these variants impact of phenotypes of livestock species could be beneficial for the production system itself and helpful, in the case of animal models, for translational medicine (Roura et al., 2011; da Silva et al., 2014; Clop et al., 2016). Besides, the definition of how these gene variants relate to each other would be profitable to uncover functional mutations and eventually define association analysis for phenotype-genotype relationship. This is particularly stimulating when referring to the possibility of extracting relevant biological information and identifying variants that can explain pig diversity in breeds coming from different ecosystems and raised under local production systems (Roura and Fu, 2017).

Taking advantage from recent whole genome sequencing (WGS) technology and from the availability of the swine genome sequence and annotated genes, we decided to perform single nucleotide polymorphism (SNP) discovery on taste receptor genomic regions, using a DNA pool-seq approach in a total of 22 pools representing European autochthonous pig breeds and commercial pig lines. Pig populations under investigation are very low characterized from a genetic point of view, and there are few data about taste sensing in these breeds, so we decided to study taste receptor genes for defining the distribution of variants in these regions and identifying breed-specific mutations possibly involved in dietary adaptation.

Materials and methods

Biological samples and DNA extraction

A total of 770 pigs were included in this work. Pigs belonged to 19 different European autochthonous pig populations and 3 Italian commercial pig lines (Duroc, Landrace and Large White). European pig breeds involved in this work (and their countries of origin) were: Apulo-Calabrese, Casertana, Cinta Senese, Mora Romagnola, Nero Siciliano, Sarda (Italy), Basque, Gascon (France), Negre Mallorquí (Spain), Alentejano, Bìsaro (Portugal), Schwäbisch-Hällisches (Germany), Black Slavonian, Turopolje (Croatia), Lietuvos vietines, Senojo tipo Lietuvos Baltosios (Lithuania), Mangulica, Moravka (Serbia) and Krškopolje (Slovenia).

For each of the breeds, blood samples were collected for at least 60 individuals. Blood samples were then lyophilized, or frozen at -20°C, until the time of DNA extraction, that was performed using NucleoSpin®Tissue commercial kit (Macherey-Nagel, Düren, Germany). After the extraction, genomic DNA quality and quantity were assessed for each sample on Qubit® 2.0 Fluorometer (Thermo Fisher Scientific) before proceeding with pools and libraries preparation for sequencing.

Sequencing and reads alignment

DNA Pool-seq approach previously described by Fontanesi et al. (2015) was used to perform Whole Genome Sequencing (WGS) on 22 DNA pools, prepared for each one of the breeds under investigation. Pools were prepared with equimolar quantity of DNA derived from 35 animals for each breed that best suited quality and quantity requirements for library preparation. A total amount of 1.0µg DNA per sample (pool) was used as input material, then genomic DNA was randomly fragmented to a size of 350bp by Covaris cracker, and fragments generated were end polished, A-

tailed and ligated with the full-length adapter for Illumina paired-end sequencing. After PCR amplification and purification, libraries were analyzed for size distribution by Agilent2100 Bioanalyzer and quantified using real-time PCR. Finally, WGS was performed on Illumina HiSeq X Ten platform (insertion size of 350bp; 150bp-long reads) and raw data were collected and filtered removing adapters and low-quality reads. Fastq files were checked by using FASTQC v.0.11.7 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and the quality check highlighted the presence of very high-quality reads. Alignments were carried out with BWA-MEM 0.7.17 (Li and Durbin, 2009) on the latest assembly of *Sus scrofa* reference genome 11.1. Bam files were obtained with SAMtools v.1.7 (Li et al., 2009). Reads were subsequently de-duplicated by using Picard v.2.1.1 (https://broadinstitute.github.io/picard).

SNP calling and statistical data analysis

From the total amount of WGS data generated by whole genome sequencing, reads mapping on selected taste receptor genes were derived and used for SNP calling. A total of 17 taste receptor genes (3 for sweet and umami and 14 for bitter taste sensing) were targeted and investigated for SNP calling. Genomic coordinates of these targeted regions are listed in Table 1. For the TAS1R family, we investigated the three genes members *TAS1R1*, *TAS1R2*, *TAS1R3*, while we focused on *TAS2R1*, *TAS2R3*, *TAS2R4*, *TAS2R7*, *TAS2R9*, *TAS2R10*, *TAS2R16*, *TAS2R20*, *TAS2R38*, *TAS2R39*, *TAS2R40*, *TAS2R41*, *TAS2R42*, *TAS2R60* genes for bitter TAS2Rs family. All these genes are annotated in the official assembly and are located on 4 different pig chromosomes (SSC 5,6,16,18). Some of them were previously reported in other studies investigating variants in commercial international pig breeds and some local ones (da Silva et al., 2014; Clop et al., 2016; Ribani et al., 2016) but they have never been investigated in most of the European autochthonous pig breeds we used for this study. Single nucleotide polymorphisms were detected using the 'mpileup' function of SAMtools (Li et al., 2009). Then, called SNPs were filtered for a quality scores ≥20 and the obtained Variant Call Format (VCF) files with all the variants for each breed were submitted to Variant Effect Predictor (VEP) web interface available in Ensembl database (release 93), to analyze data and collect information about the effect of identified polymorphisms for each pool. Information derived from single VEP files were then merged using 'Intersect' function of bedtools v.2.27.0 (Quinlan and Hall, 2010) to identify unique breed mutations or to compare overlapping features among populations under investigation. SIFT (Sorting Intolerant From Tolerant) score was defined for all the missense mutations and allele frequencies were computed by counting, both for reference and alternative alleles for each SNP in all breeds.

SNP allele frequencies for alternative alleles were used to perform a Principal Component Analysis (PCA) using "prcomp" function in R software (R Development Core Team, 2008) and to calculate Euclidean distance among pig populations for generating cluster dendrograms of the breeds using "dist" and "hclust" functions of R.

Results

The average number of reads mapped on sweet and umami receptors regions was much higher than those mapped for bitter related genes (485.292 vs 40.705). Sequencing depth for each gene was calculated to further check the sequencing quality and the reliability of the data. The value was obtained dividing the number of filtered reads for each gene by the size of the genes as reported in the Sscrofa11.1 genome version. Average values of sequencing depth ranged from 41.0 (e.g. gene *TAS2R1*) to 43.9 (e.g. *TAS1R3*; see Supplementary Table 1 for details of all breeds).

SNP calling analysis of the 17 targeted taste receptor genes detected a total of 426 SNPs that are listed in Table 2. The majority of SNPs was found in TAS1R family that accounted for a total of 289 polymorphisms, a value that is much higher than the number of SNPs identified for TAS2R family genes for which only 137 mutations were found. SNP calling was performed to detect SNPs with different effects: SNPs were classified in synonymous, missense, intron, 5'UTR, 3'UTR and start-lost variants (Fig. 1). Among detected polymorphisms, 248 were located in intronic or untranslated regions (UTRs) and 178 were in coding sequences. These latter were divided in variants with low, moderate and high impact: 82 SNPs were synonymous, 95 SNPs were missense, and 1 SNP was a start-lost mutation (the only one that was found, located in position 3 out of 1128 of *TAS2R39* coding sequence). SIFT score was obtained for every missense mutation and 9 SNPs reported SIFT values ranging from 0 to 0.05 (with amino acid substitution predicted damaging, thus considered deleterious mutations). These variants were located in 4 different taste receptor genes, and two of them (located on taste receptor type 2 member 20-like, *TAS2R2*0) are still not recorded in dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/).

