Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Scienze e Tecnologie Agrarie Ambientali e Alimentari

Ciclo XXIX

Settore Concorsuale di afferenza: 07/G1 SCIENZE E TECNOLOGIE ANIMALI

Settore Scientifico disciplinare: AGR/17 ZOOTECNICA GENERALE E MIGLIORAMENTO GENETICO

TITOLO TESI

Analysis of animal genetic resources for the identification of polymorphisms associated with phenotypic features and evolutionary aspects

Presentata da: VALERIO JOE UTZERI

Coordinatore Dottorato

Prof. Giovanni Dinelli

Relatore

Prof. Luca Fontanesi

Esame finale anno 2017

ABSTRACT

Coat colour has been one of the main target of both natural and artificial selection and in particular one of the first of domestication process. This Thesis has been focused on the detection of genetic variability of new variants in coat colour genes in different species, especially those associated with particular coat colour phenotypes and introgression of domesticated alleles in wild populations. The aim of this Thesis was to identify and analyse polymorphisms associated with phenotypic traits that have been selected during the domestication processes and may differentiate breeds or populations in several livestock species, i.e. rabbit, donkey and pig, and evaluate the potential evolutionary effects on the wild counterparts, considering the specific example of the introgression between domestic pigs and wild boar populations. In particular, two main genes involved in melanin production (TYR and TYRP1) have been characterized in donkeys and rabbits respectively and variants in these genes have been significantly associated with albinism, suggesting that rabbits and donkeys could be considered as animal models for human albinism because these variants have never been detected in humans. Another aim of this Thesis, was the possibilty to use coat colour genes as introgression markers and the results suggest that accurate monitoring and management of rural stock populations, in order to avoid backcrossing events, are needed both for the conservation of the ecologic natural equilibrium of wild environments and to make wild species less invasive for human activities. Obtained results have been also evaluated for applications that consider allele and genotype frequency differences for authentication purposes, in order to differentiate meat products originating from wild boars, considered an upper market niche, against pig meat products with a less economic value.

Contents

1 GENERAL INTRODUCTION

1.1 DOMESTICATION

Phenotypic traits of living organisms (both externally visible or not) can be explained by the sequence of their DNA and the interaction with environment (Huxley 1942).
Variability found among living organisms is founded on the differences in homologous genetic sequences between each other. These differences underlie the process of speciation and formation of ecotypes/subspecies in the wild, while if this process is human-driven it leads to animal breeds or plant varieties (Darwin 1868).
The force that has allowed the diversification of organism lineages known today is the evolution which even today continues to shape living organisms acting through "hereditary characters" described by Charles Darwin (Darwin 1859) and Alfred Russel Wallace (Wallace 1870) in the Law of Evolution; these characters were then identified as genes by Gregor Johann Mendel (Mendel 1866; 1869).

Starting from the discovery of the DNA double helix structure by Francis Harry Compton Crick, James Watson and Maurice Wilkins in 1953 and the publication of the "Central Dogma of Molecular Biology" by Francis Crick (Crick 1958; Crick 1970), technologies advantages in the last 30 years led to second and third generation DNA sequencing techniques, passing from the fundamental discovery of the PCR technique (by Kary Mullis and Michael Smith, Nobel prize for chemistry in 1993, Mullis *et al.* 1986). Nowadays, these biotechnological techniques allow to better understanding of DNA structure, functionality and the evolution of genomes of all organisms, from the simplest to the most complex ones.

In order to understand how the evolutionary forces shape genomes, considering both biotic and abiotic factors (for example food, environment, climate, sexual selection, and so on), allowing the adaptation of life forms in different contexts, domestication of plant species and animals can be considered the best model (Gupta 2004).

Domestication has been a fundamental process in human history but also in the domesticated species histories. This process in animals has often been confused with animal taming or training, which are two kind of relationships in which humans teach something to the animal with the "reward-punishment" method that activates conditioned but innatural reflexes. The result of these processes is the acquisition by the animal of some abilities not inheritable which the animal will

tend to repeat after a clear signal by humans. Domestication instead is the process that allowed humans to rear some animal species (or to cultivate some plant species) having the control on their reproduction, diet and breeding environment (Price 1999; Diamond 2002).

Humans began the domestication process choosing some particular specimens from a species with advantageous phenotypes useful for humans and they acted through artificial selection obtaining almost homogeneous groups of individuals with heritable characteristics and very different from the wild species of origin (Diamond 2002). The result of the human-driven selection, for some species very strong, induced in domesticated animals anatomical, behavioral, genetic modifications and, in some cases, physiological changes for which it is no more possible associate the domestics with the wild counterparts (Clutton-Brock 1981; Diamond 1996).

Although domestication is a recent phenomenon referring to evolution times, evident traces of it are present in domesticated plants and animals genomes (Diamond 2002; Bruford *et al.* 2003; Ross-Ibarra *et al.* 2007). Moreover, thanks to mtDNA (mitochondrial DNA) in animals and cpDNA (chloroplast DNA) in plants (both with maternal uniparental inheritance), it is possible to infer when mutation events occurred and to infer the timing of the important periods during the domestication process (Hasegawa *et al.* 1985; Rannala and Yang 2013). This is possible because these circular genomes that are located within particular organelles of eukaryotic cells and they evolve (change) with an almost constant rate and this is the reason why they can be used as molecular clocks.

However, domestication has not always been achieved in a unidirectional way by humans because in some cases has been a result of a mutualistic relationship (Zeder 2006; O'Connor 1997).

One example of domestication probably occurred thanks to a mutualistic relationship, and according to some authors it started following animal interests, is related to the first domesticated species, *Canis lupus*, and for which a lot of studies have been published (Vila *et al.* 1997; Axelsson *et al.* 2013; Thalmann *et al.* 2013; Freedman *et al.* 2014; Skoglund *et al.* 2015); but also for the cat it is hypothized a mutualism at the beginning of its domestication (Vigne *et al.* 2004; Driscoll *et al.* 2007; Hu *et al.* 2014; Montague *et al.* 2014).

Even if more information are needed to clarify the dynamics of domestication process, to date we know that domesticastion began at least 15000 years ago with *Canis lupus* and several studies suggest that there are more than one domestication centers in different geographic areas (Vila *et al.* 1997; Skoglund *et al.* 2015; Frantz *et al.* 2016).

In particular several authors suggest mainly Western Europe and East Asia as putative domestication centers and molecular analysis and archeological remains indicate that the beginnig of dog domestication could be dated from 19000 to 32000 years ago (Thalmann *et al.* 2013; Frantz *et al.* 2016; Axelsson *et al.* 2013) and that the domestication started from different wolves popuplations.

During this process, back-crossing events should not be excluded as well as the Western Europe populations replacement by East Asian dogs following human migrations (Axelsson *et al.* 2013; Frantz *et al.* 2016). The result of dog domestication is the extreme phenotipic variation among breeds for many many traits. In particular the severe selection by humans led changes in genes involved in central nervous system development, diet (bitter taste perception) and metabolic pathways related to starch, fatty acids, glucose absorption. Finally, as it is nowadays visible, changes occurred in several genes affecting coat colour (Axelsson *et al.* 2013; Schmutz *et al.* 2002).

In general domestication led to a shift in evolution, ecology, population dynamics of both *H. sapiens* and the domesticated species (crops and livestock) (MacHugh *et al.* 2017; Purugganan and Fuller 2009; Fuller *et al.* 2011); domestication occurred in coincidence with the so called Neolithic Transition about 10000 years ago in which humans became farmers and herders: in this period goats, sheeps, cattles and pigs were domesticated as first livestock in the Near East (South-West Asia) (Asouti *et al.* 2013; Connolly *et al.* 2011). In South Asia, in Indian Peninsula, about 8000 years ago humped cattle (*Bos indicus*) was domesticated while the water-buffalo (*Bubalus bubalis*) about 4500 years ago. Other important animals which were domesticated later were chicken (*Gallus gallus*) in South East Asia about 4000 years ago (Peters *et al.* 2016), horse 5500 years ago in Central Asia, donkey in Egypt about 5000 years ago and dromedarius about 3000 years ago in Arabian Peninsula (Larson *et al.* 2014; Furrel 2006; Patel and Meadow 1998; Larson and Burger 2013; Outram *et al.* 2009; Beja-Pereira *et al.* 2004; Rossel *et al.* 2008).

Domestication had great impacts on livestock morphology, physiology, metabolism and reproduction. Among mammalian domestic species several traits underwent to genetic and phenotypic changes such as coat colour variations, depigmentation, behavioral modifications in favour of tameness and docility, craniofacial morphology, alteration of metabolism and endocrine system and finally one of the most important trait: changes in female oestrus cycle and reproduction (Dobney and Larson 2006; Jensen 2014; Wilkins *et al.* 2014).

Genomics can reveal the microevolutionary aspects occurred during the domestication process, investigating some genes influencing the phenotypic changes mentioned above (Albert *et al.* 2009; MacHugh *et al.* 2017; Rubin *et al.* 2012). Recent studies revealed that genes that underwent under

artificial selection were involved in brain development or neurotransmitter signaling (Albert *et al.* 2011; Saetre *et al.* 2004; Heyne *et al.* 2014), diet modifications and genes related to reproductive traits (Axelsson *et al.* 2013; Schubert *et al.* 2014; Rubin *et al.* 2012).

A relatively new branch of Genomics is Paleogenomics, analysing ancient DNA (aDNA) from archeological subfossils, is giving new perspectives in helping scientists on the understanding of the genetic and genomic signatures of domestication in modern animals; moreover, disciplines like phylogenetics and phylogeography are very informative to detect and identify the centers of domestication from which animal domestication started. In fact, the extant genetic diversity of modern domesticated animal populations can reflect the location of original domestication centers. Populations with strong phylogeographic structure such as cattle and European pigs, which show a higher genetic diversity in Near East areas (Larson *et al.* 2005; MacHugh *et al.* 1997), confirm the hypothesis of isolation-by-distance pattern sas the result of events of founder effect during the expansion of early domestic populations.

On the other hand, there are other modern domestic animals that do not show a strong phylogeographic structure, such as horses, dromedaries and in general pigs (Almathen *et al.* 2016; Larson *et al.* 2007; Achilli *et al.* 2012; Vila *et al.* 2001). In these cases, some other microevolutionary forces occurred and shuffled the genomics signatures of domestication.

In general, comparing wild population genomics against domesticated counterparts genomics can be useful to detect and identify the putative domestication centers even if it should be taken under consideration that contemporary wild populations are not the direct ancestors of domesticated animals (MacHugh *et al.* 2017). In pigs, events of local introgression from wild boars during the early phases of domestication and the subsequent gene flows among wild boars from Europe and pigs during human migrations, masked the signature of domestication in their genomes (Larson *et al.* 2007).

In dromedaries, the wild progenitors have gone extinct (Almathen *et al.* 2016) and in this case paleogenetics could help analysing the aDNA from subfossils in order to identify and date some genomics signature of domestication. Ancient DNAs, both ancient mtDNA and ancient genomic DNA (gDNA), can give different information regarding the microevolutionary history and the domestication process: the first one is a more stable marker because of its maternal inheritance (Hutchinson *et al.* 1974), while gDNA can add more information about admixture events eventually occurred in the past or even about the putative identification of ancient progenitor species (Der Sarkissian *et al.* 2015): for example, studying ancient gDNA Edwards and colleagues (Edwards *et al.* 2010) identified in aurochs the progenitors of modern cattle breeds. Thanks to the study of both

mtDNA and gDNA in modern and ancient horses, including the whole genome sequencing (WGS) of the living last truly wild Przewalski horse (*Equus przewalskii*), recent studies showed that the genetic composition of domesticated horses has been influenced not only by *E. przewalskii*, but also by a still unknown ancient horse population, now extinct (Der Sarkissian *et al.* 2015; Schubert *et al.* 2014).

In order to detect how many and which genes have been targeted by human-driven selection during animal domestication, mainly two kind of approaches have been used: candidate gene approaches and genome-wide scans of positive selection (MacHugh *et al.* 2017). As recently demonstrated by Ludwig and collaborators (Ludwig *et al.* 2015), artificial selective pressure has not been the same along the times of domestication but it has changed depending on human socio-cultural contexts suggesting that human in different times selected different traits and phenotypes in animals.

One of the most selected trait in domesticated animals was for sure coat colour and analysis on the sequences of genes related to coat colour in ancient horses showed positive selection coefficients (Outram *et al.* 2009; Axelsson *et al.* 2013).

One of the most recent animal domestication is the rabbit one which started about 1400 years ago in the South France by monks (Carneiro *et al.* 2014; Clutton-Brock 1999). As occurred in horse domestication, the first trait selected by monks in rabbit was coat colour phenotypes and the hypothesis is that some mutations occurred in genes involved in different coat colour have been positively selected at the beginning of the rabbit domestication. According to Carneiro and colleagues (2014), no gene loss occurred during the rabbit domestication, as already highlighted in pigs and chickens (Rubin *et al.* 2010; Wright *et al.* 2010; Rubin *et al.* 2012), but the amount of many mutations in genes involved in brain and neuronal development with small effects led to behavioural modifications in rabbits in order to get used to humans and the enviroment they offered them.

While candidate gene approaches are based on the comparison between ancient and modern genes associated with domestic traits that had been selected during domestication, the genome-wide approaches include the analysis of all the genomes with the aim of detecting loci undergone under both artificial and natural selection (Carneiro *et al.* 2014). For istance, in horses 125 genomic loci have been identified as targeted of positive selection (Schubert *et al.* 2014) including regions cointaining genes involved in locomotion, skeleton development, behaviour, neuronal growth and brain plasticity. Another example is given by the comparison between auroch genome and modern cattle genome: Park and colleagues (Park *et al.* 2015) found 106 loci under positive selection which include genes involved in neurobiology, muscle development, metabolism, growth and reproductive

traits. Among these genes, *DGAT1* gene, associated with the quantitative trait locus (QTL) related to lactation, showed a high positive selection coefficient.

Domestication process caused a lot of so called “costs” leading to some deleterious genetic variations in animal genomes (Marsden *et al.* 2016; Schubert *et al.* 2014). These negative consequences are probably due to repeated demographic bottlenecks occured for founder effects.

In conclusion, in order to better understand the microevolutionary forces occurred during animal domestication process in several and different points in time and space, the future perspective might be the comparison between genomics and paleogenomics, analysing ancient and modern genomes, and now it is more possible thanks to the new technological advances of the last decade.

1.2 COAT COLOUR PHENOTYPES

Generally speaking, modern domesticated animals show a very large coat-colouration variation, which is not found in any of their wild ancestors: while wild species (especially mammals) are usually uniform in coat colour phenotypes, domesticated species are highly variable in terms of colours and colour patterns (Cieslak *et al.* 2011). For these reasons, domesticated animals can be considered as models of a rapid evolution resulting from arificial selection (Darwin, 1868) guiding the understanding of evolution and biology (Lewin, 2009). Recent studies demonstrated that the selection for coat-colour phenotypes began at first phase of domesticastion, in which several bottlenecks occurred in domesticated animals leading, first of all, to a decrease of the genetic diversity and, secondly, to a fixation of some favoured allelic variants in the new human-driven environment (Cieslak *et al.* 2011).

The first evidence of signature of selection of particular coat colour phenotypes has been published by Ludwig and collaborators (Ludwig *et al.* 2009) who genotyped 8 mutations in 6 genes modulating coat colour variation in ancient Neolithic horses. In fact, two genes (*ASIP* and *MC1R*) regulating the amount of two kind of melanins (eumelanin and pheomelanin) in skin and hairs showed high levels of positive selection coefficients suggesting that during Neolithic Transition chestnut and black colours have been strongly selected in horses.

As described by an ongoing and fascinating experiment on the domestication of silver foxes, coat colour is one of the domestic traits which can be modified by a strong selection pressure over few generation (Trut *et al.* 2009).

To date, coat colour is still one of the most important phenotype affecting the breeding schemes of several livestock (Legault, 1998) and in many breeds only one type of colouration is allowed for a breed standard characteristics (Driscoll *et al.* 2009). In pets and livestock, several breeds' names are based on their coat colour (for example the cattles Brown Swiss, the pigs Large White, the sheeps Bluefaced Leicester, the goats Belgian Fawn, the rabbits Blue Vienna, the dogs Golden Retriever, the cat Havana Brown and so on).

According to our knowledge, more than 300 loci affecting coat colour phenotype have been detected and about 150 genes are involved in pigmentation in mammals (Montoliu *et al.* 2010); each gene allele or locus can be responsible of one or more coat colour traits and they might be involved in the melanins production pathway or in their distribution (Hubbard *et al.* 2010).

Coat colour traits in mammals can be divided into patterned and non-patterned phenotypes; the first category includes all the variation related to spotting, stripes, leopard, or particular patterns, while the second group is related to all the uniform colorations from white to dark pigmented (Cieslak *et al.* 2011). The biomolecules responsible for pigmentation are two melanins, eumelanin (dark brownish to black) and pheomelanin (yellowish to reddish), and the basic coat colouration is caused by the ratio of these two pigments, determining all the colours included in a range from black to red (Hubbard *et al.* 2010). On the other hand, white phenotypes are caused by the complete lack of both melanins and are caused by allelic variations in different genes; according to which gene is non-functional or distrupted, we call white phenotypes in different ways: leucism (*KIT* and *EDNRB* genes, uniform white or patches of white colour), albinism (*TYR* gene, totally white) and the progressive greying (*STX17* gene, gradually turning into white phenotype during ageing). While leucism is defined as the total lack of pigmented cells in some areas of the body (or even all the body), albinism is the impairment of the biochemical pathway of melanin production.

Melanogenesis is the metabolic pathway for melanin production which occurrs in melanocytes; these specific cells derive from the embryonic neural crest tissue from which take origin melanoblasts, their precursors. Melanoblasts, before differentiating in melanocytes, migrate to the final destinations: skin and hair follicles (or feathers in birds).

Melanoblast specification, migration and melanocyte differentiation depend on different biochemical pathways involving genes such as *EDNRB* (endothelin receptor type B), *EDN3* (endothelin 3), *KIT* (receptor tyrosine kinase), *KITL* (KIT ligand), *MITF* (microphtalmia-associated transcription factor) and others. Mutations in these genes can cause the absence of pigmentation leading to leucism, as already observed in *KIT* gene of horses (Brooks and Bailey, 2005; Haase *et al.* 2009), pigs (Giuffra *et al.* 1999; Pielberg *et al.* 2002) and cats (Reinsh *et al.* 1999). In some

cases, *KIT* variants are associated with other coat colour phenotypes like roan, a mixture of coloured and white hairs in horses (Marklund *et al.* 1999). Another important key regulator in pigmentation is *MITF* gene, which is resposible of the transcription of essential enzymes involved in melanogenesis; mutations in this gene have pleitropic effects including lack of pigmentation, deafness and other abnormalities (Wohlke and Distel, 2009; Minvielle *et al.* 2010).