The distribution of total called SNPs (see Table 2) was extremely variable among the two families of taste receptors, with sweet and umami *TAS1R2* presenting the highest number of variants (mostly intronic) and *TAS2R4*, *TAS2R9*, *TAS2R10* and *TAS2R16* presenting the lowest number (only 5 SNPs detected for each).We calculated the SNP density for all the genes under investigation: values changed a lot (see Table 3), with a maximum of 0.016 (e.g. *TAS1R2*, *TAS2R39*) and a minimum of 0.001 (e.g. *TAS1R3*).

Considering breed distribution of SNPs, among the 22 analyzed pools, Schwabish-Hallisches and Mora Romagnola breeds reported respectively the highest (218) and the lowest (53) number of total SNPs for TAS1R genes (Fig. 2; Table 4 reporting breed by breed distributions of SNPs). A different trend was recorded for TAS2R genes (Fig. 3; Table 5 reporting breed by breed distributions of SNPs), for which Black Slavonian was the breed with the maximum number of total SNPs recorded (82) and Alentejano and Landrace the breeds with the minimum one (37).

Out of the 426 total SNPs found, only 218 were already recorded in dbSNP whereas almost half of them (208) were not previously identified. These new variants were divided in 17 synonymous, 9 missense mutations and 182 intron variants. They occurred in *TAS1R1*, *TAS1R2* and *TAS2R20* genes with this latter reporting most (200) of these novel SNPs.

Results showed that among the 426 detected polymorphisms, 64 might be described as breed specific because they were called only for a specific breed. These SNPs were missense (9) and synonymous (11) mutations and intronic variants (44). They occurred in 13 breeds, with high variability: Schwäbisch-Hällisches reported the highest number (16 SNPs) and Nero Siciliano, Bisaro, Lietuvos vietines, Moravka and Large White reported only one polymorphism each. The distribution across breeds and the genes these breed-specific SNPs belong to, are represented in Fig. 4.

Frequencies for reference and alternative alleles were calculated for all SNPs and for all the breeds in which they were found (data not shown). Results showed that 101 SNPs were found with alternative allele frequency value equal to 1 in at least one breed, suggesting that these alleles may be fixed in some of the pig populations under investigation. More in details, we found some breeds presenting frequency 1 for the mutated allele in a very high number of loci, like Mora Romagnola that reported 56 SNPs fixed for the alternative allele, and other breeds showing an opposite trend, like Krškopolje, that never reported any alternative allele with frequency 1. A brief summary of the distribution of these SNPs, considering breeds in which they are occurring and their impact, is reported in Table 6.

Allele frequencies data were calculated and used for obtaining the first two principal components (PC1 and PC2) of PCA plots. Frequencies were also used as distance matrix for calculating Euclidean distance (with 'dist' function) that was, in turn, used as a measure to find dissimilarities between clusters. By default, the 'hclust' function, follows the complete linkage clustering method that computes the largest value of dissimilarities between two elements in two clusters, as the distance between the two identified clusters.

For PCA analysis we decided to keep the two taste receptor gene categories separated and we computed four different PCA, two for each taste receptors family: the first PCA considered all the variants found for each category (289 SNPs for TAS1Rs and 137 for TAS2Rs) while the second considered only the exonic subsets (41 SNPs for TAS1Rs and 137 for TAS2Rs). The same settings for analysis were used for computing clusterization of pools.

Graphical outputs of statistical analysis are shown in figures 5 and 6. Fig. 5 is referred to sweet and umami taste receptor variants with dendrograms a) and b) representing the total amount of polymorphisms and the exonic only, respectively. The two PCA plots (data not shown) showed an almost overlapping distribution of variation in allele frequency across space, with some variables (and breeds) more correlated to each other. This is indeed reflected in dendrograms that in both cases cluster breeds in two main branches, with many sub-groups that are connected at different heights.

Fig. 6 reports the dendrogram obtained for taste receptor variants. In this case, the total number of variants corresponded to the total number of exonic polymorphisms. PCA results (data not shown) showed that there is a weak correlation among the considered variables, with pools highly separated in the plot. The dendrogram confirms this variability, with two main groups computed, but higher height of the fusion between branches and an independent branch for Black Slavonian breed.

Discussion

The recent development of genomic tools has allowed to better understand the molecular mechanisms involved in nutrient sensing. This is particularly interesting because taste perception is proved to influence eating behavior in animals and has a role in the adaptation of animal species to their ecological niches (da Silva et al., 2014). This is true overall for wild animal species or, in the case of pigs, also for autochthonous pig populations that are mostly ranged outdoor, and that consequently rely both on common farm diet but also on occasional food sources found in the environment. For this reason, local breeds, contrarily to commercial lines, whose diet is extremely controlled and nutritionally balanced, might have developed dietary mechanisms of adaptation in nutrient assumption (Roura and Fu, 2017).

Investigating for genetic variability in taste receptor genes, using biological information gathered from commercial and local pig breeds, is important for gaining genetic data about these features that in local pig populations are still uncharacterized and for discovering potential relationships between phenotype and genotype that might have physiological consequences.

Taking advantage of DNA Pool-Seq approach, we carried out a comparative study of taste receptors (both families of genes, TAS1Rs and TAS2Rs) genetic structure and we performed a SNP discovery to find analogies and differences among European local pig breeds with the possibility to compare results also with commercial pig lines data.

Pooling DNA samples is reported to be a cost-effective method for studies assessing variability and differentiation among populations (Ferretti et al., 2013; Schlötterer et al., 2014; Ribani et al., 2016; Taus et al., 2017). Main objective of DNA Pool-Seq is to identify the consistency of allele frequency differences in distinct populations and use this information to infer about selection (also in terms of evolutionary and demographic changes) in specific DNA regions (Axel et al., 2017).

Pooling strategy is considered powerful and informative, it reduces costs of sequencing and it requests a lower amount of DNA from each individual if compared to sequencing individual genomes (Anand et al., 2016). On the other hand, pool-seq derived data can be challenging to explore, for the identification of rare allele variants (because of sequencing errors) and for statistical analysis of sequence read counts in the estimation of correct genetic differentiation (Hivert et al., 2018).

We used DNA equimolar quantity of 35 animals per breed for each pool and we obtained an average sequencing depth of 40X. Reads were mapped on targeted genomic regions coding for 17 taste receptors. These genes were previously targeted for SNP discovery (da Silva et al., 2014; Clop et al., 2016; Ribani et al., 2016) with the exception of taste receptor type 2 member 20-like, *TAS2R*20, that was recently annotated as a novel pig gene in Ensembl release 93 (ENSSSCG00000038461) and in NCBI with the ID:100154902. In addition to that, to our knowledge most of the investigated breeds considered in this study were never analyzed before for SNP discovery in these genetic regions.

Our results showed that the number of variants occurring in the two families of receptors is different, with TAS1R family reporting a higher number of polymorphisms, that are mainly located in intronic regions of genes studied. A different situation was described for TAS2Rs genes that have a lower rate of mutation, but all variants occurred in coding sequences with a discrete number of them (9 out of 137) reporting deleterious SIFT values. A similar trend for bitter taste receptor variants was detected by da Silva et al. (2014) that reported that bitter taste genes had higher nucleotide diversity when compared to fatty acid or amino acid receptors.