Melanogenetic pathway occurrs in melanocytes, more precisely in melanosomes which are specialized lysosome-related organelles and contain specific components such as structural proteins, membrane transport channels and melanogenesis-related enzymes (Hearing 2011). The tyrosinase gene family, composed by *TYR* (tyrosinase), *TYRP1* (tyrosinase related protein 1) and *DCT* (dopachrome tautomerase), play the most important role in the melanin production pathways. The key enzyme of this process is tyrosinase, which catalyzes the reaction of hydrolyzation of L-tyrosine and producing Dopaquinone (DQ) after a spontaneous oxidation reaction. After that, DQ is converted in Dopachrome and subsequently in eumelanin by the tyrosinase related enzymes *DCT* and *TYRP1* (Hearing 2011). On the other hand, also pheomelanin production begins from L-tyrosine aminoacid but proceeds only if a cysteine residue is present combined with the absence of activated *TYR* (Sharma *et al.* 2002). Melanins are then deposited into the internal matrix of melanosomes and these organelles are then transported to the peripheral region of the melanocytes, thanks to a motor protein complex via microtubular networks (Cieslak *et al.* 2011).

Due to its key role in melanogenesis, disruptive mutations occurring in *TYR* gene result in a complete lack of pigmentation in skin and hairs (or feathers in birds) leading to a white coat colour phenotype defined as *TYR*-negative albinism. This phenotype in humans has been related to oculocutaneous albinism type 1 (OCA1) and several mutations in *TYR* gene have been associated with this disease (Sanabria *et al.* 2012). *TYR*-negative albinism has been described in many domesticated species such as cats (Imes *et al.* 2006), cattles (Schmutz *et al.* 2004), sheeps (Adalsteisson 1977), chickens (Chang *et al.* 2006), rabbits (Aigner *et al.* 2000) and recently even donkeys (Utzeri *et al.* 2016).

Other mutations with smaller effects on tyrosinase function have been associated with other coat colour phenotypes like Chinchilla and Himalayan colourations; while in Chinchilla the *TYR* activity is reduced (Lamoreux *et al.* 2001), in Himalayan animals *TYR* activity is regulated by temperature in which colder parts of the animal are coloured and the warmer parts (more than 25°C) are white because tyrosinase doesn't work (Lyons *et al.* 2005; Benkel *et al.* 2009).

The so called TYR-positive albinism is instead caused by distruptive mutations in *TYRP1* gene, leading to the inability to catalyze the last step of eumelanin production. This result in a typical

brown phenotype (known also as Havana phenotype) as already observed in dogs (Schmutz *et al.* 2002), cattles (Berryere *et al.* 2003), sheeps (Gratten *et al.* 2006), cats (Lyons *et al.* 2005) and rabbits (Utzeri *et al.* 2014).

As already mentioned above, the ratio between eumelanin and pheomelanin and their distribution determine the basic colour of colouration, producing several intermediate types from black to red. The ratio between the two melanins is controlled by the agouti signalling protein (*ASIP*) and melanocortin 1 receptor (*MC1R*) ligand-receptor complex (Hubbard *et al.* 2010; Ito and Wakamatsu, 2005). *MC1R* is located on the surface of the melanocytes and it is activated by hormones such as alpha melanocyte stimulating hormone (αMSH), beta melanocyte stimulating hormone (βMSH) and adenocorticotropic hormone (ACTH); *MC1R* activation induces a biochemical cascade signalling promoting eumelanin production, through *TYR* activation.

Disfunctional mutations in *MC1R* gene are associated with several coat colour variation, caused by a shift in favour of pheomelanin production, already observed in several domesticated species like rabbits (Fontanesi *et al.* 2006; 2010), pigs (Kijas *et al.* 1998; 2001), dogs (Schmutz *et al.* 2003), sheeps (Loehr *et al.* 2008) and other wild animals (Cieslak *et al.* 2011). These disfunctional alleles are generally recessive in mammals, and normally heterozygous animals show the wild type phenotype.

ASIP, instead, is an antagonist acting on *MC1R* receptor blocking the *MC1R* signalling pathway: this results in a decrease of eumalin production and a increase of pheomelanin. Mutations in *ASIP* can lead to a disequilibrium in the ratio between the two pigments: in particular, dominant alleles, causing a continue block of *MC1R* pathway, lead to yellow/reddish coat colours, while disruptive mutations (especially in homozigosity) are associated with dark colourations (Hiragaki *et al.* 2008; Royo *et al.* 2008).

In conclusion, even if animal domestication began for human requirements such as meat, milk and fur production or suitability for human activities, coat colour phenotypes have always been highly selected and rare animal colourations have always had high prestigious value, especially for an economical point of view (Icken *et al.* 2007) or in terms of biodiversity protection. The mosy likely hypothesis is that different coat colour phenotypes were a random by-product of the animal domestication process of small populations, resulted by genetic drift (bottlenecks), inbreeding and loss on natural selective constrains on coat colour (because animals were no more exposed to natural selection; Cieslak *et al.* 2011). Then, a strong human-mediated positive selection led to the fixation of the coat colour phenotypes observed today.

1.3 INTROGRESSION

Gene flows from a population (either species or subspecies) to another one through repeated interbreeding and backcrossing between the two populations is defined as introgression (Fredriksen 2016). Introgression or introgressive hybridization is a type of hybridization resulting in a complex mixture of parental genes, while simple hybridization is a more uniform mixture (Harrison 2014).

In domesticated animals introgression events can be bi-directional towards: from the domesticated populations to the wild relatives and viceversa. Introgressed wild alleles in livestock could lead to different problems in domestication studies (Meiri *et al.* 2013) but can also give some advantadges in terms of useful traits for rural breeds such as robustness, frugality, resilience, environmental and climate adversities resistance.

From an ecological point of view domestic alleles introgressed in wild gene pools could have crucial impacts on biodiversity and conservation of wilflife resources. Several studies investigated domestic allele introgression in different wild species such as dogs towards wolves (Godinho *et al.* 2011), cats towards wild cats (Fredriksen 2016), cattles towards yaks (Qi *et al.* 2010), reindeers towards their wild counterparts (Jepsen *et al*. 2002), minks towards wild minks (Kidd *et al*. 2009), sheeps towards wild sheeps (Feulner *et al.* 2013) and pigs towards wild boars (Goedbloed *et al*. 2013; Frantz *et al*. 2013).

Generally, in literature, introgression studies have been performed using neutral and uniparental genetic markers like microsatellites and mtDNA (Kikkawa *et al.* 2003; Kijas *et al*. 2001; Adams *et al.* 2003; Cieslak *et al.* 2010; Thulin *et al.* 2006) while to date, thanks to the improvements of both sequencing technologies and *in silico* analyses, it is possible compare deeply homologous nuclear regions of domesticated traits between ancient and extinct species and actual populations (Der Sarkissian *et al.* 2015; Schubert *et al.* 2014).

Among domesticated markers, genes related to coat colour phenotypes are some of the best candidates for this kind of studies because, in some cases, they have been target of the first steps of domestication process (Ludwig *et al*. 2009; Cieslak *et al.* 2011).

The detection of introgressed alleles in general can have several applications in the monitoring of wildlife stocks mainly in areas with frequent contacts between rural rearing livestock and wild relatives; contacts can occurr and can lead to human safety problems, zoonosis and agricultural issues; the gene flow from domesticated animals to their wild counterparts can also lead to

physiological modifications which can affect the ecological equilibrium of a species in the natural environment (Utzeri *et al*. See paragraph 2.4; Rhymer and Simberloff 1996; Allendorf *et al*. 2001).

Another interesting application could be the detection of food frauds and adulterations, especially for niche products or high quality foodstuffs (Fontanesi *et al*. 2014), which recently is an emerging issue due to its impact on health, economic and socio-cultural aspects.

1.4 REFERENCES

Achilli, A., Olivieri, A., Soares, P., Lancioni, H., Kashani, B. H., Perego, U. A., ... & Felicetti, M. (2012). Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. *Proceedings of the National Academy of Sciences*, *109*(7), 2449-2454.

Adalsteinsson, S. (1977). Albinism in Icelandic sheep. *The Journal of heredity*, *68*(6), 347-349.

Adams, J. R., Leonard, J. A., & Waits, L. P. (2003). Widespread occurrence of a domestic dog mitochondrial DNA haplotype in southeastern US coyotes. *Molecular Ecology*, 12(2), 541-546.

Aigner, B., Besenfelder, U., Müller, M., & Brem, G. (2000). Tyrosinase gene variants in different rabbit strains. *Mammalian Genome*, *11*(8), 700-702.

Albert, F. W., Carlborg, Ö., Plyusnina, I., Besnier, F., Hedwig, D., Lautenschläger, S., ... & Zeising, C. (2009). Genetic architecture of tameness in a rat model of animal domestication. *Genetics*, *182*(2), 541-554.

Albert, F. W., Hodges, E., Jensen, J. D., Besnier, F., Xuan, Z., Rooks, M., ... & Burbano, H. A. (2011). Targeted resequencing of a genomic region influencing tameness and aggression reveals multiple signals of positive selection. *Heredity*, *107*(3), 205-214.

Allendorf, F. W., Leary, R. F., Spruell, P., & Wenburg, J. K. (2001). The problems with hybrids: setting conservation guidelines. *Trends in ecology & evolution*, *16*(11), 613-622.

Almathen, F., Charruau, P., Mohandesan, E., Mwacharo, J. M., Orozco-terWengel, P., Pitt, D., ... & Magee, P. (2016). Ancient and modern DNA reveal dynamics of domestication and cross-continental dispersal of the dromedary. *Proceedings of the National Academy of Sciences*, 201519508.

Asouti, E., Fuller, D. Q., Barker, G., Finlayson, B., Matthews, R., Fazeli Nashli, H., ... & Fuller, D. Q. (2013). A contextual approach to the emergence of agriculture in southwest Asia: Reconstructing early neolithic plant-food production. *Current Anthropology*, *54*(3), 000-000.

Axelsson, E., Ratnakumar, A., Arendt, M. L., Maqbool, K., Webster, M. T., Perloski, M., ... & Lindblad-Toh, K. (2013). The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature*, *495*(7441), 360-364.

Beja-Pereira, A., England, P. R., Ferrand, N., Jordan, S., Bakhiet, A. O., Abdalla, M. A., ... & Luikart, G. (2004). African origins of the domestic donkey. *Science*, *304*(5678), 1781-1781.

Benkel, B. F., Rouvinen-Watt, K., Farid, H., & Anistoroaei, R. (2009). Molecular characterization of the Himalayan mink. *Mammalian Genome*, *20*(4), 256-259.

Berryere, T. G., Schmutz, S. M., Schimpf, R. J., Cowan, C. M., & Potter, J. (2003). TYRP1 is associated with dun coat colour in Dexter cattle or how now brown cow?. *Animal genetics*, *34*(3), 169-175.

Brooks, S. A., & Bailey, E. (2005). Exon skipping in the KIT gene causes a Sabino spotting pattern in horses. *Mammalian Genome*, *16*(11), 893-902.

Bruford, M. W., Bradley, D. G., & Luikart, G. (2003). DNA markers reveal the complexity of livestock domestication. *Nature Reviews Genetics*, *4*(11), 900-910.

Carneiro, M., Rubin, C. J., Di Palma, F., Albert, F. W., Alföldi, J., Barrio, A. M., ... & Younis, S. (2014). Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. *Science*, *345*(6200), 1074-1079.

Chang, C. M., Coville, J. L., Coquerelle, G., Gourichon, D., Oulmouden, A., & Tixier-Boichard, M. (2006). Complete association between a retroviral insertion in the tyrosinase gene and the recessive white mutation in chickens. *BMC genomics*, *7*(1), 19.

Cieslak, M., Pruvost, M., Benecke, N., Hofreiter, M., Morales, A., Reissmann, M., & Ludwig, A. (2010). Origin and history of mitochondrial DNA lineages in domestic horses. *PLoS One*, 5(12), e15311.

Cieslak, M., Reissmann, M., Hofreiter, M., & Ludwig, A. (2011). Colours of domestication. *Biological Reviews*, *86*(4), 885-899.

Clutton-Brock, J. (1981). Domesticated animals from early times. *Domesticated animals from early times.*

Clutton-Brock, J. (1999). *A natural history of domesticated mammals*. Cambridge University Press.

Conolly, J., Colledge, S., Dobney, K., Vigne, J. D., Peters, J., Stopp, B., ... & Shennan, S. (2011). Meta-analysis of zooarchaeological data from SW Asia and SE Europe provides insight into the origins and spread of animal husbandry. *Journal of Archaeological Science*, *38*(3), 538-545.

Crick, F. (1970). Central dogma of molecular biology. *Nature*, *227*(5258), 561-563.

Crick, F. H. (1958, January). On protein synthesis. In *Symp Soc Exp Biol* (Vol. 12, No. 138-63, p. 8).

Darwin, C. (1859). On the Origin of Species by Means of Natural Selection. *London: John Murray*. 502 pp

Darwin, C. (1868). The Variation of Animals and Plants under Domestication (Vol. 2). *O. Judd*.

Davis, S. J. M.& Valla, F. R. Evidence for domestication of the dog 12,000 years ago in the Natufian of Israel. *Nature* 276, 608–610 (1978).

Der Sarkissian, C., Ermini, L., Schubert, M., Yang, M. A., Librado, P., Fumagalli, M., ... & Petersen, B. (2015). Evolutionary genomics and conservation of the endangered Przewalski's horse. *Current Biology*, *25*(19), 2577-2583.

Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. *Nature*, 418(6898), 700-707.

Diamond, J. M. (1996). Competition for brain space. *Nature*, *382*(6594), 756.

Dobney, K., & Larson, G. (2006). Genetics and animal domestication: new windows on an elusive process. *Journal of Zoology*, *269*(2), 261-271.

Driscoll, C. A., Macdonald, D. W., & O'Brien, S. J. (2009). From wild animals to domestic pets, an evolutionary view of domestication. *Proceedings of the National Academy of Sciences*, *106*(Supplement 1), 9971-9978.

Driscoll, C. A., Menotti-Raymond, M., Roca, A. L., Hupe, K., Johnson, W. E., Geffen, E., ... & Yamaguchi, N. (2007). The Near Eastern origin of cat domestication. *Science*, *317*(5837), 519-523.

Edwards, C. J., Magee, D. A., Park, S. D., McGettigan, P. A., Lohan, A. J., Murphy, A., ... & Bradley, D. G. (2010). A complete mitochondrial genome sequence from a mesolithic wild aurochs (Bos primigenius). *PLoS One*, *5*(2), e9255.

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 science*, *98*(4), 781-784.

Fontanesi, L., Tazzoli, M., Beretti, F., & Russo, V. (2006). Mutations in the melanocortin 1 receptor (MC1R) gene are associated with coat colours in the domestic rabbit (Oryctolagus cuniculus). *Animal genetics*, *37*(5), 489-493.

Frantz, A. C., Zachos, F. E., Kirschning, J., Cellina, S., Bertouille, S., Mamuris, Z., ... & 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. *Biological Journal of the Linnean Society*, 110(1), 104-115.

Frantz, L. A., Mullin, V. E., Pionnier-Capitan, M., Lebrasseur, O., Ollivier, M., Perri, A., ... & Tresset, A. (2016). Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science*, *352*(6290), 1228-1231.

Fredriksen, A. (2016). Of wildcats and wild cats: Troubling species-based conservation in the Anthropocene. *Environment and Planning D: Society and Space*, *34*(4), 689-705.

Fuller, D. Q. (2006). Agricultural origins and frontiers in South Asia: a working synthesis. *Journal of World Prehistory*, *20*(1), 1-86.

Fuller, D. Q., Willcox, G., & Allaby, R. G. (2011). Cultivation and domestication had multiple origins: arguments against the core area hypothesis for the origins of agriculture in the Near East. *World Archaeology*, *43*(4), 628-652.

Giuffra, E., Evans, G., Törnsten, A., Wales, R., Day, A., Looft, H., ... & Andersson, L. (1999). The Belt mutation in pigs is an allele at the Dominant white (I/KIT) locus. *Mammalian Genome*, *10*(12), 1132-1136.

Godinho, R., Llaneza, L., Blanco, J. C., Lopes, S., Alvares, F., Garcia, E. J., ... & Ferrand, N. (2011). Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula. *Molecular Ecology*, *20*(24), 5154-5166.

Goedbloed, D. J., Megens, H. J., Van Hooft, P., Herrero-Medrano, J. M., Lutz, W., Alexandri, P., ... & Prins, H. H. T. (2013). Genome-wide single nucleotide polymorphism analysis reveals recent genetic introgression from domestic pigs into Northwest European wild boar populations. *Molecular Ecology*, 22(3), 856-866.

Gratten, J., Beraldi, D., Lowder, B. V., McRae, A. F., Visscher, P. M., Pemberton, J. M., & Slate, J. (2007). Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proceedings of the Royal Society of London B: Biological Sciences*, *274*(1610), 619-626.

Gupta, A. K. (2004). Origin of agriculture and domestication of plants and animals linked to early Holocene climate amelioration. *Current Science-Bangalore*, *87*, 54-59.

Haase, B., Brooks, S. A., Tozaki, T., Burger, D., Poncet, P. A., Rieder, S., ... & Leeb, T. (2009). Seven novel KIT mutations in horses with white coat colour phenotypes. *Animal genetics*, *40*(5), 623-629.

Harrison, R. G., & Larson, E. L. (2014). Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, *105*(S1), 795-809.

Hasegawa, M., Kishino, H., & Yano, T. A. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of molecular evolution*, *22*(2), 160-174.

Hearing, V. J. (2011). Determination of melanin synthetic pathways. *Journal of Investigative Dermatology*, *131*, E8-E11.

Heyne, H. O., Lautenschläger, S., Nelson, R., Besnier, F., Rotival, M., Cagan, A., ... & Petretto, E. (2014). Genetic influences on brain gene expression in rats selected for tameness and aggression. *Genetics*, *198*(3), 1277-1290.

Hiragaki, T., Inoue-Murayama, M., Miwa, M., Fujiwara, A., Mizutani, M., Minvielle, F., & Ito, S. I. (2008). Recessive black is allelic to the yellow plumage locus in Japanese quail and associated with a frameshift deletion in the ASIP gene. *Genetics*, *178*(2), 771-775.

Hu, Y., Hu, S., Wang, W., Wu, X., Marshall, F. B., Chen, X., ... & Wang, C. (2014). Earliest evidence for commensal processes of cat domestication. *Proceedings of the National Academy of Sciences*, *111*(1), 116-120.

Hubbard, J. K., Uy, J. A. C., Hauber, M. E., Hoekstra, H. E., & Safran, R. J. (2010). Vertebrate pigmentation: from underlying genes to adaptive function. *Trends in Genetics*, *26*(5), 231-239.

Hutchison, C. A., Newbold, J. E., Potter, S. S., & Edgell, M. H. (1974). Maternal inheritance of mammalian mitochondrial DNA. *Nature*, 251(5475), 536-538.

Huxley, J. (1942). Evolution. The Modern Synthesis. *Evolution. The Modern Synthesis.* 1942 pp.645 pp.

Icken, W., Bennewitz, J., & Kalm, E. (2007). Analysis of auction data for horses and influence factors on pricing. *Zuchtungskunde*, *79*(2), 111-118.

Imes, D. L., Geary, L. A., Grahn, R. A., & Lyons, L. A. (2006). Albinism in the domestic cat (Felis catus) is associated with a tyrosinase (TYR) mutation. *Animal genetics*, *37*(2), 175-178.

Ito, S., & Wakamatsu, K. (2003). Quantitative analysis of eumelanin and pheomelanin in humans, mice, and other animals: a comparative review. *Pigment cell research*, *16*(5), 523-531.