Almost half of the SNPs found were still not annotated in dbSNP and can be thus considered novel variants that need to be further investigated and validated. Some of the 426 variants were detected in a single breed and could be defined breed-specific polymorphisms. These data need to be better

investigated for their effect, since there were also breed-specific missense mutations (e.g. Apulo-Calabrese Italian breed reported 5 private missense polymorphisms).

The maximum number of detected SNPs was found for Schwäbisch-Hällisches breed with 272 variants while the breed reporting the lowest number, was the Italian Duroc commercial pig breed with 103 SNPs identified. We detected very different values among breeds, with commercial lines (e.g. Italian Landrace and Italian Large White) reporting a higher number of polymorphisms in comparison to local breeds.

This finding contrasts the results of recent studies which assigned the role of immune sensors to the taste receptors. More specifically, considering that taste receptor genes have been recently defined as sentinels of defense against infections and their role in the regulation of airway epithelial innate immunity has been documented (Lee and Cohen, 2015) it could be possible to speculate that a high variability in their genetic sequences might increase the fitness of the populations. This would mean, in general, that these pig populations could be more resilient and healthier because they can maximize the adaptability and survival to challenging environments (Clop et al., 2016). However, this hypothesis should be demonstrated evaluating the effects of these polymorphisms on disease resistance or related traits.

This descriptive study regarding SNPs discovery in pig breeds wanted to be a comparison between animals coming from different ecosystems, including commercial and autochthonous populations. Taste sensory deputed-regions were investigated since nutrient sensing is recently considered a major evolutionary system under selection pressure resulting in dietary adaptation (Roura and Fu, 2017).

Obtained data need to be further processed to investigate the evolutionary relationships among TAS1R and TAS2R genes in pigs coming from different geographical areas with different evolutionary origin and history.

Conclusion

To date, this is the first SNPs discovery study performed on genomic data coming from a very large number of European autochthonous pig breeds and comparing their genetic variability in taste receptor genes.

The results obtained for the investigated taste receptor genes updated the catalogue of variability in these chromosome regions of the *Sus scrofa* genome. Other studies are needed to understand if these

variants have any phenotypic effects. Experiments should be designed recording nutritional preferences and feeding behavior of the animals in addition to classical performance traits.

Remarks

The information reported in this chapter are currently used for preparing a full paper to be submitted. For this reason, some data are not reported/shown.

Acknowledgements

This study has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 634476 (project acronym TREASURE).
This work was realized with the collaboration of different partners and we would like to thank for their contribution, the Italian Pig Breeders Association (Associazione Nazionale Allevatori Suini – ANAS), the Department of Bio and Health Informatics, Technical University of Denmark (Kgs. Lyngby, Denmark) and all TREASURE project partners.

Table 1: Taste receptor genes investigated in this study: information regarding gene family, gene name and identifier on Ensembl and coordinates on *Sus scrofa* 11.1 reference genome.

Gene family	Gene	Ensembl ID	Gene coordinates (bp)
Sweet and umami	*TAS1R1*	ENSSSCG00000003382	SSC 6: 67,397,367-67,409,973
	TAS1R2	ENSSSCG00000039973	SSC 6: 77,350,457-77,364,843
	TAS1R3	ENSSSCG00000003341	SSC 6: 63,610,525-63,617,831
Bitter	*TAS2R1*	ENSSSCG00000028554	SSC 16: 72,720,854-72,721,753
	TAS2R3	ENSSSCG00000028394	SSC 18: 8,141,030-8,141,980
	TAS2R4	ENSSSCG00000021525	SSC 18: 8,134,381-8,135,271
	TAS2R7	ENSSSCG00000035471	SSC 5: 61,262,331-61,263,269
	TAS2R9	ENSSSCG00000000631	SSC 5: 61,253,928-61,254,863
	TAS2R10	ENSSSCG00000038925	SSC 5: 61,242,636-61,243,565
	TAS2R16	ENSSSCG00000037368	SSC 18: 24,286,430-24,287,332
	TAS2R20	ENSSSCG00000038461	SSC 5: 61,181,192-61,182,244
	TAS2R38	ENSSSCG00000031100	SSC 18: 7,982,837-7,983,843
	TAS2R39	ENSSSCG00000021954	SSC 18: 7,068,087-7,069,214
	TAS2R40	ENSSSCG00000016467	SSC 18: 7,025,731-7,026,700
	TAS2R41	ENSSSCG00000016457	SSC 18: 6,780,808-6,781,731
	TAS2R42	ENSSSCG00000032322	SSC 5: 61,144,284-61,145,234
	TAS2R60	ENSSSCG00000016458	SSC 18: 6,807,649-6,808,593

Table 2: Summary of polymorphisms detected in taste receptor genes under investigation. SNPs were classified according to their low, moderate or high impact when they were found in coding regions, and in modifiers when they occurred in non-coding ones.

			Impact						
Gene family	**Gene**	**No of SNPs**	**Low**	**Moderate**	**High**	**Modifier**			
			Synonymous	Missense	Start Lost	Intron	3'UTR	5'UTR
Sweet and umami	*TAS1R1*	51	3	5		43		
	TAS1R2	229	22	6		201		
	TAS1R3	9	5			1	2	1
		Total: 289						
Bitter	*TAS2R1*	18	6	12				
	TAS2R3	6	1	5				
	TAS2R4	5	3	2				
	TAS2R7	7	2	5				
	TAS2R9	5	1	4				
	TAS2R10	5	2	3				
	TAS2R16	5	3	2				
	TAS2R20	17	5	12				
	TAS2R38	14	7	7				
	TAS2R39	18	4	13	1			
	TAS2R40	8	2	6				
	TAS2R41	6	2	4				
	TAS2R42	14	7	7				
	TAS2R60	9	7	2				
		Total: 137						

Table 3: SNPs density calculation for genes under investigation.

Gene family	**Gene**	**Gene lenght (bp)**	**No. SNPs detected**	**SNP density**
Sweet and umami	*TAS1R1*	12.606	51	0.004
	TAS1R2	14.386	229	0.016
	TAS1R3	7.306	9	0.001
Bitter	*TAS2R1*	899	18	0.020
	TAS2R3	950	6	0.006
	TAS2R4	890	5	0.005
	TAS2R7	938	7	0.007
	TAS2R9	935	5	0.005
	TAS2R10	929	5	0.005
	TAS2R16	902	5	0.005
	TAS2R20	1.052	17	0.016
	TAS2R38	1.006	14	0.014
	TAS2R39	1.127	18	0.016

	TAS2R40	969	8	0.008
	TAS2R41	923	6	0.006
	TAS2R42	950	14	0.014
	TAS2R60	944	9	0.009

Table 4: TAS1Rs SNPs distribution among investigated breeds and the total amount of SNPs found for each population

Breed	Gene			**Total**
	TAS1R1	***TAS1R2***	***TAS1R3***	
Cinta Senese	20	32	6	58
Alentejano	21	47	6	74
Apulo Calabrese	19	57	9	85
Mangulica	18	104	6	128
Black Slavonian	16	166	6	188
Casertana	19	145	6	170
Mora Romagnola	18	29	6	53
Nero Siciliano	20	47	7	74
Krskopolje	24	79	2	105
Duroc	18	30	6	54
Negre Mallorquì	19	41	6	66
Gascon	20	126	6	152
Basque	18	39	6	63
Bisaro	19	53	6	78
Lietuvos Baltosios	18	44	3	65
Lietuvos Vietines	18	41	4	63
Turopolje	20	32	6	58
Schwabish-Hallisches	40	172	6	218
Moravka	21	178	6	205
Sarda	20	104	6	130
Large White	19	92	5	116
Landrace	19	149	6	174

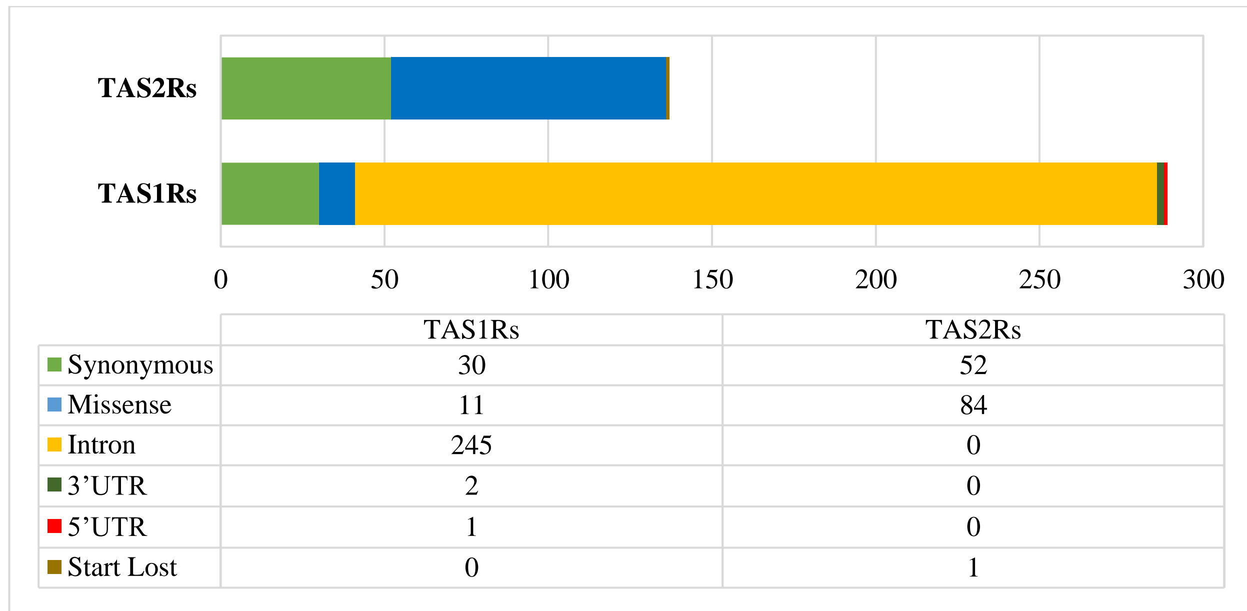
Table 5: TAS2Rs SNPs distribution among investigated breeds and the total amount of SNPs found for each population**.**

Breed	Gene														Total
	TAS2R1	***TAS2R3***	***TAS2R4***	***TAS2R7***	***TAS2R9***	***TAS2R10***	***TAS2R16***	***TAS2R20***	***TAS2R38***	***TAS2R39***	***TAS2R40***	***TAS2R41***	***TAS2R42***	***TAS2R60***	
Cinta Senese	9	2	2	2	1	0	2	9	9	5	5	0	5	0	51
Alentejano	0	0	0	4	1	0	0	10	12	0	0	0	10	0	37
Apulo Calabrese	8	2	3	5	1	4	4	17	13	6	0	3	10	1	77
Mangulica	10	2	0	4	1	1	1	16	7	5	4	0	12	0	63
Black Slavonian	13	2	1	5	3	1	3	11	7	14	7	2	12	1	82
Casertana	8	4	1	5	4	0	4	10	9	7	5	0	11	1	69
Mora Romagnola	10	5	3	4	2	1	0	2	12	13	4	0	9	3	68
Nero Siciliano	10	2	0	4	2	0	2	10	9	5	0	3	10	2	59
Krskopolje	13	2	0	3	1	0	4	6	6	7	3	4	5	1	55
Duroc	12	6	3	0	1	0	0	1	11	9	4	2	0	0	49
Negre Mallorquì	10	2	0	3	1	0	2	10	7	5	0	0	10	0	50
Gascon	8	4	3	5	4	0	1	12	10	5	0	2	11	6	71
Basque	16	5	4	4	2	0	1	10	12	5	0	2	10	1	72
Bisaro	9	2	2	3	2	0	2	11	9	8	1	5	11	5	70
Lietuvos Baltosios	9	2	3	3	3	1	4	4	7	9	3	3	8	7	66
Lietuvos Vietines	11	4	4	1	1	0	3	2	11	10	3	4	1	4	59
Turopolje	10	2	0	4	1	0	0	15	7	5	0	4	12	1	61
Schwabish-Hallisches	12	2	0	3	2	1	4	2	9	8	0	4	5	2	54
Moravka	10	4	2	1	1	0	1	5	9	12	3	3	3	1	55
Sarda	10	4	4	5	3	1	3	11	8	5	3	6	11	3	77
Large White	8	5	4	2	2	0	3	2	6	13	5	4	8	3	65
Landrace	10	2	0	2	4	0	2	1	1	5	0	1	8	1	37

Table 6: SNPs with fixed alternative allele (frequency =1) distribution considering breeds and impact of polymorphisms.

Breed	No of SNPs	Missense	Synonymous	Intron variants	3'UTR
Cinta Senese	21	5	4	12	0
Alentejano	19	0	6	13	0
Apulo Calabrese	10	0	2	8	0
Mangulica	30	20	8	2	0
Black Slavonian	4	1	1	2	0
Casertana	6	1	1	4	0
Mora Romagnola	56	8	9	39	0
Nero Siciliano	8	6	1	1	0
Krskopolje	0	0	0	0	0
Duroc	9	3	3	3	0
Negre Mallorquì	26	9	4	13	0
Gascon	17	4	2	11	0
Basque	27	6	5	16	0
Bisaro	5	1	1	3	0
Lietuvos Baltosios	7	0	1	6	0
Lietuvos Vietines	18	11	1	6	0
Turopolje	30	9	5	14	2
Schwäbisch-Hällisches	10	7	2	1	0
Moravka	3	0	1	2	0
Sarda	1	0	1	0	0
Large White	12	8	2	2	0
Landrace	17	10	3	4	0

Figure 1: SNPs distribution in the two families of taste receptor genes analyzed.