Jensen, P. (2014). Behavior genetics and the domestication of animals. *Annu. Rev. Anim. Biosci.*, *2*(1), 85-104.

Jepsen, B. I., Siegismund, H. R., & Fredholm, M. (2002). Population genetics of the native caribou (Rangifer tarandus groenlandicus) and the semi-domestic reindeer (Rangifer tarandus tarandus) in Southwestern Greenland: Evidence of introgression. *Conservation Genetics*, *3*(4), 401-409.

Kidd, A. G., Bowman, J., Lesbarreres, D., & Schulte-Hostedde, A. I. (2009). Hybridization between escaped domestic and wild American mink (Neovison vison). *Molecular Ecology*, *18*(6), 1175-1186.

Kijas, J. M. H., & Andersson, L. (2001). A phylogenetic study of the origin of the domestic pig estimated from the near-complete mtDNA genome. *Journal of Molecular Evolution*, 52(3), 302-308.

Kijas, J. M. H., 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*(2), 779-785.

Kijas, J. M. H., Wales, R., Törnsten, A., Chardon, P., Moller, M., & Andersson, L. (1998). Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. *Genetics*, *150*(3), 1177-1185.

Kikkawa, Y., Takada, T., Nomura, K., Namikawa, T., Yonekawa, H., & Amano, T. (2003). Phylogenies using mtDNA and SRY provide evidence for male-mediated introgression in Asian domestic cattle. *Animal Genetics*, 34(2), 96-101.

Lamoreux, M., Wakamatsu, K., & Ito, S. (2001). Interaction of major coat color gene functions in mice as studied by chemical analysis of eumelanin and pheomelanin. *Pigment Cell & Melanoma Research*, *14*(1), 23-31.

Larson, G., & Burger, J. (2013). A population genetics view of animal domestication. *Trends in Genetics*, *29*(4), 197-205.

Larson, G., Albarella, U., Dobney, K., Rowley-Conwy, P., Schibler, J., Tresset, A., ... & Bălăçsescu, A. (2007). Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proceedings of the National Academy of Sciences*, *104*(39), 15276-15281.

Larson, G., Dobney, K., Albarella, U., Fang, M., Matisoo-Smith, E., Robins, J., ... & Rowley-Conwy, P. (2005). Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science*, *307*(5715), 1618-1621.

Larson, G., Piperno, D. R., Allaby, R. G., Purugganan, M. D., Andersson, L., Arroyo-Kalin, M., ... & Doust, A. N. (2014). Current perspectives and the future of domestication studies. *Proceedings of the National Academy of Sciences*, *111*(17), 6139-6146.

Legault, C. (1998). Genetics of colour variation. *The Genetics of the pig*, 51-69.

Lewin, H. A. (2009). It's a Bull's Market. *Science*, *324*(5926), 478-479.

Loehr, J., Worley, K., Moe, J., Carey, J., & Coltman, D. W. (2008). MC1R variants correlate with thinhorn sheep colour cline but not individual colour. *Canadian Journal of Zoology*, *86*(2), 147-150.

Ludwig, A., Pruvost, M., Reissmann, M., Benecke, N., Brockmann, G. A., Castaños, P., ... & Slatkin, M. (2009). Coat color variation at the beginning of horse domestication. *Science*, *324*(5926), 485-485.

Ludwig, A., Reissmann, M., Benecke, N., Bellone, R., Sandoval-Castellanos, E., Cieslak, M., ... & Pruvost, M. (2015). Twenty-five thousand years of fluctuating selection on leopard complex spotting and congenital night blindness in horses. *Phil. Trans. R. Soc. B*, *370*(1660), 20130386.

Lyons, L. A., Foe, I. T., Rah, H. C., & Grahn, R. A. (2005a). Chocolate coated cats: TYRP1 mutations for brown color in domestic cats. *Mammalian Genome*, *16*(5), 356-366.

Lyons, L. A., Imes, D. L., Rah, H. C., & Grahn, R. A. (2005). Tyrosinase mutations associated with Siamese and Burmese patterns in the domestic cat (Felis catus). *Animal genetics*, *36*(2), 119-126.

MacHugh, D. E., Larson, G., & Orlando, L. (2017). Taming the Past: Ancient DNA and the Study of Animal Domestication. *Annual Review of Animal Biosciences*, (0).

MacHugh, D. E., Shriver, M. D., Loftus, R. T., Cunningham, P., & Bradley, D. G. (1997). Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (Bos taurus and Bos indicus). *Genetics*, *146*(3), 1071-1086.

Marklund, S., Moller, M., Sandberg, K., & Andersson, L. (1999). Close association between sequence polymorphism in the KIT gene and the roan coat color in horses. *Mammalian Genome*, *10*(3), 283-288.

Marsden, C. D., Ortega-Del Vecchyo, D., O'Brien, D. P., Taylor, J. F., Ramirez, O., Vilà, C., ... & Lohmueller, K. E. (2016). Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. *Proceedings of the National Academy of Sciences*, *113*(1), 152-157.

Meiri, M., Huchon, D., Bar-Oz, G., Boaretto, E., Horwitz, L. K., Maeir, A. M., ... & Finkelstein, I. (2013). Ancient DNA and population turnover in southern levantine pigs-signature of the sea peoples migration? *Scientific reports*, *3*, 3035.

Mendel G (1869): Ueber einige aus kunstlicher Befruchtung gewon- nenen Hieracium-Bastarde. Verhandl naturf Verein Briinn 8 (Abhandl.):26-31.

Mendel, J. G. (1866). Versuche über Plflanzenhybriden Verhandlungen des naturforschenden Vereines in Brünn, Bd. IV für das Jahr, 1865. Abhandlungen: 3–47. Translated in: Druery, CT, and Bateson, W. 1901. *Experiments in plant hybridization*, 1-32.

Minvielle, F., Bed'Hom, B., Coville, J. L., Ito, S. I., Inoue-Murayama, M., & Gourichon, D. (2010). The" silver" Japanese quail and the MITF gene: causal mutation, associated traits and homology with the" blue" chicken plumage. *BMC genetics*, *11*(1), 15.

Montague, M. J., Li, G., Gandolfi, B., Khan, R., Aken, B. L., Searle, S. M., ... & Driscoll, C. A. (2014). Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *Proceedings of the National Academy of Sciences*, *111*(48), 17230-17235.

Mullis, K., Faloona, F., Scharf, S., Saiki, R. K., Horn, G. T., & Erlich, H. (1986, January). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. In *Cold Spring Harbor symposia on quantitative biology* (Vol. 51, pp. 263-273). Cold Spring Harbor Laboratory Press.

O'Connor, T. P. (1997). Working at relationships: another look at animal domestication. Antiquity, 71(271), 149-156.

Outram, A. K., Stear, N. A., Bendrey, R., Olsen, S., Kasparov, A., Zaibert, V., ... & Evershed, R. P. (2009). The earliest horse harnessing and milking. *Science*, *323*(5919), 1332-1335.

Pang, J. F. et al. mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Mol. Biol. Evol*. 26, 2849–2864 (2009).

Park, S. D., Magee, D. A., McGettigan, P. A., Teasdale, M. D., Edwards, C. J., Lohan, A. J., ... & Chamberlain, A. T. (2015). Genome sequencing of the extinct Eurasian wild aurochs, Bos primigenius, illuminates the phylogeography and evolution of cattle. *Genome biology*, *16*(1), 234.

Patel, A. K., & Meadow, R. (1998). The exploitation of wild and domestic water buffalo in prehistoric northwestern South Asia. *Archaeozoology of the Near East III*, 180-199.

Peters, J., Lebrasseur, O., Deng, H., & Larson, G. (2016). Holocene cultural history of Red jungle fowl (Gallus gallus) and its domestic descendant in East Asia. *Quaternary Science Reviews*, *142*, 102-119.

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*(1), 305-311.

Price, E. O. (1999). Behavioral development in animals undergoing domestication. *Applied Animal Behaviour Science*, *65*(3), 245-271.

Purugganan, M. D., & Fuller, D. Q. (2009). The nature of selection during plant domestication. *Nature, 457(7231), 843-848. e*, *457*(7231), 843-848.

Qi, X. B., Jianlin, H., Wang, G., Rege, J. E. O., & Hanotte, O. (2010). Assessment of cattle genetic introgression into domestic yak populations using mitochondrial and microsatellite DNA markers. *Animal genetics*, *41*(3), 242-252.

Rannala, B., & Yang, Z. (2013). Molecular clock dating. *The Princeton guide to evolution (ed. JB Losos)*, 67-74.

Reinsch, N., Thomsen, H., Xu, N., Brink, M., Looft, C., Kalm, E., ... & Leyhe, B. (1999). A QTL for the degree of spotting in cattle shows synteny with the KIT locus on chromosome 6. *Journal of Heredity*, *90*(6), 629-634.

Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, *27*(1), 83-109.

Rossel, S., Marshall, F., Peters, J., Pilgram, T., Adams, M. D., & O'Connor, D. (2008). Domestication of the donkey: Timing, processes, and indicators. *Proceedings of the National Academy of Sciences*, *105*(10), 3715-3720.

Ross-Ibarra, J., Morrell, P. L., & Gaut, B. S. (2007). Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proceedings of the National Academy of Sciences*, *104*(suppl 1), 8641-8648.

Royo, L. J., Alvarez, I., Arranz, J. J., Fernández, I., Rodríguez, A., Pérez-Pardal, L., & Goyache, F. (2008). Differences in the expression of the ASIP gene are involved in the recessive black coat colour pattern in sheep: evidence from the rare Xalda sheep breed. *Animal genetics*, *39*(3), 290-293.

Rubin, C. J., Megens, H. J., Barrio, A. M., Maqbool, K., Sayyab, S., Schwochow, D., ... & Archibald, A. L. (2012). Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences*, *109*(48), 19529-19536.

Rubin, C. J., Zody, M. C., Eriksson, J., Meadows, J. R., Sherwood, E., Webster, M. T., ... &

Hallböök, F. (2010). Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*, *464*(7288), 587-591.

Saetre, P., Lindberg, J., Leonard, J. A., Olsson, K., Pettersson, U., Ellegren, H., ... & Jazin, E. (2004). From wild wolf to domestic dog: gene expression changes in the brain. *Molecular Brain Research*, *126*(2), 198-206.

Schmutz, S. M., Berryere, T. G., & Goldfinch, A. D. (2002). TYRP1 and MC1R genotypes and their effects on coat color in dogs. *Mammalian Genome*, *13*(7), 380-387.

Schmutz, S. M., Berryere, T. G., Ciobanu, D. C., Mileham, A. J., Schmidtz, B. H., & Fredholm, M. (2004). A form of albinism in cattle is caused by a tyrosinase frameshift mutation. *Mammalian Genome*, *15*(1), 62-67.

Schmutz, S. M., Berryere, T. G., Ellinwood, N. M., Kerns, J. A., & Barsh, G. S. (2003). MC1R studies in dogs with melanistic mask or brindle patterns. *Journal of Heredity*, *94*(1), 69-73.

Schubert, M., Jónsson, H., Chang, D., Der Sarkissian, C., Ermini, L., Ginolhac, A., ... & Fumagalli, M. (2014). Prehistoric genomes reveal the genetic foundation and cost of horse domestication. *Proceedings of the National Academy of Sciences*, *111*(52), E5661-E5669.

Sharma, S., Wagh, S., & Govindarajan, R. (2002). Melanosomal proteins–role in melanin polymerization. *Pigment cell research*, *15*(2), 127-133.

Skoglund, P., Ersmark, E., Palkopoulou, E., & Dalén, L. (2015). Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. *Current Biology*, *25*(11), 1515-1519.

Skoglund, P.,Gotherstrom,A.&Jakobsson, M. Estimation of population divergence times from non-overlapping genomic sequences: examples from dogs and wolves. *Mol. Biol. Evol*. 28, 1505–1517 (2011).

Stritzel, S., Wöhlke, A., & Distl, O. (2009). A role of the microphthalmia-associated transcription factor in congenital sensorineural deafness and eye pigmentation in Dalmatian dogs. *Journal of Animal Breeding and Genetics*, *126*(1), 59-62.

Thulin, C. G., Fang, M., & O Averianov, A. (2006). Introgression from Lepus europaeus to L. timidus in Russia revealed by mitochondrial single nucleotide polymorphisms and nuclear microsatellites. *Hereditas*, 143(2006), 68-76.

Trut, L., Oskina, I., & Kharlamova, A. (2009). Animal evolution during domestication: the domesticated fox as a model. *Bioessays*, *31*(3), 349-360.

Utzeri, V. J., Bertolini, F., Ribani, A., Schiavo, G., Dall'Olio, S., & Fontanesi, L. (2016). The albinism of the feral Asinara white donkeys (Equus asinus) is determined by a missense mutation in a highly conserved position of the tyrosinase (TYR) gene deduced protein. *Animal genetics*, *47*(1), 120-124.

Utzeri, V. J., Ribani, A., & Fontanesi, L. (2014). A premature stop codon in the TYRP1 gene is associated with brown coat colour in the European rabbit (Oryctolagus cuniculus). *Animal genetics*, *45*(4), 600-603.

Vigne, J. D., Guilaine, J., Debue, K., Haye, L., & Gérard, P. (2004). Early taming of the cat in Cyprus. *Science*, *304*(5668), 259-259.

Vilà, C., Leonard, J. A., Götherström, A., Marklund, S., Sandberg, K., Lidén, K., ... & Ellegren, H. (2001). Widespread origins of domestic horse lineages. *Science*, *291*(5503), 474-477.

Vilà, C., Savolainen, P., Maldonado, J. E., Amorim, I. R., Rice, J. E., Honeycutt, R. L., ... & Wayne, R. K. (1997). Multiple and ancient origins of the domestic dog. *Science*, *276*(5319), 1687-1689.

vonHoldt, B. M. et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* 464, 898–902 (2010).

Wallace, A. R. (1870). Contributions to the theory of natural selection: a series of essays. *Macmillan and Company*.

Wilkins, A. S., Wrangham, R. W., & Fitch, W. T. (2014). The "domestication syndrome" in mammals: A unified explanation based on neural crest cell behavior and genetics. *Genetics*, *197*(3), 795-808.

Wright, D., Rubin, C. J., Martinez Barrio, A., Schütz, K., Kerje, S., Brändström, H., ... & Andersson, L. (2010). The genetic architecture of domestication in the chicken: effects of pleiotropy and linkage. *Molecular ecology*, 19(23), 5140-5156.

Zeder, M. A. (2006). Central questions in the domestication of plants and animals. *Evolutionary Anthropology: Issues, News, and Reviews*, *15*(3), 105-117.

1.5 AIM

The aim of this Thesis was to identify and analyse polymorphisms associated with phenotypic traits that have been selected during the domestication processes and may differentiate breeds or populations in several livestock species, i.e. rabbit, donkey and pig, and evaluate the potential evolutionary effects on the wild counterparts, considering the specific example of the introgression between domestic pigs and wild boar populations. Obtained results have been also evaluated for applications that consider allele and genotype frequency differences for authentication purposes.

In particular in the paragraphs 2.1 and 2.2 we focused our attention on two key genes involved in the melanogenesis pathway identifying causative mutations of two types of albinism: in rabbits it has been detected a stop codon gain mutation in TYRP1 gene associated with the *Brown locus* in Havana rabbits while in the second paper it has been identified the causative missense mutation leading to a severe albinism form in the semi-feral population of Asinara White Donkeys.

In paragraphs 2.3 and 2.4 two different points of view highlighted the possibility to use the coat colour related gene MC1R and NR6A1 gene as meat product authentication markers and introgression markers respectively.

2 PAPERS

2.1 A premature stop codon in the *TYRP1* gene is associated with brown coat colour in the European rabbit (*Oryctolagus cuniculus*)

V.J. Utzeri*, A. Ribani*, and L. Fontanesi*,†

PUBLISHED in ANIMAL GENETICS - **DOI:** 10.1111/age.12171

* Department of Agricultural and Food Sciences (DISTAL), Division of Animal Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

† Centre for Genome Biology, University of Bologna, 40126 Bologna, Italy

Running head: *TYRP1* mutation and *brown* locus in rabbit.

Summary

Classical genetic studies in European rabbits (*Oryctolagus cuniculus*) suggested the presence of two alleles at the *brown* coat colour locus: a wild type *B* allele that gives dense black pigment throughout the coat and a recessive *b* allele that in homozygous condition (*b*/*b* genotype) produces brown rabbits that are unable to develop black pigmentation. In several other species this locus is determined by mutations in the *tyrosinase-related protein 1* (*TYRP1*) gene, encoding a melanocyte enzyme needed for the production of dark eumelanin. In this study we investigated the rabbit *TYRP1* gene as a strong candidate for the rabbit *brown* coat colour locus. A total of 3846 bp of the *TYRP1* gene were sequenced in eight rabbits of different breeds and identified 23 single nucleotide polymorphisms (SNPs; 12 in intronic regions, 5 in exons and 6 in the 3'-untranslated region) and an insertion/deletion of 13 bp, in the 3'-untranslated region, organized in a few haplotypes. A mutation in exon 2 (g.41360196G>A) leads to a premature stop codon at position 190 of the deduced aminoacid sequence (p.W190ter). Therefore, translation predicts a truncated TYRP1 protein lacking almost completely the tyrosinase domain. Genotyping 203 rabbits of 32 different breeds identified this mutation only in brown Havana rabbits. Its potential functional relevance in disrupting the *TYRP1* protein and its presence only in brown animals, strongly argue for this nonsense mutation being a causative mutation for the recessive *b* allele at the *brown* locus in *Oryctolagus cuniculus*.

Key words: *brown* locus, *tyrosinase-related protein 1*, domestic breed, mutation, pigmentation, *OCA3*.

Text

Several coat colour loci described by classical genetic studies at the beginning of the last century in the domestic rabbit (*Oryctolagus cuniculus*) have been recently characterized at the molecular level. In particular, mutations in the melanocortin 1 receptor (*MC1R*) gene determine different alleles at the *extension* (*E*) locus series (Fontanesi *et al.* 2006; 2010b), the causative mutation of the nonagouti *a* allele of the *agouti* locus is determined by a frameshift mutation in the agouti signaling protein (*ASIP*) gene (Fontanesi *et al.* 2010a), a deletion in the melanophilin (*MLPH*) gene causes the *dilute* coat colour phenotype (Fontanesi *et al.* 2014) and mutations in the tyrosinase (*TYR*) genes determine alleles of the *albino* (*C*) locus (Aigner *et al.* 2001).

However, the *Brown* (*B*) coat colour locus has not yet been characterized at the molecular level in rabbit. Early genetic studies suggested the presence of two alleles at this locus: a wild type *B* allele that gives dense black pigment throughout the coat and a recessive *b* allele that in homozygous condition (*b*/*b* genotype) produces brown rabbits that are unable to develop black pigmentation (Castle 1930; Robinson 1958; Searle 1968). Early comparative genetic studies on coat colour loci suggested the presence of a homolog locus in several mammals (Searle 1968). The molecular characterization of the *brown* locus started in mice when the *tyrosinase-related protein 1* (*Tyrp1*) gene was assigned to the chromosome region in which this locus was mapped (Jackson 1988). Then, mutations in this gene, encoding a melanocyte enzyme needed for the production of dark eumelanin, were shown to determine brown coat colour in mice (Zdarsky *et al.* 1990; Smyth *et al.* 2006) as well as in several other species like dogs (Schmutz *et al.* 2002), cats (Lyons *et al.* 2005; Schmidt-Küntzel *et al.* 2005), cattle (Berryere *et al.* 2003; Mohanty *et al.* 2008), sheep (Gratten *et al.* 2007) and pigs (Ren *et al.* 2011).