	TAS1Rs	TAS2Rs
Synonymous	30	52
Missense	11	84
Intron	245	0
3’UTR	2	0
5’UTR	1	0
Start Lost	0	1

Figure 2: Schematic representation of SNPs distribution in TAS1R members in all the breeds investigated

TAS1Rs

0 50 100 150 200 250

Cinta Senese
Alentejano
Apulo Calabrese
Mangulica
Black Slavonian
Casertana
Mora Romagnola
Nero Siciliano
Krskopolje
Duroc
Negre Mallorquì
Gascon
Basque
Bisaro
Lietuvos Baltosios
Lietuvos Vietines
Turopolje
Schwabish-Hallisches
Moravka
Sarda
Large White
Landrace

TAS1R1 TAS1R2 TAS1R3

Figure 3: Schematic representation of SNPs distribution TAS2Rs members in all the breeds investigated

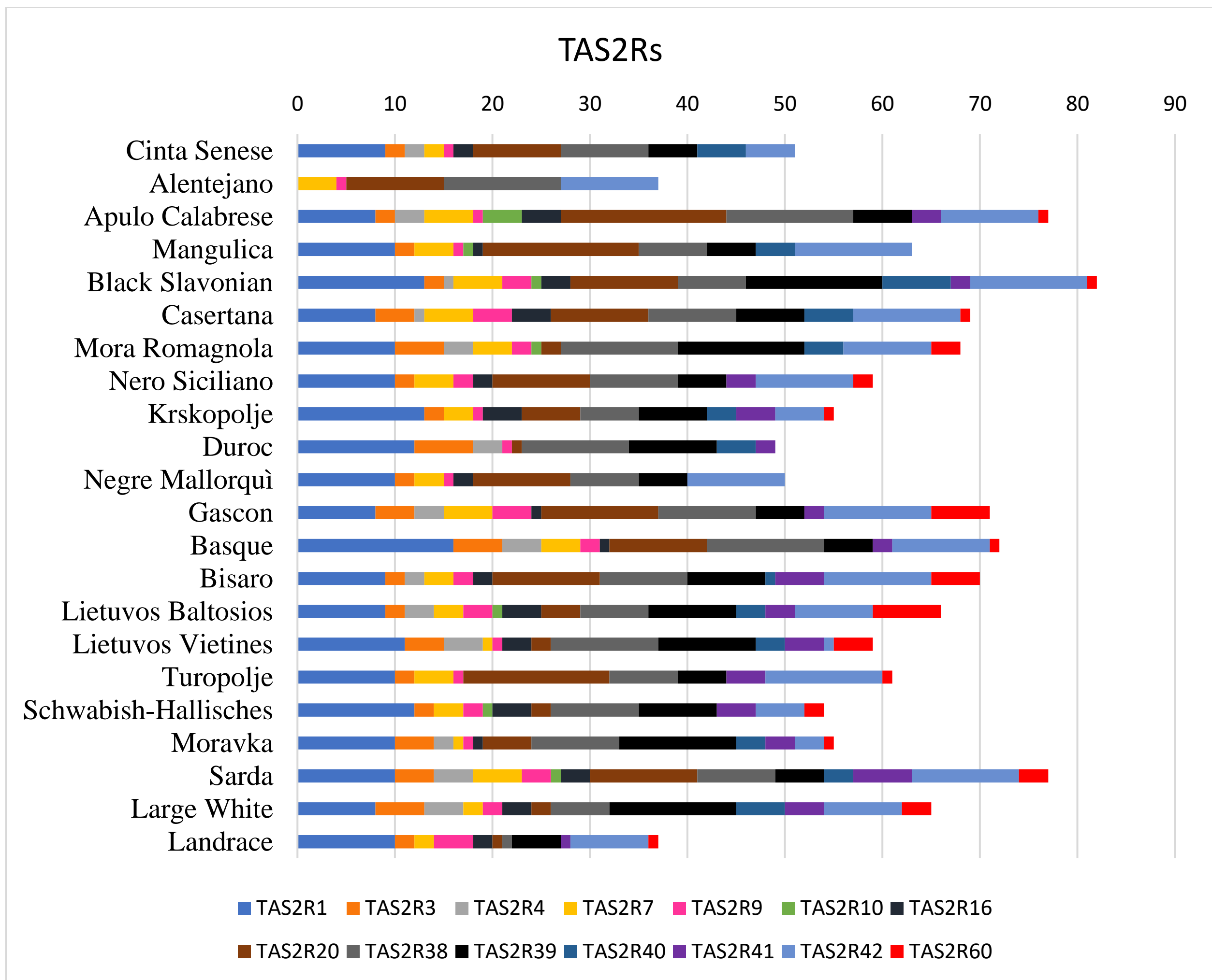


Figure 4: Representation of breed-specific polymorphisms detected across taste receptor genes.

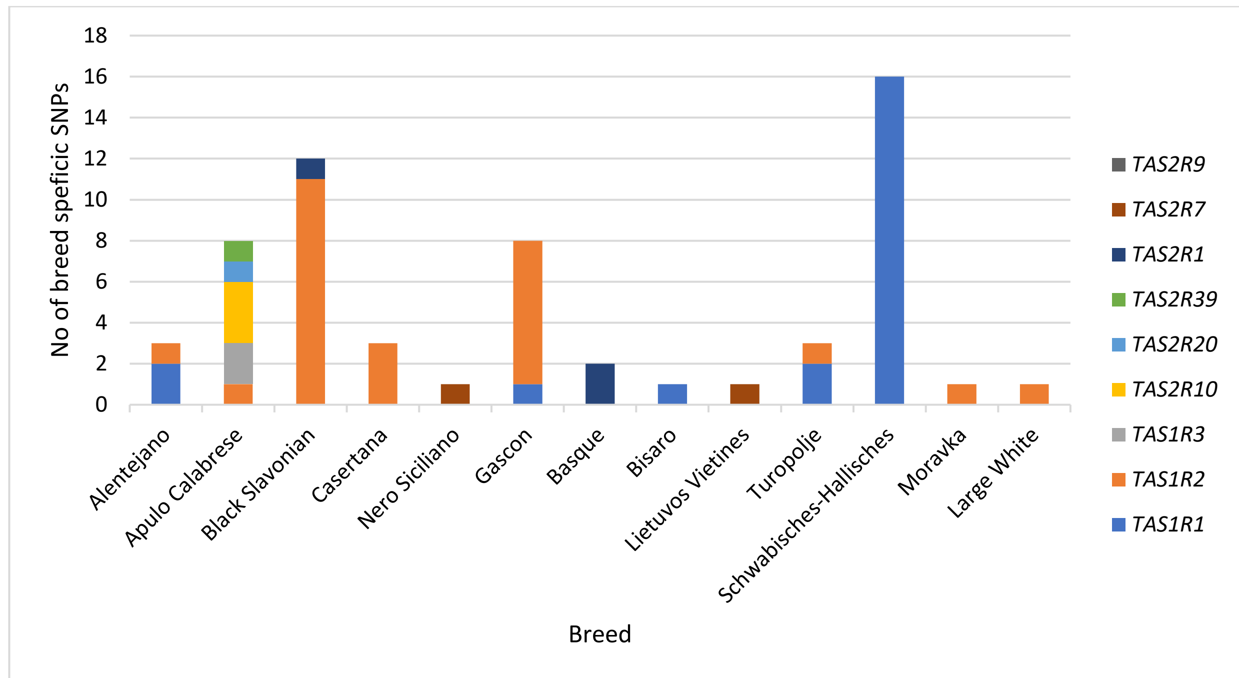
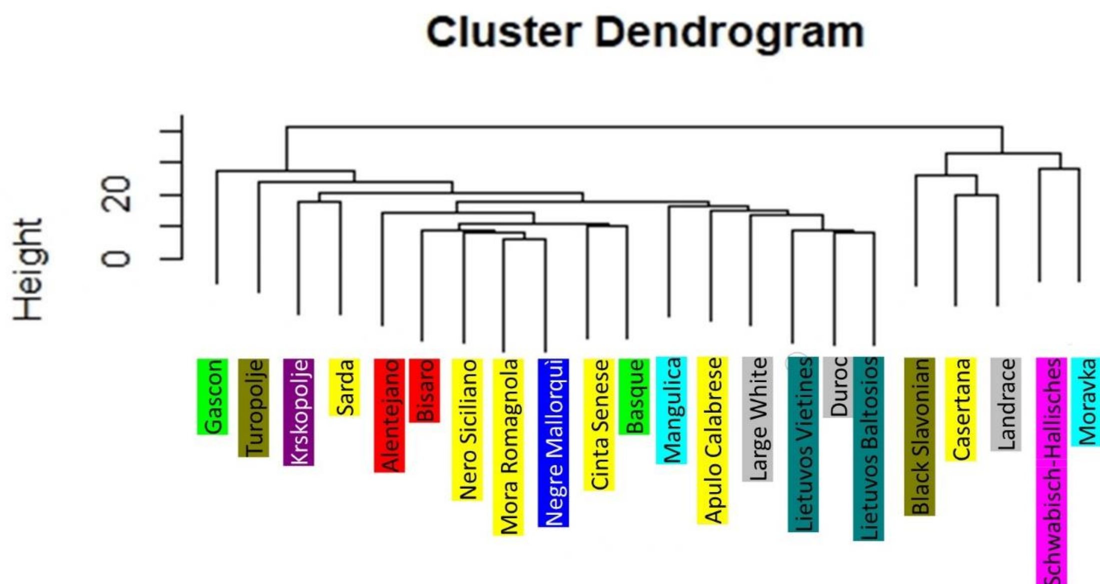


Figure 5: Sweet and umami taste receptor genes: a) cluster dendrogram computed considering all variants, b) cluster dendrogram computed using exonic variants only. In the picture, breeds highlighted with the same colour belong to the same country.

a)



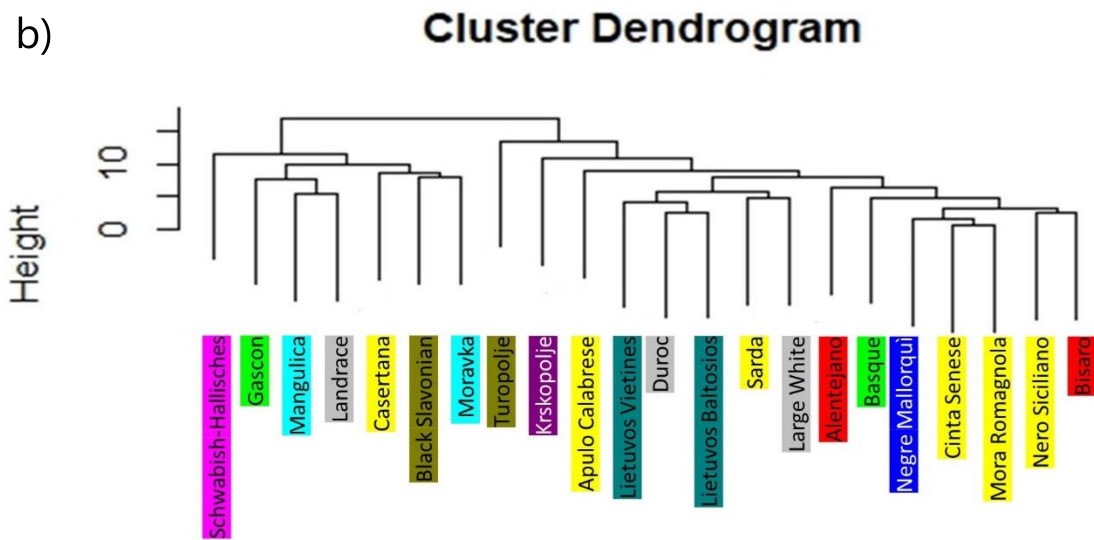
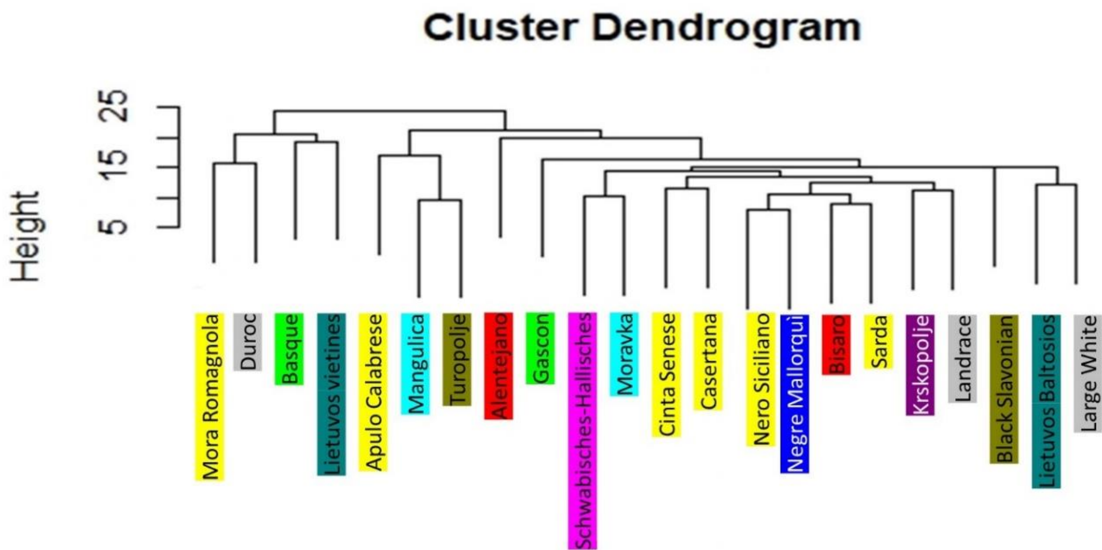


Figure 6: Bitter taste receptor genes cluster dendrogram, considering all the variants (that are, all exonic polymorphisms). In the picture, breeds highlighted with the same colour belong to the same country.



References

Axel, R., Wiberg, W., Gaggiotti, O.E., Morrissey, M.B., Ritchie, M.G., 2017. Identifying consistent allele frequency differences in studies of stratified populations. Methods Ecol Evol 8, 1899–1909. doi:10.1111/2041-210X.12810

Anand, S., Mangano, E., Barizzone, N., Bordoni, R., Sorosina, M., Clarelli, F., Corrado, L., Martinelli Boneschi, F., D'Alfonso, S., De Bellis, G., 2016. Next Generation Sequencing of Pooled Samples: Guideline for Variants' Filtering. Sci. Rep. 6, 33735. doi:10.1038/srep33735

Avau, B., Depoortere, I., 2016. The bitter truth about bitter taste receptors: beyond sensing bitter in the oral cavity. Acta Physiol. 216, 407–420. doi:10.1111/apha.12621

Bachmanov, A.A., Beauchamp, G.K., 2007. Taste receptor genes. Annu. Rev. Nutr. 27, 389–414. doi:10.1146/annurev.nutr.26.061505.111329

Chaudhari, N., Roper, S.D., 2010. The cell biology of taste. J. Cell Biol. 190, 285–296. doi:10.1083/jcb.201003144