In this study we investigated the rabbit *TYRP1* gene as a strong candidate for the *brown* coat colour locus in *Oryctolagus cuniculus*. We designed eight primer pairs (Table S1) on the sequence of the *TYRP1* gene (Ensembl accession number ENSOCUG00000004221) included on rabbit chromosome 1 of the oryCun2.0 genome assembly provided by the Broad Institute within the

Mammalian Genome Project, http://www.broadinstitute.org/science/projects/mammals-models/rabbit/rabbit-genome-sequencing-project (Lindblad-Toh *et al.*. 2011). Primers were used to amplify all recognized coding exons, portions of the intronic regions positioned downstream and upstream to the exons and 5'- and 3'- untranslated regions (Table S1) in eight rabbits of different coat colours: 3 Havana (with brown coat colour, expected to have *b/b* genotype), 1 Belgian hare, 1 Champagne d'Argent, 1 Chinchillà, 2 wild rabbits (Table 1). Obtained PCR products were sequenced using the Big Dye v3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) as previously described (Fontanesi *et al.* 2014). Sequencing reactions were loaded on an ABI3100 Avant sequencer (Applied Biosystems). All sequences were visually inspected and aligned with the help of the CodonCode Aligner software (http://www.codoncode.com/aligner).

A total of 3846 bp of the *TYRP1* gene (1614 bp of exons; 1488 bp of introns; 80 bp of the 5'-untranslated region or 5'-flanking region; 644 bp of the 3'-untranslated region or 3'-flanking region) were sequenced in the eight selected rabbits and identified 23 single nucleotide polymorphisms (SNPs; 12 in intronic regions, 5 in exons and 6 in the 3'-untranslated region) and an insertion/deletion of 13 bp, in the 3'-untranslated region, organized in a few haplotypes (Table S2). Of the 5 exonic polymorphisms, 4 were synonymous substitutions in exons 1, 2, 5, and 6. The most interesting mutation was in exon 2 (g.41360196G>A) that leads to a premature stop codon at position 190 of the deduced protein sequence (p.W190ter; Fig. 1) that, in the wild type form, contains 537 predicted aminoacids (Ensembl accession number: ENSOCUP00000003649). Therefore, translation predicts a truncated TYRP1 protein lacking almost completely the tyrosinase domain, as identified by Pfam. This mutation was identified only in the sequenced brown Havana rabbits.

Based on these first evidences this premature stop codon in exon 2 could be a strong candidate for the *brown* locus in rabbit. To further evaluate this hypothesis we genotyped this mutation by PCR-RFLP (Table S1 and Fig. S1) in a total of 203 rabbits from 32 breeds/populations with different coat colours including a few additional rabbits of the Havana breed (Table 1). All

animals of this breeds were confirmed to be homozygous for the mutation causing a stop codon in exon 2, whereas all other animals did not carry this mutated allele. Considering all together animals that did not have a brown coat colour that were genotyped at this locus (all breeds/populations except Havana), association between the premature stop codon genotype and brown coat colour was highly significant (P=1.61E-14; Fischer exact test). Even if we could not formally exclude the presence of other mutations in linkage disequilibrium with the g.41360196G>A SNP, its functional relevance in disrupting the TYRP1 protein that is the product of the most important candidate gene for this coat colour locus, and its presence only in brown animals, strongly argue for this nonsense mutation being a causative mutation for the recessive *b* allele at the *brown* locus in *Oryctolagus cuniculus*. Crossbreeding experiment could be carried out to confirm the cosegragation of the *brown* locus with this mutation. However, it is interesting to note that mutations in exon 2 of the *TYRP1* gene causing brown coat colour phenotypes occur also in other species. In particular, a missense mutation (p.C86Y) has been reported in mouse (Zdarsky *et al.* 1990) and a missense mutation (p.A3G) of the chocolate haplotype and a nonsense mutation at codon 100 causing the cinnamon allele of the B locus have been identified in cats (Lyons *et al.* 2005; Schmidt-Küntzel *et al.* 2005).

Based on these data and on results we obtained for other coat colour genes in rabbits, it could be possible to understand the main genetic factors affecting coat colour in the Havana breed. The eumelanic background of this breed, that is derived by the non-agouti mutation we identified in the *ASIP* gene (Fontanesi *et al.* 2010a), supported by a wild type allele at the *extension* locus (Fontanesi *et al.* 2006), is then modified by the mutation in the *TYRP1* gene that prevents the production of black pigments.

Mutations in the *TYRP1* gene have been associated with oculocutaneous albinism (OCA) type 3 in humans (Boissy *et al.* 1996). The identification of this premature stop codon mutation in rabbit provides an additional animal model for this disease and a natural way to investigate the functional relevance and activity of the N terminal part of the TYRP1 protein.

Acknowledgments

We thank Associazione Nazionale Coniglicoltori Italiani (ANCI) for the help in sampling rabbits of different breeds. This work was funded by University of Bologna RFO funds and has been carried out in the framework of the COST Action TD1101 “A Collaborative European Network on Rabbit Genome Biology – RGB-Net”.

References

Aigner B., Besenfelder U., Müller M. & Brem G. (2000) Tyrosinase gene variants in different rabbit strains. *Mammalian Genome* **11**, 700-2.

Berryere T.G., Schmutz S.M., Schimpf R.J., Cowan C.M. & Potter J. (2003) TYRP1 is associated with dun coat colour in Dexter cattle or how now brown cow? *Animal Genetics* **34**, 169-75.

Boissy R.E., Zhao H., Oetting W.S., Austin L.M., Wildenberg S.C., Boissy Y.L., Zhao Y., Sturm R.A., Hearing V.J., King R.A. & Nordlund J.J. (1996) Mutation in and lack of expression of tyrosinase-related protein-1 (TRP-1) in melanocytes from an individual with brown oculocutaneous albinism: a new subtype of albinism classified as "OCA3". *American Journal of Human Genetics* **58**, 1145-56.

Castle W.E. (1930) *The Genetics of Domestic Rabbit*. Cambridge Harvard University Press, London, UK.

Fontanesi L., Tazzoli M., Beretti F. & Russo V. (2006) Mutations in the melanocortin 1 receptor (*MC1R*) gene are associated with coat colours in the domestic rabbit (*Oryctolagus cuniculus*). *Animal Genetics* **37**, 489-93.

Fontanesi L., Tazzoli M. & Russo V. (2007) Non-invasive and simple methods for sampling DNA for PCR analysis of melanocortin 1 receptor (*MC1R*) gene mutations: a technical note. *World Rabbit Science* **15**, 121-6.

Fontanesi L., Forestier L., Allain D., Scotti E., Beretti F., Deretz-Picoulet S., Pecchioli E., Vernesi C., Robinson T.J., Malaney J.L., Russo V. & Oulmouden A. (2010a) Characterization of the rabbit agouti signaling protein (*ASIP*) gene: Transcripts and phylogenetic analyses and identification of the causative mutation of the nonagouti black coat colour. *Genomics* **95**, 166-75.

Fontanesi L., Scotti E., Colombo M., Beretti F., Forestier L., Dall'Olio S., Deretz S., Russo V., Allain D. & Oulmouden A. (2010b) A composite six bp in-frame deletion in the melanocortin 1 receptor (*MC1R*) gene is associated with the Japanese brindling coat colour in rabbits (*Oryctolagus cuniculus*). *BMC Genetics* **11**, 59.

Fontanesi L., Scotti E., Allain D. & Dall'Olio S. (2014) A frameshift mutation in the melanophilin gene causes the *dilute* coat colour in rabbit (*Oryctolagus cuniculus*) breeds. *Animal Genetics*, DOI: 10.1111/age.12104.

Gratten J., Beraldi D., Lowder B.V., McRae A.F., Visscher P.M., Pemberton J.M. & Slate J. (2007) Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proceedings of the Royal Society B: Biological Sciences* **274**, 619-26.

Jackson I.J. (1988) A cDNA encoding tyrosinase-related protein maps to the mouse brown locus. Proceedings of the National Academy of Sciences of the USA 85, 4391-6.

Lindblad-Toh K., Garber M., Zuk O., *et al*.. (2011) A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**, 476-82.

Lyons L.A., Foe I.T., Rah H.C. & Grahn R.A. (2005) Chocolate coated cats: TYRP1 mutations for brown color in domestic cats. *Mammalian Genome* **16**, 356-66.

Mohanty T.R., Seo K.S., Park K.M., Choi T.J., Choe H.S., Baik D.H. & Hwang I.H. (2008) Molecular variation in pigmentation genes contributing to coat colour in native Korean Hanwoo cattle. *Animal Genetics* **39**, 550-3.

Ren J., Mao H., Zhang Z., Xiao S., Ding N. & Huang L. (2011) A 6-bp deletion in the *TYRP1* gene causes the brown colouration phenotype in Chinese indigenous pigs. *Heredity* **106**, 862-8.

Robinson R. (1958) Genetic studies of the rabbit. *Bibliographia Genetica* **17**, 229-558.

Schmidt-Küntzel A., Eizirik E., O'Brien S.J. & Menotti-Raymond M. (2005) Tyrosinase and tyrosinase related protein 1 alleles specify domestic cat coat color phenotypes of the albino and brown loci. *Journal of Heredity* **96**, 289-301.

Schmutz S.M., Berryere T.G. & Goldfinch A.D. (2002) TYRP1 and MC1R genotypes and their effects on coat color in dogs. *Mammalian Genome* **13**, 380-7.

Searle A.G. (1968) *Comparative Genetics of Coat Colour in Mammals*. Logos Press, London, UK.

Smyth I.M., Wilming L., Lee A.W., *et al.* (2006) Genomic anatomy of the Tyrp1 (brown) deletion complex. *Proceedings of the National Academy of Sciences of the USA* **103**, 3704-9.

Zdarsky E., Favor J. & Jackson I.J. (1990) The molecular basis of brown, an old mouse mutation, and of an induced revertant to wild type. *Genetics* **126**, 443-9.

Table 1. Rabbits genotyped for the *TYRP1* g.41360196G>A mutation.

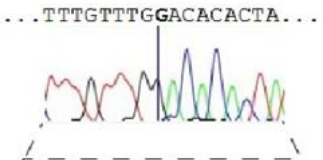
Breeds/Populations (no. of rabbits)	**Coat colour [proposed *brown* genotype]**	**g.41360196G>A genotypes [1]**		
		GG	**GA**	**AA**
Alaska (3)	Self black [*B/-*]	3	-	-
Ariete Nano (7)	Wild-grey [*B/-*]	7	-	-
Balck and Tan (8)	Black fire [*B/-*]	8	-	-
Belgian Hare (10)	Reddish laced with black [*B/-*]	10	-	-
Blue Vienna (7)	Dark blue [*B/-*]	7	-	-
Bristle White (4)	Black with white [*B/-*]	4	-	-
Burgundy Fawn (3)	Fawn [*B/-*]	3	-	-
Californian (11)	White whit black markings [*B/-*]	11	-	-
Champagne d'Argent (8)	Silver and black [*B/-*]	8	-	-
Checkered Giant (10)	White with black markings [*B/-*]	10	-	-
Commercial Hybrid (11)	White albino [-/-]	11	-	-
Dutch (7)	White with black markings [*B/-*]	7	-	-
English Spot (7)	White with black markings [*B/-*]	7	-	-
Ermine (4)	White with blue eyes [-/-]	4	-	-
Fairy Pearly (3)	Pearlyng grey [*B/-*]	3	-	-
Giant Chinchilla (6)	Chinchilla [*B/-*]	6	-	-
Giant Grey (10)	Wild-grey [*B/-*]	10	-	-
Giant White (5)	White albino [-/-]	5	-	-
Gold Saxony (3)	Red [*B/-*]	3	-	-

Havana (8)	Chocolate brown [*b/b*]	-	-	8
Hotot (2)	White with black markings [*B/-*]	2	-	-
Japanese (3)	Japanese brindling [*B/-*]	3	-	-
Leprino di Viterbo (4)	Wild-grey [*B/-*]	4	-	-
Lop (4)	Wild-grey [*B/-*]	4	-	-
Mini Silver (4)	Fawn with silvering [*B/-*]	4	-	-
New Zealand Red (4)	Red [*B/-*]	4	-	-
New Zealand White (8)	White albino [*B/-*]	8	-	-
Rhinelander (7)	White with black and yellow markings [*B/-*]	7	-	-
Satin (3)	Black or broken [*B/-*]	3	-	-
Silver (9)	Black with silvering [*B/-*]	9	-	-
White Vienna (2)	White-blue eyes [-/-]	2	-	-
Wild rabbits - Sardinia (18)	Wild-grey [*B/-*]	18	-	-
Total (203)		**195**		**8**

[1] Genotype AA is homozygous for the premature stop codon in exon 2 of the *TYRP1* gene.

Figure 1. Part of exon 2 sequence of the wild type and brown *TYRP1* alleles. The g.41360196G>A mutation causing a premature stop codon is indicated.

Wild type allele

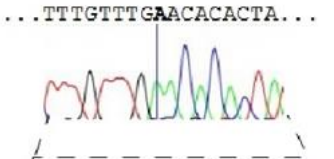



```
TTT GAG AAT ATT TCC ATT TAT AAC TAC TTT GTT TGG ACA CAC TAT TAC
 F   E   N   I   S   I   Y   N   Y   F   V   W   T   H   Y   Y

TCA GTC AAA AAG ACT TTT CTT GGT CCG GGG CAG GAA AGC TTT GGT GAA
 S   V   K   K   T   F   L   G   P   G   Q   E   S   F   G   E

GTT GAT TTT TCT CAT GAA GGA CCG GCT TTC CTC ACA TGG CAC AGG TAC
 V   D   F   S   H   E   G   P   A   F   L   T   W   H   R   Y
```

***Brown* allele**




```
TTT GAG AAT ATT TCC ATT TAT AAC TAC TTT GTT TGA ACA CAC TAT TAC
 F   E   N   I   S   I   Y   N   Y   F   V   *

TCA GTC AAA AAG ACT TTT CTT GGT CCG GGG CAG GAA AGC TTT GGT GAA

GTT GAT TTT TCT CAT GAA GGA CCG GCT TTC CTC ACA TGG CAC AGG TAC
```

Table S1. PCR primers, PCR conditions and use of the amplified fragments of the rabbit *TYRP1* gene.

Primer pair name	Primer forward	Primer reverse	Annealing temperature[1]	Gene region/bp[2]	Use
Ex1	TCCTTTACAGCTGGACTTTGG	CTGGAAGAACAGCTGGCAAT	56	5'-flanking, exon 1, intron 1/573	Sequencing
Ex2	CAGAGAACACATGTACACATAACCA	TCCTTTACAGACGAGATGCC	57	intron 1, exon 2, intron 2/ 500	Sequencing
Ex3	TGGCCATGCACTTACTGGT	GCATCTAGCATATTCAAGGCA	56	intron 3, exon 3, intron 4 /500	Sequencing
Ex4	GGACATGGTCACTTCAATATTCTC	GATTTGGTTATAAATAACGGCTCAA	58	intron 3, exon 4, intron 4 / 497	Sequencing
Ex5	GCAAAATTTGGCCTGAAAAG	TTTCCGAGGATTCAGAGCC	54	intron 4, exon 5, intron 5 / 459	Sequencing
Ex6	TTCCCTGTTCTGCTTTATGAAC	GACCTGGCATGCATTGAGTT	57	intron 5, exon 6, intron 6 / 461	Sequencing
Ex7_1	GTCAATGTAGTCAACAAAAATTCAA	GCTTCAGGCATATTTTCCAGA	56	intron 6, first part of exon 7 / 488	Sequencing
Ex7_2	ACAAAACCACCTGGTTGGAA	TACAGGGGAGACCTGGAAAA	57	part of exon 7, 3' UTR / 495	Sequencing
Ex2_mutation	TATTTCCATTTATAACTACTTTGTTCG[3]	TCCTTTACAGACGAGATGCC[4]	55	part of exon 2, intron 2 / 214	PCR-RFLP (*Taq*I) of the g.41360196G>A SNP[5]

[1] Annealing temperature in °C. PCR was carried out using a 2720 thermal cycler (Life Technologies, Carlsbad, CA, USA) in a 20 μL reaction volume containing ~50 ng genomic DNA, 1 U DNA EuroTaq DNA polymerase (EuroClone Ltd., Paington, Devon, UK), 1X Euro Taq PCR buffer, 2.5 mM dNTPs, 10 pmol of each primer and 2.5 mM of $MgCl_2$. PCR profile was as follows: 5 min at 95 °C; 35 amplification cycles of 30 sec at 95 °C, 30 sec at the specific annealing temperature for each primer pair, 30 sec at 72 °C; 5 min at 72 °C.

[2] Fragment length in bp.

[3] Underlined nucleotide is a mismatched base created to insert an artificial restriction site for *Taq*I.

[4] Reverse primer is the same of Ex2 primer pair.

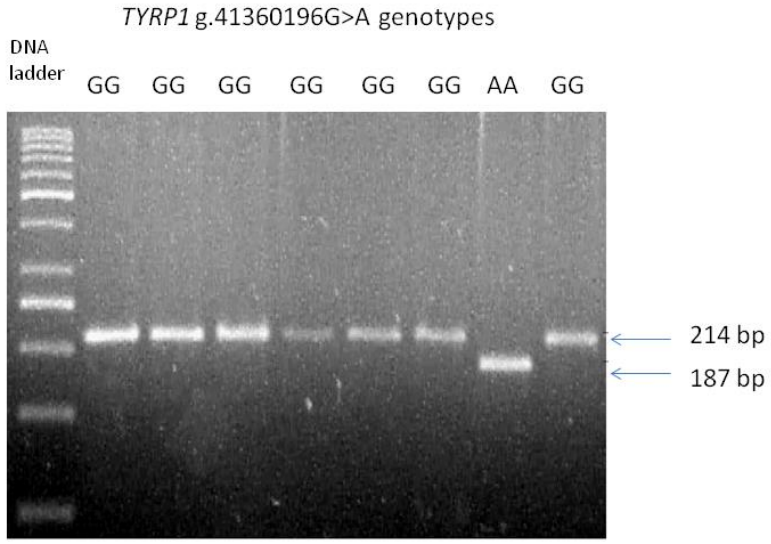
[5]Amplified DNA fragment (5 μL of PCR product) was digested overnight at 65 °C with 2 U of *Taq*I restriction enzyme (Thermo Scientific - Fermentas, Vilnius, Lithuania) in a total of 25 μL of reaction volume with 1X reaction buffer. Resulting DNA fragments were electrophoresed in TBE 1X 2.5% agarose gels. DNA bands were visualized with 1X GelRed Nucleic Acid Gel Stain (Biotium Inc., Hayward, CA, USA). Allele G resulted in an undigested fragment of 214 bp, whereas allele A was detected by the occurrence of two fragments of 187 and 27 bp (Fig. S1).

Table S2. Polymorphisms identified in the rabbit *TYRP1* gene.

Polymorhisms[1]	**Type of polymorphism**	**TYRP1 gene region[2]**	**Haplotype in oryCun2.0**	**Haplotype in wild rabbits[3]**	**Haplotype in Belgian hare[3]**	**Haplotype in Chinchilla rabbits[3]**	***Brown* haplotype (Havana rabbits)**	**Effect**
g.41361171T>C	SNP	Exon 1	T	T	C	T	C	Synonymous
g.41361090G>A	SNP	Intron 1	G	G	G	G	A	-
g.41361088C>T	SNP	Intron 1	C	C	C	C	A	-
g.41359983C>T	SNP	Exon2	C	T	C	C/T	C	Synonymous
g.41360196G>A	**SNP**	**Exon 2**	**G**	**G**	**G**	**G**	**A**	**Nonsense mutation (a stop codon is created); p.W190ter**
g.41356777C>T	SNP	Intron 4	C	T	C	C	T	-
g.41353202G>A	SNP	Intron 3-4	G	A	A	G	A	-
g.41352892C>T	SNP	Intron 4	C	C	C/T	C	C	-
g.41350144T>G	SNP	Intron 4	T	G	G	T/G	G	-
g.41350127A>G	SNP	Intron 4	A	G	G	A	A	-
g.41350092A>G	SNP	Intron 4	A	G	G	G/A	G	
g.41350041A>G	SNP	Exon 5	A	A/G	A	A/G	A	Synonymous
g.41349875G>A	SNP	Intron 5	G	A	G	G	G	-
g.41347640A>T	SNP	Intron 5	A	T	A	A	A	-
g.41347518C>T	SNP	Exon 6	C	T	C	C/T	T	Synonymous
g.41347337G>A	SNP	Intron 6	G	G	G	G	A	-
g.41346626G>T	SNP	Intron 6-7	G	T	T	G/T	T	-
g.41346154A>G	SNP	3'-UTR	A	G	G	G	G	-
g.41346116A>C	SNP	3'-UTR	A	C	C	C	C	-
g.41346021A>G	SNP	3'-UTR	A	G	A	A	A	-
g.41346014A>G	SNP	3'-UTR	A	A	G	A/G	G	-
g.41345986insGGGGGCCGGCGAA	Indel	3'-UTR	-	-	insGGGGGCCGGCGAA	insGGGGGCCGGCGAA	insGGGGGCCGGCGAA	-
g.41345873G>A	SNP	3'-UTR	G	G		G/A	A	-
g.41345864G>A	SNP	3'-UTR	G	G		G	A	-

[1] Positions are relative to coordinate systems in Ensembl for the oryCun2.0 rabbit genome version.
[2] Gene regions are relative to the *TYRP1* gene entry ENSOCUG00000004221 in Ensembl oryCun2.0 genome version. No information was available to clearly define 5'-UTR and 3'-UTR regions.
[3] Not possible to establish phases for the two alleles as sequenced animals were heterozygous at several polymorphic positions. A few additional haplotypes were identified by sequencing the *TYRP1* gene in other breeds (data not shown).

Figure S1. PCR-RFLP patterns from rabbits with different genotypes (GG and AA) at the *TYRP1* g.41360196G>A mutation. L = DNA ladder; GG = homozygous *B/B* at the brown locus; AA = homozygous *b/b* at the brown locus. The fragment of 27 bp derived by the digestion of the 214 bp amplicon is not visible in this gel.



2.2 The albinism of the feral Asinara white donkeys (*Equus asinus*) is determined by a missense mutation in a highly conserved position of the tyrosinase (*TYR*) gene deduced protein

V.J. Utzeri*, F. Bertolini*, A. Ribani*, G. Schiavo*, S. Dall'Olio* and L. Fontanesi*

PUBLISHED in ANIMAL GENETICS – **DOI**: 10.1111/age.12386

* Department of Agricultural and Food Sciences, Division of Animal Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

Running title: *TYR* gene mutation and albinism in donkey

Summary

A feral donkey population (*Equus asinus*), living in the Asinara island National Park (north-west to Sardinia, Italy), includes a unique white albino donkey sub-population or colour morph that is a major attraction of this park. Disrupting mutations in the tyrosinase (*TYR*) gene are known to cause recessive albinisms in humans (Oculocutaneous Albinism Type 1 or OCA1) and other species. In this study, we analysed the donkey *TYR* gene as a strong candidate to identify the causative mutation of the albinism of these donkeys. The *TYR* gene was sequenced from 13 donkeys (7 Asinara white albino and 6 coloured animals). Seven single nucleotide polymorphisms were identified. A missense mutation (c.604C>G or p.H202D) in a highly conserved amino acid position (even across kingdoms), that disrupts the first copper binding site (CuA) of functional protein, was identified in homozygous condition (G/G or D/D) in all Asinara white albino donkeys and in the albino son of a trio (the grey parents had genotype C/G or H/D), supporting the recessive mode of inheritance of this mutation. Genotyping 82 donkeys confirmed that Asinara albino donkeys had genotype G/G whereas all other coloured donkeys had genotype C/C or C/G. Across populations association between the c.604C>G genotypes and the albino coat colour was highly significant (P=6.17E-18). The identification of the causative mutation of the albinism in the Asinara white donkeys might open new perspectives to study the dynamics of this putative deleterious allele in a feral population and to manage this interesting animal genetic resource.

Keywords: Equid; Coat colour; Pigmentation; Asinara island; Population genetics; OCA1; deleterious mutation.

Main text

Asinara (whose one of its middle age etymology seems to recall the meaning of "land of the donkeys") is a small Mediterranean island (about 52 km^2) closely located north-west to Sardinia (Fig. 1A). This island was inhabited till 1885. Then it was closed as it became a quarantine Italian site and subsequently a highly secured prisoner colony. The island was reopened to the public in 1999 after the constitution of the Asinara National Park in 1998 (http://www.parcoasinara.org: Gazzetta Ufficiale Repubblica Italiana 1997). Among the species living in the island, the Asinara white donkey (Fig. 1B) or "Asino dell'Asinara" (*Equus asinus*) is the most representative and peculiar component of the Park's fauna and is the symbol of Asinara. The origin of the Asinara white donkeys is uncertain and based only on legends from which it could be possible to date back the occurrence of the first white donkeys before the closure of the island in the XIX century (Ministero di Agricoltura, Industria e Commercio 1905; Vinceti 2007). The Asinara white donkey sub-population or colour morph (accounting ~100-120 animals) lives together with coloured (usually grey) donkeys that can be attributed to the Asino Sardo population (Pinna *et al.* 1993). All these donkeys can mate producing a hybrid population (whose number of heads is not known). The whole donkey population of the Asinara island (white and coloured) can be considered a feral population as no human direct intervention has been managing these animals for more than one century (Kugler & Broxahm 2014). The animals with white coat colour are also considered a donkey breed in a critical status by the Food and Agriculture Organization (Sherf 2000) and by the register of equine and asinine Italian local breeds (Ministero Politiche Agricole e Forestali 2009). Despite the uniqueness of the Asinara white donkeys, just few authors investigated this sub-population. As far as we know, only two studies analysed these animals at the DNA level using microsatellites to evaluate genetic variability together with other donkey breeds (Cosseddu *et al.* 2001; Colli *et al.* 2013). Pinna *et al.* (1993) described the Asinara white donkeys at the morphological level and reported that these animals resemble those of the Asino Sardo breed in terms of size and body shape, confirming their genetic closeness determined by microsatellite markers (Colli *et al.* 2013). The differentiating trait is only the complete white coat colour without pigmentation in the skin, hair, eyelashes, eyebrows and eyes that are light blue as also described for several forms of human Oculocutaneous Albinism Type 1A and 1B (OCA1A and OCA1B) defects (e.g. Grønskov *et al.* 2007; Fig. 1C). These donkeys have low visual acuity and during sunny hours they hide inside the disused buildings of the prisoner colony. These traits and their elusive behaviour to sun rays indicate that Asinara white donkeys are affected by albinism (Pinna *et al.* 1993). The albinism in these animals is one of the few cases of this type of pigmentation defect that is maintained in a wild or feral vertebrate population (Protas *et al.* 2006; Xu *et al.* 2013), as fitness

is expected to be lower, especially in a Mediterranean sunny environment.

In many different species, the *albino* locus allelic series (formally identified as the *C* locus; Searle 1968) is determined by mutations in the tyrosinase (*TYR*) gene, that lead to completely white coat colour and lack of pigmentation in case of disrupting mutations (Aigner *et al.* 2000; Oetting 2000; Beermann *et al.* 2004; Schmutz *et al.* 2004; Blaszczyk *et al.* 2005; Imes *et al.* 2006; Blaszczyk *et al.* 2007; Anistoroaei *et al.* 2008), determining the recessive *c* allele(s) (Searle 1968). Tyrosinase (EC 1.14.18.1) is the key enzyme involved in the melanogenesis process in which both melanins (eumelanims and pheomelanins) are produced. This enzyme has an active site composed by a pair of antiferromagnetically coupled copper ions, CuA and CuB, which are coordinated by six histidine residues, three per each copper binding site (Harald & Heinz 2006; Kanteev *et al.* 2015). Removal of only one of the copper-binding histidine residues results in loss of the corresponding copper ion, thereby abolishing enzyme activity (e.g. Jackman *et al.* 1991).

In this study we used a candidate gene approach to identify the causative mutation of the albinism in Asinara white donkeys. For this aim, six primer pairs (Table S1) were designed on the assembled donkey *TYR* gene (Bertolini *et al.* 2015) and used to amplify and sequence by Sanger and Ion Torrent sequencing technologies (as described in Fontanesi *et al.* 2015) all coding exons, portions of the intronic regions (downstream and upstream to the exons), 5'- and 3'- untranslated regions of the donkey *TYR* gene in 13 animals of different coat colours (7 Asinara white donkeys, expected to have *c/c* genotype at the albino locus; and 6 coloured donkeys: 2 grey Asinara donkeys, phenotypically considered as Asino Sardo donkeys; 1 Asino Sardo donkey; 1 Martina Franca; 1 Sicilian Grey; and 1 Ragusano; EMBL accession numbers LN880531 and LN880532). Seven single nucleotide polymorphisms (SNPs) were identified (Table S2). Four SNPs were in exonic regions (3 in exon 1 and 1 in exon 2) and the remaining polymorphisms were in intronic regions (2 in intron 2 and 1 in intron 4; Table S2). Of the four missense mutations, two (c.274G>A or p.V83I in exon 1 and c.987G>A or p.E316K in exon 2) were identified only in heterozygous condition in one coloured donkey (Ragusano). SIFT score (Kumar *et al.* 2009) indicated that these two amino acid substitutions are tolerated (Table S2). For the c.18G>C or p.L6F mutation, the genotype for three coloured donkeys of different breeds (Martina Franca, Grigio Siciliano and Ragusano) was G/G (L/L) whereas it was heterozygous G/C (L/F) in two grey donkeys sampled in the Asinara island (resembling Asino Sardo donkeys) and homozygous C/C (F/F) in the third grey Asino Sardo donkey sampled in Sardinia. Genotype C/C or F/F was fixed in all Asinara white albino donkeys as well (Table S2). SIFT analysis indicated that this missense mutation is not deleterious (P=0.48). The second missense mutation (c.604C>G or p.H202D; Fig. 1D) identified only in donkeys from Asinara island, that was homozygous D/D in all sequenced white donkeys, had a highly significant

SIFT score (P<0.001) supporting a deleterious effect of this substitution (Table S2). The amino acid at position 202 of the wild type TYR protein is one of the three highly conserved histidine positions of the first copper binding site (CuA) of the TYR catalytic domain (Fig. 1E). This histidine is the second copper-binding histidine residue within the CuA site (indicated as H2A) that is always present at this position in all tyrosinase protein sequences available, even across kingdoms (Fig. 1E; García-Borrón & Solano 2002; Claus & Decker 2006). The 3D structure of the wild type and mutated donkey TYR proteins obtained following the homology modelling strategy (template protein: PDB entry 4P6R of *Bacillus megaterium*; Goldfeder *et al.* 2014) with MODELLER software (version 9.14; Eswar *et al.* 2006) confirmed the disruptive effect of the p.H202D substitution (Fig. S1).

According to sequencing data, as grey donkeys sampled in the Asinara island were heterozygous at the c.604C>G (p.H202D) missense mutation, it was possible to suppose a recessive mode of inheritance of the effect of the mutated allele, as expected for mutations causing albinism (Searle 1968). Mendelian recessive inheritance of this mutation was strengthened by sequencing and genotyping (Table S1) a trio family sampled in the Asinara island composed by a grey father (genotype C/G or H/D), a grey mother (C/G or H/D) and a white albino foal (G/G or D/D; Fig. 1D).

To further confirm the role of the p.H202D substitution, the c.604C>G mutation was genotyped (Table S1) in a total of 65 donkeys (including the animals already sequenced to confirm the sequencing determined genotype) from 8 coloured breeds or populations in addition to a total of 17 feral Asinara white albino donkeys (Table 1). All Asinara white albino donkeys were homozygous for the mutated allele. Only four grey Asino Sardo donkeys were heterozygous (three sampled in the Asinara National Park, two of which already described for the sequencing, and one from a farm in the province of Sassari, in the North of Sardinia). Considering all genotyped donkeys of different breeds and populations and the occurrence of homozygous G/G (D/D) animals only in albino donkeys, across populations association between the genotype at the c.604C>G mutation and the albino phenotype was highly significant (P=6.17E-18; two tailed chi square test).

The phylogenetic tree produced including the five donkey *TYR* haplotypes (obtained from the sequenced donkeys using PHASE program v. 2.1; Stephens *et al.* 2001) and the horse sequence (Wade *et al.* 2009), generated with the UPGMA method available in the MEGA6 software (Tamura *et al.* 2013), supported the hypothesis that the albino mutation occurred in a “grey” haplotype also present in Sardinia donkeys (Fig. S2). This hypothesis might exclude the legendary origin of the white donkeys of the Asinara island that derived them from white Egyptian donkeys imported by Marchese di Mores, Duke of Asinara Island, in the XIX century, or from a wreck of a ship directed in France in the same period (Vinceti 2007).

The isolation of the Asinara donkey population and the consequent putative high inbreeding level might have been the causes of the increased frequency of the *TYR* mutated allele in the Asinara island donkey population. The presence of many small uninhabited tumbledown buildings derived by the previous destinations of the island that are used as shelters by the white donkeys during the sunniest period of the year and the low activity of these animals during the daylights might reduce the negative effects of this mutation. However, we cannot be sure whether these hypotheses are sufficient to explain the conservation of a mutation determining a potential deleterious effect in a free-living population (Page-McCaw *et al.* 2004). We did not investigate whether the albino *TYR* haplotype is in linkage disequilibrium with other variant(s) conferring advantages in a wild, marginal and harsh environment.

The identification of the causative mutation of the albinism in the Asinara white donkeys adds a new natural animal model for human OCA1 defects and might open new perspectives to study the dynamics of this putative deleterious allele in a feral population and to manage this interesting animal genetic resource that is the symbol of the Asinara National Park.

Acknowledgements

We thank Giovannatonio Pilo (veterinary of the Asinara National Park), Dr. Guy D'Hallewin (CNR Sassari), Marco Ghionda (ANAS), Asineria of Gombola, Asinara National Park personnel for their help during the sampling and Samuele Bovo and Paolo Manghi (University of Bologna) who helped with the 3D modelling. This work has been possible with the collaboration of the Asinara National Park and with the RFO2014 funds provided by the University of Bologna.

References

Aigner B., Besenfelder U., Müller M. & Brem G. (2000) Tyrosinase gene variants in different rabbit strains. *Mammalian Genome* **11**, 700-2.

Anistoroaei R., Fredholm M., Christensen K. & Leeb T. (2008) Albinism in the American mink (*Neovison vison*) is associated with a tyrosinase nonsense mutation. *Animal Genetics* **39**, 645-8.

Beermann F., Orlow S.J. & Lamoreux M.L. (2004) The Tyr (albino) locus of the laboratory mouse. *Mammalian Genome* **15**, 749-58.

Bertolini F., Scimone C., Geraci C., Schiavo G., Utzeri V.J., Chiofalo V. & Fontanesi L. (2015) Next generation semiconductor based sequencing of the donkey (*Equus asinus*) genome provided comparative sequence data against the horse genome and a few millions of single nucleotide polymorphisms. *PLoS One* **10**, e0131925.

Blaszczyk W.M., Arning L., Hoffmann K.P. & Epplen J.T. (2005) A Tyrosinase missense mutation causes albinism in the Wistar rat. *Pigment Cell Research* **18**, 144-5.

Blaszczyk W.M., Distler1 C., Dekomien G., Arning L., Hoffmann K.P. & Epplen J.T. (2007) Identification of a *tyrosinase* (*TYR*) exon 4 deletion in albino ferrets (*Mustela putorius furo*). *Animal Genetics* **38**, 421-23.

Claus H. & Decker H. (2006) Bacterial tyrosinases. *Systematic and Applied Microbiology* **29**, 3-14.

Colli L., Perrotta G., Negrini R., *et al.* (2013) Detecting population structure and recent demographic history in endangered livestock breeds: the case of the Italian autochthonous donkeys. *Animal Genetics* **44**, 69-78.

Cosseddu G.M., Fraghi A., Mura L., Carta A., Cherchi R. & Pau S. (2001) Genetic relationships among donkey populations living in Sardinia. An analysis using molecular markers. *Ippologia* **12**, 25-33.

Eswar N., Marti-Renom M.A., Webb B., *et al.* (2006) Comparative Protein Structure Modeling With MODELLER. *Current Protocols in Bioinformatics* Suppl. **15**, 5-6.

Fontanesi L., Bertolini F., Scotti E., *et al.* (2015) Next generation semiconductor based-sequencing of a nutrigenetics target gene (*GPR120*) and association with growth rate in Italian Large White pigs. *Animal Biotechnology* **26**, 92-7.

García-Borrón J.C, & Solano F. (2002) Molecular anatomy of tyrosinase and its related proteins: beyond the histidine-bound metal catalytic center. *Pigment Cell Research* **15**, 162-73.

Gazzetta Ufficiale della Repubblica Italiana (1997) Disposizioni per lo sviluppo e la qualificazione degli interventi e dell'occupazione in campo ambientale. GU n. 239 del 13-10-1997, Legge 344 art. 4, 8 Ottobre 1997.

Goldfeder M., Kanteev M., Isaschar-Ovdat S., Adir N. & Fishman A. (2014) Determination of tyrosinase substrate-binding modes reveals mechanistic differences between type-3 copper proteins. *Nature Communication* **5**, 4505.

Grønskov K., Ek J & Brondum-Nielsen K. (2007) Oculocutaneous albinism. *Orphanet Journal of Rare Diseases* **2**, 43.

Harald C. & Heinz D. (2006) Bacterial tyrosinases. *Systematic and Applied Microbiology* **29**, 3–14

Imes D.L., Geary L.A., Grahn R.A. & Lyons L.A. (2006) Albinism in the domestic cat (*Felis catus*) is associated with a *tyrosinase* (*TYR*) mutation. *Animal Genetics* **37**, 175-8.

Jackman M.P., Hajnal A. & Lerch K. (1991) Albino mutants of *Streptomyces glaucescens* tyrosinase. *Biochemical Journal* **274**, 707-13.