Cirera, S., Clop, A., Jacobsen, M.J., Guerin, M., Lesnik, P., Jørgensen, C.B., Fredholm, M., Karlskov-Mortensen, P., 2018. A targeted genotyping approach enhances identification of variants in taste receptor and appetite/reward genes of potential functional importance for obesity-related porcine traits. Anim. Genet. 49, 110–118. doi:10.1111/age.12641

Clop, A., Sharaf, A., Castelló, A., Ramos-Onsins, S., Cirera, S., Mercadé, A., Derdak, S., Beltran, S., Huisman, A., Fredholm, M., van As, P., Sánchez, A., 2016. Identification of protein-damaging mutations in 10 swine taste receptors and 191 appetite-reward genes. BMC Genomics 17, 685. doi:10.1186/s12864-016-2972-z

Clouard, C., Meunier-Salaün, M.C., Val-Laillet, D., 2012. Food preferences and aversions in human health and nutrition: how can pigs help the biomedical research? animal 6, 118–136. doi:10.1017/S1751731111001315

da Silva, E.C., de Jager, N., Burgos-Paz, W., Reverter, A., Perez-Enciso, M., Roura, E., 2014. Characterization of the porcine nutrient and taste receptor gene repertoire in domestic and wild populations across the globe. BMC Genomics 15, 1057. doi:10.1186/1471-2164-15-1057

Dong, D., Jones, G., Zhang, S., 2009. Dynamic evolution of bitter taste receptor genes in vertebrates. BMC Evol. Biol. 9, 12. doi:10.1186/1471-2148-9-12

Drayna, D., 2005. Human Taste Genetics. Annu. Rev. Genomics Hum. Genet. 6, 217–235. doi:10.1146/annurev.genom.6.080604.162340

Ferretti, L., Ramos-Onsins, S.E., Pérez-Enciso, M., 2013. Population genomics from pool sequencing. Mol. Ecol. 22, 5561–5576. doi:10.1111/mec.12522

Fontanesi, L., Bertolini, F., Scotti, E., Schiavo, G., Colombo, M., Trevisi, P., Ribani, A., Buttazzoni, L., Russo, V., Dall'Olio, S., 2015. Next Generation Semiconductor Based-Sequencing of a Nutrigenetics Target Gene (GPR120) and Association with Growth Rate in Italian Large White Pigs. Anim. Biotechnol. 26, 92–97. doi:10.1080/10495398.2014.881369

Groenen, M.A.M., Archibald, A.L., Uenishi, H., Tuggle, C.K., Takeuchi, Y., Rothschild, M.F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H.-J., Li, S., Larkin, D.M., Kim, H., Frantz, L.A.F., Caccamo, M., Ahn, H., Aken, B.L., Anselmo, A., Anthon, C., Auvil, L., Badaoui, B., Beattie, C.W., Bendixen, C., Berman, D., Blecha, F., Blomberg, J., Bolund, L., Bosse, M., Botti, S., Bujie, Z., Bystrom, M., Capitanu, B., Carvalho-Silva, D., Chardon, P., Chen, C., Cheng, R., Choi, S.-H., Chow, W., Clark, R.C., Clee, C., Crooijmans, R.P.M.A., Dawson, H.D., Dehais, P., De Sapio, F., Dibbits, B., Drou, N., Du, Z.-Q., Eversole, K., Fadista, J., Fairley, S., Faraut, T., Faulkner, G.J., Fowler, K.E., Fredholm, M., Fritz, E., Gilbert, J.G.R., Giuffra, E., Gorodkin, J., Griffin, D.K., Harrow, J.L., Hayward, A., Howe, K., Hu, Z.-L., Humphray, S.J., Hunt, T., Hornshøj, H., Jeon, J.-T., Jern, P., Jones, M., Jurka, J., Kanamori, H., Kapetanovic, R., Kim, J., Kim, J.-H., Kim, K.-W., Kim, T.-H., Larson, G., Lee, K., Lee, K.-T., Leggett, R., Lewin, H.A., Li, Y., Liu, W., Loveland, J.E., Lu, Y., Lunney, J.K., Ma, J., Madsen, O., Mann, K., Matthews, L., McLaren, S., Morozumi, T., Murtaugh, M.P., Narayan, J., Truong Nguyen, D., Ni, P., Oh, S.-J., Onteru, S., Panitz, F., Park, E.-W., Park, H.-S., Pascal, G., Paudel, Y., Perez-Enciso, M., Ramirez-Gonzalez, R., Reecy, J.M., Rodriguez-Zas, S., Rohrer, G.A., Rund, L., Sang, Y., Schachtschneider, K., Schraiber, J.G., Schwartz, J., Scobie, L., Scott, C., Searle, S., Servin, B., Southey, B.R., Sperber, G., Stadler, P., Sweedler, J. V., Tafer, H., Thomsen, B., Wali, R., Wang, J., Wang, J., White, S., Xu, X., Yerle, M., Zhang, G., Zhang, J., Zhang, J., Zhao, S., Rogers, J., Churcher, C., Schook, L.B., 2012. Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491, 393–398. doi:10.1038/nature11622

Hivert, V., Leblois, R., Petit, E.J., Gautier, M., Vitalis, R., 2018. Measuring Genetic Differentiation from Pool-seq Data. Genetics 210, 315–330. doi:10.1534/genetics.118.300900

Kiuchi, S., Yamada, T., Kiyokawa, N., Saito, T., Fujimoto, J., Yasue, H., 2006. Genomic structure of swine taste receptor family 1 member 3, TAS1R3, and its expression in tissues. Cytogenet. Genome Res. 115, 51–61. doi:10.1159/000094801

Kosiol, C., Vinař, T., da Fonseca, R.R., Hubisz, M.J., Bustamante, C.D., Nielsen, R., Siepel, A., 2008. Patterns of Positive Selection in Six Mammalian Genomes. PLoS Genet. 4, e1000144. doi:10.1371/journal.pgen.1000144

Lee, R.J., Cohen, N.A., 2015. Taste receptors in innate immunity. Cell. Mol. Life Sci. 72, 217–236. doi:10.1007/s00018-014-1736-7

Li, D., Zhang, J., 2014. Diet Shapes the Evolution of the Vertebrate Bitter Taste Receptor Gene Repertoire. Mol. Biol. Evol. 31, 303–309. doi:10.1093/molbev/mst219

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–60. doi:10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. doi:10.1093/bioinformatics/btp352

Lindemann, B., 1996. Taste reception. Physiol. Rev. 76, 719–766. doi:10.1152/physrev.1996.76.3.719

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic

features. Bioinformatics 26, 841–842. doi:10.1093/bioinformatics/btq033

R Development Core Team (2008) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ribani, A., Bertolini, F., Schiavo, G., Scotti, E., Utzeri, V.J., Dall'Olio, S., Trevisi, P., Bosi, P., Fontanesi, L., 2016. Next generation semiconductor based sequencing of bitter taste receptor genes in different pig populations and association analysis using a selective DNA pool-seq approach. Anim. Genet. 48, 97–102. doi:10.1111/age.12472

Roura, E., Fu, M., 2017. Taste, nutrient sensing and feed intake in pigs (130 years of research: then, now and future). Anim. Feed Sci. Technol. 233, 3–12. doi:10.1016/J.AniFeedSci.2017.08.002

Roura, E., Koopmans, S.-J., Lallès, J.-P., Le Huerou-Luron, I., de Jager, N., Schuurman, T., Val-Laillet, D., 2016. Critical review evaluating the pig as a model for human nutritional physiology. Nutr. Res. Rev. 29, 60–90. doi:10.1017/S0954422416000020

Schlötterer, C., Tobler, R., Kofler, R., Nolte, V., 2014. Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. Nat. Rev. Genet. 15, 749–763. doi:10.1038/nrg3803

Shi, P., Zhang, J., 2006. Contrasting Modes of Evolution Between Vertebrate Sweet/Umami Receptor Genes and Bitter Receptor Genes. Mol. Biol. Evol. 23, 292–300. doi:10.1093/molbev/msj028

Taus, T., Futschik, A., Schlötterer, C., 2017. Quantifying Selection with Pool-Seq Time Series Data. Mol. Biol. Evol. 34, 3023–3034. doi:10.1093/molbev/msx225

Zhao, G.Q., Zhang, Y., Hoon, M.A., Chandrashekar, J., Erlenbach, I., Ryba, N.J.P., Zuker, C.S., 2003. The receptors for mammalian sweet and umami taste. Cell 115, 255–66.