Kanteev M., Goldfeder M. & Fishman A. (2015) Structure-function correlations in tyrosinases. *Protein Science*, doi: 10.1002/pro.2734.

Kugler W. & Broxham E. (2014) The Ecological Value of Feral Livestock Popultions in Europe. Overview, situation and development of a network for management wild livestock populations, Final Report. SAVE-Project, St. Gallen, Switzerland.

Kumar P., Henikoff S. & Ng P.C. (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols* **4**, 1073-81.

Ministero delle Politiche Agricole, Alimentari e Forestali (2010) Registro Anagrafico delle razze equine ed asinine a limitata diffusione. Prot. 0015944, Uscita 14/07/2010.

Ministero di Agricoltura, Industria e Commercio. Divisione Industria. Sezione Pesca (1905) Bollettino Ufficiale. Nuova Serie. Italy.

Oetting W.S. (2000) The tyrosinase gene and oculocutaneous albinism type 1 (OCA1): A model for understanding the molecular biology of melanin formation. *Pigment Cell Research* **13**, 320-5.

Page-McCaw P.S., Chung S.C., Muto A., *et al.* (2004) Retinal network adaptation to bright light requires tyrosinase. *Nature Neuroscience* **7**, 1329-36.

Pinna W., Cappio Borlino A., Vacca G.M. & Lai G. (1993) Morphology of adult white donkeys of Asinara. *Bollettino della Societa Italiana di Biologia Sperimentale* **69**, 595-600.

Protas M.E., Hersey C., Kochanek D., *et al.* (2006) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics* **38**, 107-11.

Schmutz S.M., Berryere T.G., Ciobanu D.C., Mileham A.J., Schmidtz B.H. & Fredholm M. (2004) A form of albinism in cattle is caused by a tyrosinase frameshift mutation. *Mammalian Genome* **15**, 62-7.

Searle A.G. (1968) *Comparative Genetics of Coat Colour in Mammals*. Logos Press, London, UK.

Sherf B.D. (2000) *World Watch List for Domestic Animal Diversity*. 3rd Edition, FAO-UNEP. Rome.

Stephens M., Smith N.J. & Donnelly P. (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* **68**, 978-89.

Tamura K., Stecher G., Peterson D., Filipski A. & Kumar S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**, 2725-9.

Vinceti S. (2007) *Parco Nazionale dell'Asinara*. Armando, Roma.

Wade C.M., Giulotto E., Sigurdsson S., *et al.* (2009) Genome sequence, comparative analysis, and

population genetics of the domestic horse. *Science* **326**, 865-7.

Xu X., Dong G.X., Hu X.S., *et al.* (2013) The genetic basis of white tigers. *Current Biology* **23**, 1031-5.

Table 1. Distribution of c.604C>G (p.H202D) genotypes obtained from PCR-RFLP and sequencing analyses among the investigated breeds.

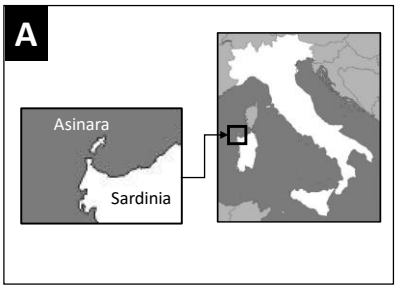
Donkey breeds/populations	**No. of donkeys**	c.604C>G genotypes[2]		
		C/C	C/G	G/G
Amiata (coloured)	2	2	-	-
Asinara (white albino)	17	-	-	17
Asino Sardo (coloured)[1]	7	3	4	-
Coloured hybrids (coloured)	13	13	-	-
Martina Franca (coloured)	13	13	-	-
Pantesco (coloured)	1	1	-	-
Ragusano (coloured)	19	19	-	-
Romagnolo (coloured)	2	2	-	-
Sicilian Grey (coloured)	8	8	-	-
Total	82	61	4	17

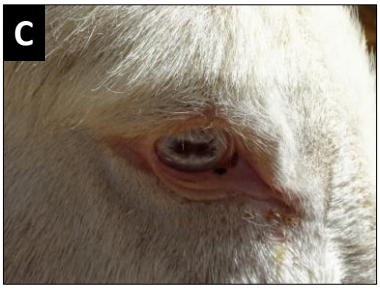
[1] Including also grey donkeys sampled in the Asinara island (see text for details).

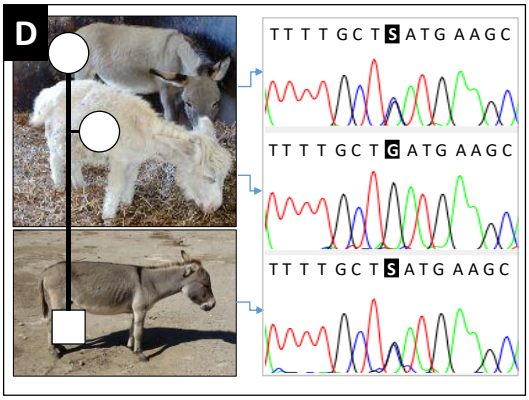
[2] The number of donkeys with the corresponding genotype is reported. Genotypes are indicated for the c.604C>G single nucleotide polymorphism: allele C corresponds to the deduced amino acid H and allele G corresponds to the deduced amino acid D for the missense mutation indicated as protein position (p.H202D).

Figure 1. Geographical position of the Asinara island, phenotypic details of white Asinara donkeys and the causative mutation determining their albinism.

A) The geographical location of the Asinara island. **B)** Asinara white albino donkey. **C)** A particular of the depigmented eye of an Asinara white albino donkey. **D)** Recessive Mendelian inheritance of the albino phenotype demonstrated in a trio. Two grey parents (with heterozygous genotype H/D at the p.H202D site or C/G in the nucleotide sequence at the c.604C>G nucleotide position, indicated with S according to the IUPAC nomenclature) gave birth to an albino donkey (D/D genotype or G/G at the nucleotide position). Microsatellite analysis (data not shown) confirmed the relationship among the three donkeys. **E)** Alignment of the donkey tyrosinase protein region containing the p.H202D substitution with the corresponding region in different species. The grey region indicated with an arrow corresponds to the position of the pH202D substitution in donkeys (H2A position in the CuA site). The other arrow indicates the histidine of the H3A position in the CuA site (the H1A position is not included in this alignment). Protein accession numbers for the sequences used in the alignment are: *Equus caballus*, F6YIA2; *Homo sapiens*, P14679; *Mus musculus*, P11344; *Bos taurus*, Q8MIU0; *Oryctolagus cuniculus*, G1SYA0; *Gallus gallus*, P55024; *Xenopus laevis*, F7CL37; *Danio rerio*, F1QDZ4; *Ipomoea batatas*, Q9MB14; *Neurospora crassa*, P00440; *Bacillus megaterium*, B2ZB02. Numbers in the alignments indicate the starting and ending amino acid residues of the corresponding protein.







E

H2A H3A

Equus asinus wild type	193	EVWKNIDFAHEAPGFLPWHRVFLLLWEQEIQK	228
Equus asinus albino	193	.........D.....	228
Equus caballus	193	..........S.....	228
Homo sapiens	193	.I.RD.....A.....L...R.....	224
Mus musculus	193	.I.RD.....L.....RE	224
Bos taurus	193	...RD.....L.....	224
Oryctolagus cuniculus	193	...RD.....L.....	224
Gallus gallus	193	N..RD.....A.....R....	224
Xenopus laevis	197	AL.RD.....A.V....Y...H..H....	228
Danio rerio	193	N..AD.....SAA.....Y..F..H..R.	225
Neurospora crassa	88	SSGFGGYCT.SSIL.IT...PY.A.Y..ALYA	119
Ipomoea batatas	189	DYPDKEIQV.NSWL.F.F..WY.YFY.RILG.	220
Bacillus megaterium	51	PPGSDRNA..MSSA.....EY..RF.RDL.S	81

Supplementary Material

Figure S1. 3D modelled structure of the A) wild type (allele H – His at position 202) and B) mutated (allele D – Asp at position 202) TYR proteins in the CuA and CuB copper binding sites. The changed residue is evidenced. MODELLER software (version 9.14; Eswar *et al.* 2006) was used to build the 3D structure of the wild type and mutated donkey TYR proteins following the homology modelling strategy and selecting the model with the lowest objective function. The template protein (PDB entry 4P6R of *Bacillus megaterium*; Goldfeder *et al.* 2014) was selected with BLASTP (http://www.uniprot.org/blast/) on UNIPROT database (release March 26, 2015) and aligned with the algorithm for local alignment on the LALIGN server (http://www.ebi.ac.uk/Tools/psa/lalign/). The computed model had a sequence identity of 30.69% (derived from structural superimposition) and 0.065-nm root mean square deviation (RMSD) to the selected template. The resultant model was used as the template for the mutated variant protein with 0.021 nm RMSD to the wild type model. The PROCHECK validation (Ramachandran Plot) gave 98.8% residues in favoured regions when the wild type protein was used to reconstruct the mutated protein, indicating that the prediction was reliable.

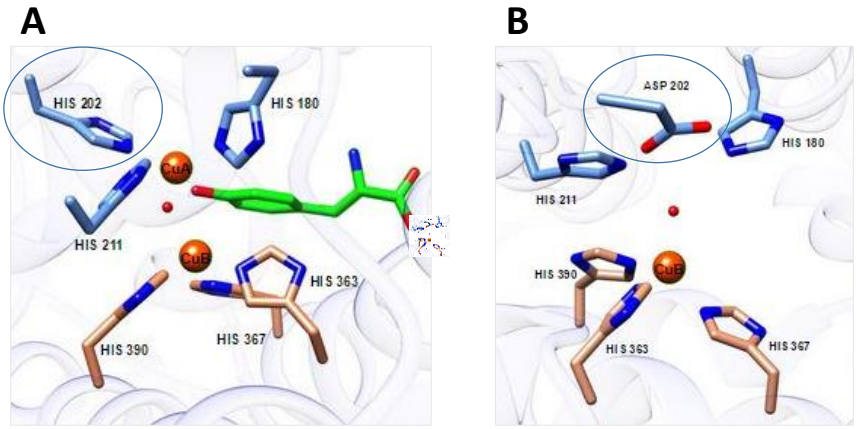


Figure S2. Phylogenetic tree of the donkey *TYR* gene haplotypes. Haplotypes are named according to the breed of the donkey from which it was sequenced. In the case of Martina Franca haplotype, the same haplotype was also identified in Sicilian Grey and Ragusano. Ragusano haplotype was only identified in the sequenced animal of this breed. Three haplotypes from grey animals (resembling Asino Sardo donkeys) were identified. Haplotypes are reported with the informative nucleotide positions. The position determining the p.H202D is evidenced in the haplotype. Numbers at the nodes indicate bootstrapping frequency. The horse sequence has been included in this tree (additional single nucleotide variations for this sequence are not shown).

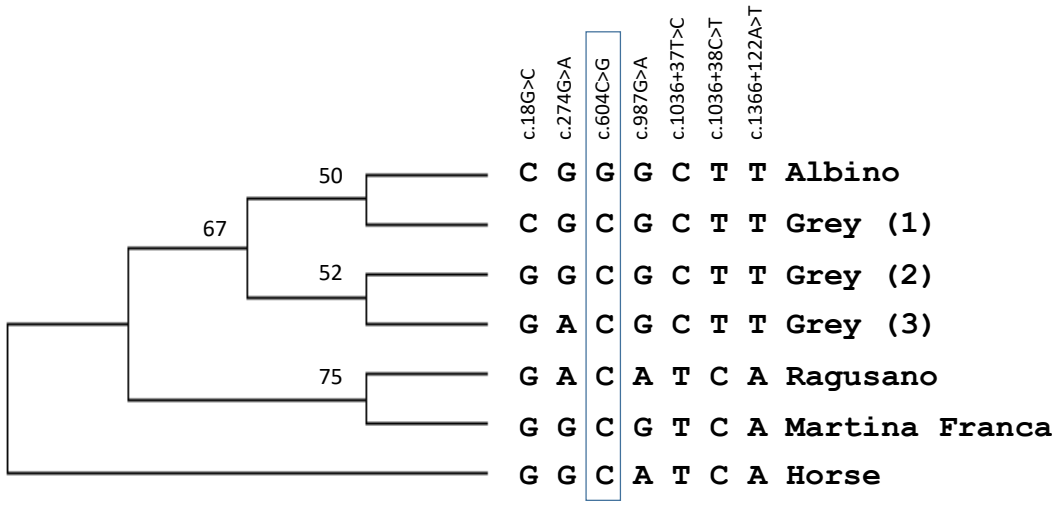


Table S1. PCR primers used in this study, sequencing and genotyping.

Primer pair name[1]	Forward primer (5‘-3‘)	Reverse primer (5‘-3‘)	Amplified region	Primer position[2]	T[3]	Amplified region size (bp)	Use[5,6]
TYR1a	GTCTTAGCCAAAACATGTGATA	TTTGTTCTTCTCTGGAACACTC	5'-flanking region, 5‘-UTR and part of CDS of Exon 1	F: 57115465-57115486 R: 57114989-57115010	56	498	Sequencing
TYR1b[4]	ACGACTCTTGGTGAGAAGAAG	GCTATTGTAGGTCCAGCAGAC	Part of Exon 1 and part of intron 1	F: 57115023-57115043 R: 57114543-57114563	57.5	501	Sequencing, PCR-RFLP
TYR2	CAATCCAAATATGGTAATCAGG	ATATGCCTCATGTCCTTACTGA	Part of intron 1, Exon 2 and part of intron 2	F: 57106418-57106439 R: 57105991-57106012	58	449	Sequencing
TYR3	TGCAGTTGTAAACCAACTAAGA	GTTTCTCTCGCATAAAACATCT	Part of intron 2, Exon 3 and part of intron 3	F: 57078708-57078729 R: 57078323-57078344	54	407	Sequencing
TYR4	CTGAATCTGAATAATCCTTTCCT	ATGTTAAGCCGCGTAGATATAA	Part of intron 3, Exon 4 and part of intron 4	F: 57033738-57033759 R: 57033352-57033373	54	409	Sequencing
TYR5	TGGTGATGACGATGATAGTAAA	GGAAAACTGTCAAAAAGATGAA	Part of intron 4, CDS of Exon 5 and 3'-UTR	F: 57024265-57024286 R: 57023793-57023814	53	494	Sequencing

[1] PCR primers were designed on the donkey *TYR* gene sequence, assembled by aligning Ion Proton reads (Bertolini *et al.* 2015) on the horse *TYR* gene sequence (Ensembl accession number ENSECAG00000016425) available in the EcuCab2.0 horse genome version (Wade *et al.* 2009), as previously described (Bertolini *et al.* 2015).

[2] Start and end positions of the forward (F) and reverse (R) primers on chromosome 7 of the EquCab2.0 horse genome.

[3] Annealing Temperature in the PCR (in °C). DNA was extracted from hair roots using the Wizard® Genomic DNA Purification Kit (Promega Corporation, Madison, USA). PCRs were cycled in a 2720 Life Technologies thermal cycler (Life Technologies, Foster City, CA, USA) with the following profile: 5 min at 95 °C; 35 amplification cycles of 30 sec at 95 °C, 30 sec at the appropriate annealing temperature (Table S1), 30 sec at 72 °C; 10 min at 72 °C. The final reaction volume was of 20 µL and included: about 50 ng of template DNA; the Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, UK); 10 pmol of each primer.

[4] Primer pair TYRb was used to sequence the region including the albino mutation in the donkeys of the trio.

[5] Sequencing. For Asinara white and Asino Sardo donkeys, obtained PCR products were purified with ExoSAP-IT (USB Corporation, Cleveland, Ohio, USA) and then sequenced using the Sanger method with the BrightDye® Terminator Cycle Sequencing Kit (NIMAGEN, Nijmegen, the Netherlands). Sequencing reactions were loaded on an ABI3100 Avant sequencer (Applied Biosystems). All sequences were visually inspected and aligned with the help of CodonCode Aligner (version 5.1.5) software (http://www.codoncode.com/aligner) and assembled with MEGA6 software (Tamura *et al.* 2013). PCR fragments obtained from the other donkeys were sequenced using the Ion Torrent Personal Genome Machine (PGM) (Life Technologies, together with many other PCR products of other origin (data not shown). Amplified products were treated with ExoSAP-IT (USB Corporation, Cleveland, Ohio, USA) and equimolar pools of amplicons for each animal were obtained in order to construct libraries. For each library, 200 ng of amplified DNA was fragmented using restriction enzymes and ligated with a specific barcode using the Ion Xpress™ Plus Fragment Library and Ion Xpress™ Barcode Adapters 1-16 kits (Life Technologies). Then, the protocol was carried out according to the instruction for the Ion Torrent Personal Genome Machine (Life Technologies) sequencing, using the Ion One Touch™ 200 Template Kit (Life Technologies), with the Ion PGM Sequencing 200 Kit (Life Technologies) and an Ion 314 Chip (Life Technologies) as previously detailed (Fontanesi *et al.* 2015).

[6] Genotyping. The missense mutation (c.604C>G) identified by sequencing was genotyped by PCR-RFLP. Briefly, genomic DNA from individual samples was amplified with primer pair TYRb and the obtained DNA fragment (5 μL of PCR product) was digested overnight at 37 °C with 2 U of *Pag*I restriction enzyme (Thermo Scientific - Fermentas, Vilnius, Lithuania) in a total of 20 μL of reaction volume with 1X reaction buffer. Resulting DNA fragments were electrophoresed in TBE 1X 2.5% agarose gels. DNA bands were visualized with 1X GelRed Nucleic Acid Gel Stain (Biotium Inc., Hayward, CA, USA). Allele G resulted in an undigested fragment of 501 bp, whereas allele C was detected by the occurrence of two

fragments of 259 and 242 bp.

Table S2. *TYR* gene polymorphisms and genotypes of the sequenced donkeys.

Polymorphisms	**Gene region**	**Effect (SIFT)**	**Coloured donkeys[1]**						**Asinara white albino donkeys[1]**							**Horse[2]**
			Martina Franca	Grigio Siciliano	Ragusano	Grey (n. 18)	Grey (n. 20)	Grey (n. 22)	n. 5	n. 8	n. 10	n. 11	n. 15	n. 16	n. 17	
c.18G>C	Exon1	p.L6F (0.48)	G/G	G/G	G/G	G/C	C/C	G/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	g.57115370G
c.274G>A	Exon1	p.V83I (0.32)	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	g.57115141G
c.604C>G	Exon1	p.H202D (0.00)	C/C	C/C	C/C	C/G	C/C	C/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	g.57114784C
c.987G>A	Exon 2	p.E316K (0.29)	G/G	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	g.57106216A
c.1036+37T>C	Intron 2	-	T/T	T/T	T/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	g.57106089T
c.1036+38C>T	Intron 2	-	C/C	C/C	C/C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	g.57106088C
c.1366+122A>T	Intron 4	-	A/A	A/A	A/A	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	g.57033427A

[1] The laboratory number of the sequenced animals is reported. Grey n. 20 is a grey Asino Sardo donkey sampled in Sardinia. Grey n. 18 and 22 were sampled in the Asinara island.

[2] The corresponding nucleotide position in the chromosome 7 of the EquCab2.0 horse genome is reported.