Supplementary material

Supplementary Table 1: Sequencing depth calculated for targeted gene regions for each of the breeds. A total of 22 pools were sequenced and they were renamed from 1 to 22. Average number of reads and average gene sequencing depth are also reported.

Gene	Average reads number	Average sequencing depth	Pool name[1]																					
			1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**
TAS1R1	537600	42.6	41.3	43.3	45.2	40.9	40.4	46.1	42.7	38.2	42.9	41.2	41.6	41.4	39.6	45.6	41.4	43.1	44.9	42.6	42	41.5	46.7	44.8
TAS1R2	597052	41.5	39.7	42.9	46	41	38.7	42.9	39.2	37	40	39.7	41.7	39.9	35.5	42.2	41	43.1	45.9	42.8	43	40.7	44.8	44.3
TAS1R3	321226	43.9	41.2	44.6	50.7	43.2	42	47.5	50	47.5	41.7	42.3	44.8	40.8	38.2	43.4	40.5	40.6	44.6	46.1	45.2	38.3	46.2	47.1
TAS2R1	40705	41	41.2	42.5	46.7	40.4	43.8	46.5	32.6	42.2	46	39.4	46.4	37.6	34.3	40.8	32.9	37.8	38	43.3	40.4	41.9	44.5	42.4
TAS2R3	41093	43.2	44.8	43.1	42.5	43.3	43.6	43.9	38.8	42.9	46.4	49.5	38.8	43.2	38.4	45.4	44.2	47.4	43.7	35.6	37.4	47.8	44.5	44.7
TAS2R4	38433	43.1	38.9	45.2	51	44.4	45.2	41.9	44.9	42	41.8	44	38.5	40.5	38.8	42.5	43.4	35.3	47.3	46.1	45.2	40.9	47.1	43.4
TAS2R7	41284	43.9	41.4	42.1	39.3	44.2	45	47.5	45	36.2	43.7	47.4	43.1	44.7	44.2	43.7	43.7	44.6	40.7	46.6	44.4	49.4	41.5	48.1
TAS2R9	38848	41.5	48.7	39.5	46.6	33.6	42.5	42.2	43.1	34.1	35.8	39.1	45.3	40.4	36.5	40.6	42.1	37.1	41.4	45	43.6	48.4	41.3	45.3
TAS2R10	39283	42.2	45.1	41.2	44.4	44.8	45.2	49.4	40.1	43.1	39.7	34.4	47.1	44.5	37	39.4	38.8	45.3	36	42.2	45.2	43.2	39	43.5

TAS2R16	38064	42.1	42	45.5	46.6	39.3	45.6	45.8	41.5	31.4	38.8	40.1	36.1	47.7	45	42.3	40.9	46	43.7	39.4	41.1	41.8	41.2	44.6
TAS2R20	46218	43.9	46.9	44.7	49.1	50.6	44.4	45.1	38.1	37.3	36.8	40.7	42.6	38.1	42	45.3	46	40.1	43.9	52	40.5	52	44.1	44.6
TAS2R38	41868	41.6	40.3	38.8	44.2	41.4	39.2	40.9	41.6	32.2	46.7	37.5	42.1	40	40.6	39.9	37.5	38.9	46.9	40.6	48.4	39.6	49.6	47.2
TAS2R39	48020	42.6	42.1	40	45	45.6	43	43.6	35.8	34.5	42.1	41.7	43.9	40.7	40.8	46.6	40.8	44.9	41.2	48.2	49.6	43.3	42.1	40.6
TAS2R40	41615	42.9	41.4	42.2	47.7	40.6	43.4	37.4	38.6	39.6	39.1	40.3	39.5	47	38.5	44	41.6	50.1	41.5	47	38.1	45.9	46.7	52.6
TAS2R41	38616	41.8	40.7	43	47.3	45.9	35.5	39.9	46.3	36.8	38	38.5	40.2	37.2	33.4	47.5	41	41.4	47.7	42.7	41.5	43.8	48.1	42.4
TAS2R42	40429	42.5	46.7	38.9	44.3	42.6	40.1	44.2	35.2	34.7	41.4	37.6	44.8	41.2	41.6	41.7	40.4	43.5	40.7	41.2	40.3	51.5	55.4	46.4
TAS2R60	39186	41.4	42.6	46.6	47.7	37.3	38.3	43.8	31.4	36.2	48.2	34.8	36.2	37	37.8	41.2	41.6	44.3	42.7	42.7	45.6	43.1	44.6	47.8

1 Each pool corresponded to a breed: 1, Cinta Senese; 2, Alentejano; 3, Apulo Calabrese; 4, Mangulica; 5, Black Slavonian; 6, Casertana; 7, Mora Romagnola; 8, Nero Siciliano; 9, Krskopolje; 10, Duroc; 11, Negre Mallorquì; 12, Gascon; 13, Basque; 14, Bisaro; 15, Lietuvos Baltosios; 16, Lietuvos Vietines; 17, Turopolje; 18, Schwäbisch-Hällisches; 19, Moravka; 20, Sarda; 21, Large White; 22, Landrace.

General conclusions

This thesis provided some insights into the genetic architecture of European commercial and local pig populations related to several production and morphological traits that might important to define breed specific traits and economical relevant aspects (i.e. disease resistance vs production traits and feed preference). The potential practical applications of the obtained results derived from the molecular analysis of three main categories of genes: i) genes associated to genetic resistance to the most relevant porcine infectious diseases, ii) evolutionary-relevant genes associated with phenotypic changes over the pig domestication process and iii) genes coding for taste receptors responsible for pig nutrient sensing. Analyses were focused on these specific genes because each of the investigated topic they represent, strongly impact pig breeding, in terms of production system and management of genetic resources, including the impact of wild boar populations over domestic stocks.

Taken together, our results can be considered a basis for the use of genetic variability within and among pig populations. At the same time, the large amount of produced data (especially for taste receptor genes) represents a profitable source of information for comparative purposes and it opens the path to further research aiming to better describe the genetic potential of commercial and untapped pig breeds. Another outcome of these analyses could be the identification of breed specific genomic features for the development of DNA-based tools for traceability and authentication of mono-breed products which would be needed for a sustainable conservation of these genetic resources. Mining at the genome level the variability segregating in commercial and local pig populations could provide additional information to understand the genetic basis of complex and economically relevant traits.