2.3 Differentiation of meat from European wild boars and domestic pigs using polymorphisms in the *MC1R* and *NR6A1* genes

L. Fontanesi[a,*], A. Ribani[a], E. Scotti[a], V.J. Utzeri[a], N. Veličković[b], S. Dall'Olio[a]

PUBLISHED in MEAT SCIENCE – **DOI**: 10.1016/j.meatsci.2014.07.026

[a] Department of Agricultural and Food Sciences, Division of Animal Sciences, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

[b] Department of Biology and Ecology, Faculty of Sciences, University of Novi Sad, Trg Dositeja Obradovica 2, 21000 Novi Sad, Serbia

Running title: SNPs for wild boar and domesticated pig meat

Abstract

Wild boar meat cannot be easily distinguished from domestic pig meat, especially in processed products, thus it can be fraudulently substituted with cheaper domestic pork. In this study we genotyped polymorphisms in two genes (*MC1R*, affecting coat colour and *NR6A1*, associated with number of vertebrae) in 293 domestic pigs of five commercial breeds, 111 wild boars sampled in Italy, and 90 in Slovenia and other Western Balkan regions. Allele and genotype frequency data were used to set up a DNA-based method to distinguish meat of wild boars and domestic pigs. Genotyping results indicated that domesticated genes were introgressed into wild boar populations. This complicated determination of the origin of the meat and caused a high error rate if markers of only one gene were used. The combined use of polymorphisms in the two analysed genes substantially reduced false negative results.

Highlights

►Wild boar meat represents an increasing niche market. ►DNA-based systems are needed to distinguish wild boar meat from domestic pig meat. ►Introgression of domesticated genes into wild boar genomes complicates the question. ►We set up an efficient method based on the analysis of *MC1R* and *NR6A1* polymorphisms.

Key words: authentication; coat colour; pork; SNP; *Sus scrofa*; vertebral number.

1. Introduction

Assuring authenticity of food is becoming a key issue in most production chains and food niche markets. For meat products, the species of origin and, within species, the breed of origin of mono-breed labeled products represent two levels of authentication that are required to protect consumers and producers from fraud (Ballin, 2010; D'Alessandro et al. 2007; Fontanesi, 2009). Financial incentives drive fraudulent substitution of products of lower value for meat of higher value. For example, meat of game animal species that is in generally considered a delicacy is sold at a higher price compared to meat coming from related domesticated species. Among game species, wild boar meat represents an increasing niche market that derives from hunting of the expanding natural populations of wild boars throughout Europe and from farming of wild boars. Wild boar meat is a premium product that usually has lower intramuscular fat content, higher flavor, and darker color than meat from domestic pigs (Sales, & Kotrba, 2013). However, wild boar meat can be fraudulently substituted with cheaper domestic pork because, especially in processed products, the two meats cannot be easily distinguished.

Quite a large number of DNA-based methods have been developed for the authentication of meats of game animal species (reviewed in Fajardo, González, Rojas, García, & Martín, 2010). A few methods have also been proposed to differentiate wild boar meat from pork of domestic pigs using microsatellites (Conyers, Allnutt, Hird, Kaye, & Chisholm, 2012), mitochondrial DNA markers (Alves et al. 2009), a single nucleotide polymorphism (SNP) in the melanocortin 1 receptor (*MC1R*) gene (Kijas et al. 1998; Mayer, & Hochegger, 2011), or a combination of mitochondrial DNA polymorphisms and SNPs in the *MC1R* gene (Fajardo et al. 2008). The use of microsatellites relies on a probabilistic assignment of a meat product to wild boar or domestic pigs if a reference dataset, including a large number of domestic breeds genotyped for a panel of microsatellite markers, is previously constructed (Conyers et al. 2012). Despite the acceptable level of assignment for single blind samples, this method is not precise when mixtures of pork and wild boar meat are analyzed (Conyers et al. 2012), and might not be easily implemented as a routine method because of

difficulties in reading microsatellite profiles. Mitochondrial D-loop sequences did not allow direct discrimination between wild boar and domestic pig meats (Fajardo et al. 2008). Other mitochondrial DNA markers identified by Alves et al. (2009) provided information about the maternal origin of samples from Iberian pigs and Iberian wild boars but they were not useful for authentication purposes.

The *MC1R* gene, located on porcine chromosome 6, is responsible for the *Extension* coat color locus in *Sus scrofa.* Several *MC1R* mutations are associated with different coat colors in this species (Kijas et al.. 1998; Kijas, Moller, Plastow, & Andersson, 2001). The wild type allele (E^+), carried by European wild boars, is determined by a haplotype with a unique sequence derived by the combination of a few nucleotide differences from other *MC1R* alleles that cause different coat color phenotypes (alleles E^{D1} and E^{D2} reported in pigs with black coat color; E^P reported in white and spotted pigs; and *e*, the recessive red allele, fixed in red breeds, such as Duroc). A few diagnostic sites in the *MC1R* gene can be used to distinguish different alleles and, in turn, to identify the breed/subspecies of origin of meat samples (D'Alessandro et al. 2007). However, the diagnostic power of the wild type *MC1R* allele (E^+) is not absolute because some wild boars have other alleles at the *MC1R* locus rather than being homozygous for the E^+ allele. This is probably due to crossbreeding with domestic pigs and/or introgression of domesticated genetic pools (Babicz, Pastwa, Skrypczak, & Buczyński, 2013; Koutsogiannouli, Mountou, Sarafidou, Stamatis, & Mamuris, 2010).

Another phenotypic trait that differentiates wild boars and European commercial pigs is the vertebral number. Wild boars have 19 thoracic and lumbar vertebrae whereas European commercial pigs have 21-23 vertebrae (King & Roberts, 1960). Recently, Mikawa et al. (2007) suggested that a missense substitution (p.Pro192Leu) in the nuclear receptor subfamily 6, group A, member 1 (*NR6A1*) gene, mapped on porcine chromosome 1, is the causative mutation of a QTL affecting the number of vertebrae in pigs. Wild boars might carry only the wild type allele (p.192Pro) whereas commercial pigs may be fixed for the mutated allele (Rubin et al. 2012). Therefore this

polymorphism could potentially be useful to distinguish meat of wild boars from meat of commercial pigs.

In this study we investigated *MC1R* and *NR6A1* gene polymorphisms in domestic pigs of different breeds and wild boars sampled in Italy, Slovenia and Western Balkan regions and obtained allele and genotype frequency data that were used to set up a DNA-based method to distinguish meat coming from wild boars and domestic commercial pigs.

2. Materials and methods

2.1. Samples

A total of 494 animals were sampled and genotyped. Of these animals, 201 were wild boars: 111 from the Emilia Romagna region (Northern Italy, representing a South Central European region); 90 from Slovenia (n = 15) and six Western Balkan regions (Bosnia, n = 8; Croatia, n = 10; Former Yugoslav Republic of Macedonia - FYRM, n = 16; Herzegovina, n = 8; Montenegro, n = 16; Serbia, n = 17), representing South East European samples. Collection of wild boar samples was opportunistic, derived by hunting or via inspection of carcasses/hunted animals by forest policemen, and without bias towards locations, sex and age of the animals. The remaining 293 animals were from 5 commercial domestic pig breeds (Italian Large White, n = 101; Italian Landrace, n = 59; Italian Duroc, n = 72; Belgian Landrace, n = 11; Piétrain, n = 50).

2.2. DNA marker analyses

Genomic DNA was extracted from blood, hair root, muscle or ear tissues using a standard phenol-chloroform protocol (Sambrook, Fritsch, & Maniatis, 1989) or the Wizard® Genomic DNA Purification kit (Promega Corporation, Madison, WI, USA). PCR primers used to amplify porcine *MC1R* and *NR6A1*gene regions are reported in Table 1. Two DNA fragments were amplified for the *MC1R* gene. One *MC1R* fragment of 196 bp (from positions 214 to 409 of the coding region) included a diagnostic site (c.367G>A) that differentiates allele E^+ from alleles E^{D2} and E^P, but not

from allele E^{D1}. This polymorphism was analysed by PCR-RFLP with the restriction enzyme *Bsp*HI (recognized sequence: TCATGA) that cuts alleles E^{D2} and E^{P} whereas the fragments obtained from alleles E^{+} and E^{D1} are undigested (Table 1). The second *MC1R* fragment of 154 bp (from positions 619 to 772 of the coding region) contains two polymorphic sites (c.727G>A and c.729G>A) that can distinguish all alleles from allele *e*, and all alleles from alleles *e* and E^{D1}, respectively. These two polymorphic sites can be analysed by digesting the amplified fragment with *Hha*I (recognized sequence: GCGC; allele *e* is not digested, whereas all other alleles are digested into two fragments) and *Bst*UI (recognized sequence: CGCG; alleles E^{D1} and *e* are not digested, whereas all other alleles are digested into two fragments) (Table 1). The combination of the three genotyped *MC1R* single nucleotide polymorphisms (SNPs) makes it possible to identify alleles E^{+}, E^{D1} and *e*, whereas alleles E^{P} and E^{D2} cannot be separated from each other and henceforth are indicated as one allele E^{PD2}. For the purpose of this study, it was not necessary to distinguish these two alleles of domesticated pigs. The *NR6A1* polymorphism (g.299084751C>T or p.Pro192Leu) was genotyped by PCR-RFLP using the restriction enzyme *Msp*I that cuts the amplified fragment when the wild type allele is present (Table 1).

PCR was carried out in a 20 µL reaction volume containing ~50-100 ng genomic DNA, 1 U DNA EuroTaq DNA polymerase (EuroClone Ltd., Paignton, Devon, UK), 1X Euro Taq PCR buffer, 2.0 mM dNTPs, 10 pmol of each primer and 2.5-3.0 mM of $MgCl_2$ (Table 1). A 2720 thermal cycler (Life Technologies, Carlsbad, CA, USA) was used for amplification reactions with the following temperature profile: 5 min at 95 °C; 35 amplification cycles of 30 sec at 95 °C, 30 sec at 55 or 61 °C (Table 1), 30 sec at 72 °C; 5 min at 72 °C. Obtained PCR products were digested overnight at 37 °C with the specific restriction enzyme (Table 1) and then were electrophoresed on 3.0-3.5% agarose gels and visualized with 1X GelRed Nucleic Acid Gel Stain (Biotium Inc., Hayward, CA, USA).

2.3. Data analysis

Allele and genotype frequencies were calculated for each breed/population at the two investigated loci. Hardy–Weinberg equilibrium was evaluated with the HWE software program (Linkage Utility Programs, Rockefeller University, New York, NY). Genotypic disequilibrium between the two investigated loci was calculated using FSTAT v. 2.9.3 (Goudet, 1995). Probability to incorrectly assign an unknown meat sample to domestic pigs (error rate) was calculated using the frequency of occurrence of wild boars carrying domestic alleles, considering one locus, the other, or both loci.

3. Results and discussion

3.1. Allele and genotype frequencies of the *MC1R* gene

Genotyping results for *MC1R* polymorphisms in wild boars sampled in South Central and South East European regions showed that not all animals were homozygous for the wild type E^+ allele (Table 2). However, this allele was the most frequent in the wild boars sampled in both regions (0.867 in South East Europe and 0.905 in South Central Europe). The E^{PD2} allele(s) was detected in both regions (0.133 and 0.063, respectively). Two other alleles, *e* and E^{D1}, were detected in Italian wild boars (0.027 and 0.005, respectively). All wild boars analysed were carriers of at least one E^+ allele except for two Italian wild boars (genotype E^{PD2}/e) and one animal sampled in South East Europe (genotype E^{PD2}/E^{PD2}). Both wild boars populations were in HWE at the *MC1R* polymorphic locus (P>0.05).

Allele E^+ was not detected in any investigated domestic pigs of commercial breeds. Domestic pigs of the different breeds had their expected *MC1R* genotype (Fontanesi et al. 2010): all pigs of white (Italian Large White, Italian Landrace and Belgian Landrace) and spotted (Pietrain) breeds had genotype E^{PD2}/E^{PD2}; all Italian Duroc had genotype *e/e*.

These results indicate that introgression of domesticated genes occurred in different wild boars populations modifying the genetic structure of these populations that originally might have

been fixed for the wild type E^+ allele. Introgression of domesticated *MC1R* alleles in wild boars has been previously reported in other European populations (Koutsogiannouli et al. 2010; Frantz et al. 2013). Wild boars that were sampled for this study were described as having all features that characterize European wild boars.

3.2. Allele and genotype frequencies of the *NR6A1* gene

All pigs of the five commercial breeds studied were homozygous for the *NR6A1* g.299084751T (p.192Leu) allele that is considered a causative mutation for the increased number of vertebrae in domestic pigs (Mikawa et al. 2007). Most of wild boars were homozygous for the wild type allele g.299084751C (p.192Pro) (Table 3). However, a few wild boars carried the mutated allele: one and ten animals sampled in South East Europe were homozygous T/T and heterozygous C/T, respectively; four wild boars sampled in Italy had genotype C/T (Table 3). Therefore, the frequency of the domesticated allele in the wild boar populations sampled in the two European areas was 0.034. Results obtained for the *NR6A1* gene confirm to some extent results obtained by genotyping *MC1R* polymorphisms in wild boars; introgression of domesticated genes in European wild boars occurred even if the frequency of the domesticated allele in wild boars is lower for the *NR6A1* gene than that for the *MC1R* gene. Both wild boars populations were in HWE at the *NR6A1* polymorphic site ($P>0.05$).

3.3. Usefulness of *MC1R* and *NR6A1* polymorphisms to differentiate wild boar and domestic pig meat

In this context, a DNA-based test could be useful if it can differentiate not only meat of pure wild boars from meat of pure domestic pigs, but also meat obtained from F1 animals derived by crossbreeding of wild boars with domestic pigs. This is a common practice used to increase production and performance traits, compared to pure wild boars, while maintaining meat quality traits similar to those of wild boars (Razmaite, Kerziene, & Jatkauskiene, 2009). According to the

genetic structure of the investigated wild boar populations, genetic tests based on the analysis of *MC1R* polymorphisms that have been already proposed for this aim (Kijas et al. 1998; Fajardo et al. 2008; Mayer & Hochegger, 2011), might give a high number of false negatives if we would use this method to identify true F1 animals and attribute a deterministic value to the test. The high error rate related to the strict interpretation of the results obtained with *MC1R* polymorphisms derives from the high frequency of wild boars (20.9%) carrying not only the E^+ allele but also other domestic alleles at this locus (Table 2). In addition, it would be quite difficult to interpret results in the rare cases in which wild boars, or animals classified as wild boars for their features, carry only domesticated alleles (3 out of 201 genotyped animals; Table 2). Other investigations of the level of introgression of domesticated genes in wild boar populations sampled in Greece and Northern Europe came to the same conclusion (Koutsogiannouli et al. 2010; Frantz et al. 2013).

For all these reasons, it is important to analyse other DNA markers and to set up simple and useful DNA-based tests that may add more information. Mitochondrial DNA polymorphisms are not completely informative for the aim of distinguishing European wild boars from domestic pigs (Fajardo et al. 2008; Alves et al. 2009). Therefore we considered a causative mutation for the number of vertebrae in *Sus scrofa* (*NR6A1* g.299084751C>T or p.Pro192Leu) that has been selected during the domestication process in this species. The domestication process selected animals with favorable phenotypic traits, such as in this case, an increased length that in turn is associated with increased meat content and reproduction performances of the animals (Mikawa et al. 2007; Rubin et al. 2012). We found that 7.46% of European wild boars carried at least one copy of the mutated domesticated allele and the observed total frequency was 0.034 (Table 3).

None of the analysed wild boars carried domesticated alleles at both investigated genes (*MC1R* and *NR6A1*) that were not in genotypic disequilibrium (P=0.620, s.e.= 0.014). Therefore, the combined use of the two genes can be further used to refine results obtained from one or the other gene.

Considering only the *MC1R* gene, the error rate (wild boars attributed erroneously to F1 hybrids of wild boars x domestic pigs) can be calculated using the frequency of the wild boars carrying domesticated alleles (42/201 = 0.209, summing up the results of the two European regions). The error rate was a little bit higher for the South East European wild boar population (23/90 = 0.256) than the Italian wild boars (19/111 = 0.171). Considering only the *NR6A1* polymorphism, the error rate was still important for the South East European populations (11/90 = 0.122), even though it was lower than that derived by the *MC1R* gene alone, and negligible for the Italian population (4/111 = 0.036), with an overall error rate across the two groups of 15/201= 0.075. Combinations of the two polymorphisms gave a combined error rate of 0.209 x 0.075 = 0.016, considering all samples and 0.256 x 0.122 = 0.031 and 0.171 x 0.036 = 0.006 considering genotyping results obtained from the South East and South Central European wild boars, separately.

We conclude that the combined use of the two analysed genes is more precise than the use of only *MC1R* polymorphisms. The use of only the *MC1R* gene to authenticate wild boar meat is biased by the introgression of domesticated alleles. However, error rate might depend on the level of introgression and, in turn, on the origin of wild boars (different populations) that might not have been hybridized with domestic pigs to the same degree. In addition, as the *MC1R* and *NR6A1* markers are in genotypic equilibrium, it could be possible that by chance wild boars could carry at both loci domesticated alleles with a frequency that is equal to the combined error rate reported above. Therefore, the use of additional DNA markers should further reduce the error rate. For example, the Illumina PorcineSNP60 BeadChip that can genotype more than 60,000 single nucleotide polymorphisms (SNPs) in the *Sus scrofa* genome has already been used to evaluate the level of genetic introgression of the domestic pig genome into wild boar populations, demonstrating the powerful potential of this tool (Goedbloed et al. 2013). It is more convenient for practical reasons to analyse the markers tested in our work that have been chosen to discriminate wild boars and domestic pigs. In addition to the higher cost of Illumina PorcineSNP60 BeadChip, genotyping analysis compared to the cost of the few simple PCR-RFLP tests described here, small laboratories

usually do not have direct access to high throughput genotyping platforms and a small number of samples are usually tested to discriminate the subspecies of origin of the meat.

4. Conclusions

This study confirmed that wild boar populations have been introgressed, at least in part, with domesticated genomes, complicating the possibility to authenticate the subspecies of origin of the meat (wild boars *vs* domestic pigs). Using only *MC1R* gene markers to authenticate wild boar meat can produce a high rate of false negative assignments of the meat to crossbred animals or domestic pigs. The use of an additional marker (*NR6A1* g.299084751C>T or p.Pro192Leu) that has been shown to control a phenotypic trait, that distinguishes commercial domestic pig breeds from wild boars, can improve the efficiency of the test to identify fraudulent substitution of wild boar meat with domestic pig meat obtained from these breeds. It will be interesting to determine genotype frequencies of the *MC1R* and *NR6A1* markers in local pig breeds to evaluate if the combined use of the investigated polymorphisms could be applied to distinguish wild boar meat from meat of local pig breeds. However, it could be expected that fraudulent substitution of meat declared to be of wild boar origin would likely occur with less expensive meat from commercial pig breeds. Other useful markers could be identified using high throughput genotyping platforms. These polymorphisms might be integrated with DNA markers genotyped in this study to further improve the efficiency of DNA-based tests to determine the origin of meat from the two subspecies. However, practical aspects (i.e. cost of the analyses, number of samples to be processed) should be considered if a DNA-based test is applied for this purpose.

Acknowledgements

We thank Maurizio Musolesi (Polizia Provinciale di Bologna), Savigni Silvano and several hunters for providing wild boar samples. We would like to thank Dr. Rose Mage (NIH) for English

correction of the text. This study has been funded by University of Bologna RFO program and by MiPAAF (Innovagen) project.

References

Alves, E., Fernández, A.I., Fernández-Rodríguez, A., Pérez-Montarelo, D., Benitez, R., Ovilo, C., Rodríguez, C., & Silió, L. (2009). Identification of mitochondrial markers for genetic traceability of European wild boars and Iberian and Duroc pigs. *Animal, **3***, 1216-1223.

Babicz, M., Pastwa, M., Skrzypczak, E., & Buczyński, J.T. (2013). Variability in the melanocortin 1 receptor (MC1R) gene in wild boars and local pig breeds in Poland. *Animal Genetics, **44***, 357-358.

Ballin, N.Z. (2010). Authentication of meat and meat products. *Meat Science, **86***, 577-587.

Conyers, C.M., Allnutt, T.R., Hird, H.J., Kaye, J., & Chisholm, J. (2012). Development of a microsatellite-based method for the differentiation of European Wild Boar (*Sus scrofa scrofa*) from domestic pig breeds (*Sus scrofa domestica*) in food. *Journal of Agricultural and Food Chemistry, **60***, 3341-3347.

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. *Veterinary Research Communication, **31**, Suppl. 1*, 389-392.

Fajardo, V., González, I., Martín, I., Rojas, M.A., Hernández, P.E., García, T., & Martín, R. (2008). Differentiation of European wild boar (*Sus scrofa scrofa*) and domestic swine (*Sus scrofa domestica*) meats by PCR analysis targeting the mitochondrial D-loop and the nuclear melanocortin receptor 1 (*MC1R*) genes. *Meat Science, **78***, 314-322.

Fajardo, V., González, I., Rojas, M.A., García, T., & Martín, R. (2010). A review of current PCR-based methodologies for the authentication of meats from game animal species. *Trends in Food Science & Technology, **21***, 408-421.

Fontanesi, L. (2009). Genetic authentication and traceability of food products of animal origin: new developments and perspectives. *Italian Journal of Animal Science, **8**, Suppl. 2*, 9-18.

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. *Animal Genetics*, ***41***, 478-492.

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. *Biological Journal of the Linnean Society*, **110**, 104-115.

Goedbloed, D.J., Megens, H.J., Van Hooft, P., Herrero-Medrano, J.M., Lutz, W., Alexandri, P., Crooijmans, R.P., Groenen, M., Van Wieren, S.E., Ydenberg, R.C., & Prins, H.H. (2013). Genome-wide single nucleotide polymorphism analysis reveals recent genetic introgression from domestic pigs into Northwest European wild boar populations. *Molecular Ecology, **22***, 856-866.

Goudet, J. (1995). FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *Journal of Heredity, **86***, 485-486.

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***, 779-785.

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***, 1177-1785.

King, G.W.B., & Roberts, R.C. (1960). Carcass length in the bacon pig: Its association with vertebrae numbers and prediction from radiographs of the young pig. *Animal Production*, ***2***, 59-65.

Koutsogiannouli, E.A., Montou, 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. *Mammalian Biology*, ***75***, 69-73.

Mayer, W., & Hochegger, R. (2011). Discrimination of two alleles of the melanocortin receptor 1 gene to discern European wild boar (*Sus scrofa scrofa*) and domestic pig (*Sus scrofa domestica*) in meat products by real-time PCR. *European Food Research and Technology*, ***232***, 687-692.

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 Research*, *17*, 586-593.

Razmaite, V., Kerziene, S., & Jatkauskiene, V. (2009). Body and carcass measurements and organ weights of lithuanian indigenous pigs and their wild boar hybrids. *Animal Science Papers and Reports*, **27**, 331-342.

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., & Andersson, L. (2012). Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences of the USA*, ***109***, 19529-19536.

Sales, J., & Kotrba, R. (2013). Meat from wild boar (*Sus scrofa* L.): a review. *Meat Science*, ***94***, 187-201.

Sambrook, J., Fritsch, E.F., & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. Second edition. Cold Spring Harbor Laboratory Press, NY.

Table 1. PCR primers, PCR conditions and PCR-RFLP protocols used to analyse polymorphisms in the *MC1R* and *NR6A1* genes.

Primer pair names/genes	**Primer sequences (5'-3')**	**Amplified region (length in bp)**	**PCR conditions[a]**	**PCR-RFLP[b]**
MC1R_1	Forward: CTGCACTCGCCCATGTACTA Reverse: AGCAGAGGCTGGACACCAT	196	61/3.0	*BspH*I (c.367G = 196 bp in E^{+}, E^{D1} and *e*; c.367A = 154 + 42 bp in E^{D2} and E^{P})
MC1R_2	Forward: GCGGGTACTGTACGTCCACAT Reverse: CCCAGCAGAGGAGGAAGAC	154	61/3.0	*Hha*I (c.727G = 108 + 46 bp in E^{+}, E^{D1}, E^{D2} and E^{P}; c.727A = 154 bp in *e*); *Bst*UI (c.729G = 109 + 47 bp in E^{+}, E^{D2} and E^{P}; c.729A = 154 bp in E^{D1} and *e*)
NR6A1	Forward: GGTATCCTGAGCACCCAGTC Reverse: ACCTGGAGGACAGTGTGGAG	203	55/2.5	*Msp*I (g.299084751C = 180 + 23 bp; g.299084751T = 203 bp)

[a] Annealing temperature/[$MgCl_2$].

[b] Restriction Fragment Length Polymorphism. 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.

Table 2. Allele frequencies and genotypes identified at the *MC1R* locus in wild boar populations sampled in two European areas.

Origin	**N. of animals**	**Allele frequency** [c]				**Genotypes (No. of animals)**					
		E^+	E^{PD2}	E	E^{D1}	E^+/E^+	E^+/E^{PD2}	E^+/e	E^{PD2}/e	E^{PD2}/E^{PD2}	E^+/E^{D1}
SEE [a]	90	0.867	0.133	-	-	67	22	-	-	1	-
SCE [b]	111	0.905	0.063	0.027	0.005	92	12	4	2	-	1
Total	201	0.888	0.095	0.015	0.002	159	34	4	2	1	1

[a] SEE = South East Europe (Slovenia and Western Balkan regions).

[b] SCE = South Central Europe (North of Italy).

[c] alleles E^P and E^{D2} cannot be separated from each other and henceforth are indicated as one allele E^{PD2}.

Table 3. Allele frequencies and genotypes identified at the *NR6A1* g.299084751C>T (p.Pro192Leu) polymorphism in wild boar populations sampled in two European areas.

Origin (country)[a]	**N. of animals**	**Allele frequency**		**Genotypes and n. of animals**		
		C	**T**	**CC**	**CT**	**TT**
SEE [a]	90	0.933	0.067	79	10	1
SCE [b]	111	0.991	0.009	107	4	-
Total	201	0.966	0.034	186	14	1

[a] SEE = South East Europe (Slovenia and Western Balkan regions).

[b] SCE = South Central Europe (North of Italy).

2.4 High frequency of *MC1R* and *NR6A1* domestic allele introgression in European wild boar populations: a threat for the genetic integrity of these wildlife resources?

Valerio Joe Utzeri[1§], Mihajla Djan[2§], Anisa Ribani[1§], Stefania Dall'Olio[1], Emilio Scotti[1], Leonardo Nanni Costa[1], Nevena Veličković[2], Luca Fontanesi[1*]

UNDER REVIEW

[1]Dipartimento di Scienze e Tecnologie Agroalimentari, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

[2]Department of Biology and Ecology, Faculty of Sciences, University of Novi Sad, Trg Dositeja Obradovica 2, 21000 Novi Sad, Serbia

§ These authors contributed equally to this work

Running title: Domestic *MC1R* and *NR6A1* alleles in wild boars

Abstract

Fixation of alleles during the domestication process has contributed to modify phenotypic traits that are useful to distinguish domestic pigs from wild boars. If domestic alleles are then introgressed into wild boars, they could alter relevant traits in these populations. In this study, we genotyped a total of 229 wild boars (113, 26 and 90 sampled in North of Italy, Sardinia and South-East Europe, respectively) for polymorphisms in two major genes, melanocortin 1 receptor (*MC1R*) and nuclear receptor subfamily 6 group A member 1 (*NR6A1*), that affect two domesticated traits in pigs, coat colour and number of vertebrae. Three (E^{D2}, E^{P} and *e*) and two ($E^{D2}$and E^{P}) *MC1R* domestic alleles were identified in North Italian and South-East European wild boars. The most frequent domestic alleles were E^{D2} (0.11) and E^{P} (0.05) in the two populations, respectively. No domestic *MC1R* allele was identified in the Sardinian samples. The *NR6A1* domestic allele was identified in wild boars of all three regions. The highest frequency was in Sardinia (0.12). Combining information from the two genes, signals of introgression of domestic alleles were observed in 15.58% (Northern Italy) and 34.44% (South-East Europe) wild boars. Our results showed a higher frequency of domestic alleles in European wild boars compared to previous reports, raising concerns for the genetic integrity of these wildlife resources. 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.

Keywords: Allele frequency**;** gene flow; *Sus scrofa*; coat colour; number of vertebrae

Introduction

Wild boars are the ancestors of the domestic pigs. Reproduction barriers between wild boars and domestic pigs that were introduced with the constitution of the modern breeds in Europe contributed to separate the two forms of *Sus scrofa*, at least in more advanced agricultural systems. Artificial directional selection shaped the genome of commercial pig breeds, that became cosmopolitan, increasing the frequency and, subsequently, fixing genetic variants associated to domesticated traits (e.g. coat colour, length of the animals, growth rate, reproduction performances, meat production) and losing undesired or inferior alleles (in terms of economic value for the breeder objectives), that were the common counterparts in the wild populations. In particular, coat colour has been one of the first traits that was fixed and that evidently marked domestic populations (e.g., Fang et al. 2009). Mutations in the melanocortin 1 receptor (*MC1R*) gene (the *Extension* locus), encoding for a seven transmembrane protein controlling the production of eumelanin and pheomelanin, have been associated with different coat colours in *Sus scrofa* (Kijas et al. 1998, 2001). The wild type allele (E^+, indicated also as allele *0101*; Fang et al. 2009) is the typical form in European wild boars. Several other alleles are present in domestic breeds: alleles E^{D1} (indicated as alleles *0201*, *0202* and *0203* by Fang et al. 2009) and E^{D2} (or allele *0301*; Fang et al. 2009) determine the dominant black coat colour and are of Asian and European origin, respectively; allele E^P (identified also as alleles *0501*, *0502* and *0503*; Fang et al. 2009) is usually reported in spotted and completely white pigs; allele *e* is the recessive allele determining the red coat colour of the Duroc breed and other red pigs. Polymorphisms in the *MC1R* gene have been already analysed to evaluate if European wild boar populations carry domestic alleles (Gongora et al. 2003; Koutsogiannouli et al. 2010; Frantz et al. 2012; 2013; Canu et al. 2016; Fulgione et al. 2016). Among the free ranging wild boars extensively investigated, domestic *MC1R* alleles were detected in Greece (in 5% of the investigated wild boars; 6 out of 119; Koutsogiannouli et al. 2010) and Belgium/Luxembourg (0.5%, 1 animal out of 196; Frantz et al. 2013). Other studies detected *MC1R*

alleles in wild boars sampled in different countries but the low number of investigated animals or the sampling strategy could just provide a preliminary picture of their distribution (Canu et al. 2016; Fulgione et al. 2016). Fajardo et al. (2008) used variants in this gene, together with mtDNA markers, to differentiate meat from wild boars and domestic pigs. Similarly, Fontanesi et al. (2014) used markers in the *MC1R* gene for the same purpose, coupled with another diagnostic mutation in the nuclear receptor subfamily 6 group A member 1 (*NR6A1*) gene. This mutation (p.P192L), fixed in commercial breeds for the domestic allele, is associated with an increased number of vertebrae compared to pure wild boars (21-23 *vs* 19 vertebrae) that are fixed for the wild type allele (Mikawa et al.. 2007; Rubin et al. 2012). The pig is one of the few vertebrate species for which variability in the number of vertebrae has been described (Freeman 1939). Domestic pigs have been selected for a higher number of vertebrae that is associated with increased length of the animals, more meat, higher number of teats and, in turn, increased reproduction performances of the sows compared to their wild boar relatives (Borchers et al. 2004; Mikawa et al. 2007). Both coat colour and number of vertebrae are traits that have been clearly modified during the domestication and genetic improvement processes and for which genetic bases have been reported and can be easily monitored (Fang et al. 2009; Rubin et al. 2012). A few other studies evaluated the level and extent of introgression between domestic pigs and wild boars using anonymous microsatellites, single nucleotide polymorphisms or mtDNA haplotypes (Alves et al. 2010; Scandura et al. 2011a; Frantz et al. 2012, 2013; Goedbloed et al. 2013a, 2013b). Introgression might be considered a relevant issue for wild boars, having implications in conservation genetics and population genetic monitoring of a species that in many areas is nowadays considered a pest (Scandura et al. 2011b). Introgressed domestic genes in wild boars might produce hybrid vigor and alter relevant traits (e.g. reproduction performances, behavior, disease resistance), that on one hand could have negative impacts on the related ecosystems (Gethöffer et al. 2007; Verhoeven et al. 2011; Canu et al. 2014; Fulgione et al. 2016) and on the other hand may provide the opportunity to evaluate the effect of

natural selection on artificially selected (domestic) alleles that reached wild populations through interdemic gene flow (Fulgione et al. 2016).

In this study we analysed mutations in genes (*MC1R* and *NR6A1*), that directly provide information on phenotypic traits useful to distinguish wild boars from domestic pigs, extending our previous work from a different perspective (Fontanesi et al. 2014), to evaluate the level of introgression of genes that might impact on fitness and contribute to explain the phenotypic shifts of wild boar populations.

Materials and methods

A total of 229 wild boars were sampled. Collection was opportunistic, derived by hunting or via inspection of carcasses/hunted animals by forest policemen. We received descriptive information from the hunters and forest policemen that these animals did not show any evident differences from common phenotypes. For only one wild boar white spots in the legs were reported. Of these wild boars, 139 were sampled in Italy: 113 in Emilia-Romagna [Bologna Appennini area; 110 of them already included in the study of Fontanesi et al. (2014); see also note to Table 1] and 26 in Sardinia (Cagliari and Sassari provinces) regions. These Italian populations can be considered two geographically isolated populations that were thereafter named as populations from North of Italy and Sardinia. The remaining 90 wild boars (included in the study of Fontanesi et al. 2004) were from western and central Balkan countries (8 from Bosnia, 10 from Croatia, 16 from the Former Yugoslavian Republic of Macedonia, 8 from Herzegovina, 16 from Montenegro, 17 from Serbia and 15 from Slovenia) that altogether were thereafter named as South-East (SE) European population (Figure 1).

DNA was extracted from muscle, ear cartilaginous tissues or hair roots using a standard phenol-chloroform protocols or the Wizard (R) Genomic DNA Purification kit (Promega

Corporation, Madison, WI, USA). 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 (E^+, E^{D1}, E^{D2}, E^P and *e*) described by Kijas et al.. (1998, 2001); one missense mutation in the *NR6A1* gene (g.299084751 C>T or p.P192L) that is the causative mutation of the QTL for number of vertebrae, identified on porcine chromosome 1 (Mikawa et al.. 2007).

PCR conditions and primers were as reported in Table S1.

PCR-RFLP methods were applied for genotyping *MC1R* and *NR6A1* polymorphisms, as described in Fontanesi et al.. (2014). In addition, one more fragment of *MC1R* gene that includes the indel that discriminates allele E^P from allele E^{D2} was amplified. Variability in this amplicon (Table S1) was detected by fragment length analysis on a capillary sequencer (ABI PRISM 3100 Avant Genetic Analyzer, Applied Biosystems).

Basic parameters of genetic variability (number and frequencies of alleles, observed and expected heterozigosity) and deviations from Hardy-Weinberg equilibrium were calculated using ARLEQUIN 3.5.1.2 (Excoffier & Lischer 2010). Furthermore, χ^2 test was used to evaluate differences among populations for allele and genotype frequencies distributions

Results and discussion

Four different *MC1R* alleles were detected in wild boars: E^+, E^{D2}, E^P and *e*. Figure 1 reports *MC1R* allele and genotype frequencies in the different wild boar populations. The wild type allele (E^+) was the only one detected in the Sardinian wild boars. Its frequency was 0.87 in SE Europe and 0.91 in North of Italy in which domestic alleles were identified (E^{D2} and E^P in SE Europe; E^{D2}, E^P and *e* in North of Italy). In the former population, a total of 22 animals (about 25%, mainly sampled in the Former Yugoslavian Republic of Macedonia, Serbia and Slovenia) carried at least

one domestic allele in heterozygous condition with the wild type allele (21 animals) or in homozygous condition (E^{D2}/E^{D2}) in one wild boar. The most frequent domestic allele in these animals was E^{D2} (0.11). In North of Italy, about 16% of the wild boars carried at least one domestic allele in heterozygous condition with the E^{+} allele except in two animals that had genotype E^{P}/e. The most frequent domestic allele in this population was E^{P} (0.05). The $\chi 2$ test revealed statistically significant differences in *MC1R* genotype frequencies between SE European and Northern Italian (P=0.0044) and SE European and Sardinian wild boar populations (P=0.030).

The domestic allele of the *NR6A1* g.299084751C>T polymorphism (allele T) was identified in all three wild boar populations (Figure 1). Interestingly, the higher frequency (0.12) was observed in the Sardinian population in which 5 out of 26 wild boars carried at least one domestic allele (4 in heterozygous conditions and one homozygous T/T). Among the SE European population, 10 wild boars were heterozygous C/T, one was homozygous T/T and all other animals were homozygous for the wild type allele. The North of Italy population had the lower frequency of allele T (0.02) as only four animals carried one domestic allele (genotype C/T). Comparison of allele frequencies between populations indicated a significantly lower frequency ($P<0.05$, $\chi 2$ test) of allele T in North of Italy than in all other regions. No deviations from Hardy Weinberg equilibrium were observed in any population/locus.

A few other studies already showed that European wild boar populations have experienced, to a various extent, introgression from domestic pigs, raising concerns about the management of wild boar genetic resources (e.g. Scandura et al. 2011b; Frantz et al. 2013; Goedbloed et al. 2013a, 2013b). Introgression may derive from accidental crossbreeding with free-ranging domestic pigs or from deliberate crosses in captive farming systems and then release of crossbred animals.

Considering that farmed wild boar stocks might have been subject of artificial selection against visible domestic phenotypes during their farm history (Goedbloed et al. 2013a), these animals might be selected for the E^{+} allele that is expected to maintain the wild boar coat colour

phenotype. Other *MC1R* alleles that are expected to have a detectable impact on coat colour might be eliminated from farmed wild boars. If this is true, we could speculate that introgression of *MC1R* domestic alleles we observed in North of Italy and SE Europe might be mainly accidental, derived by local domestic pig breeds that are usually raised in extensive or semi-extensive ways in free ranging systems. However, the quite high frequency of domestic *MC1R* alleles observed in both regions (about 25% and 16% of the animals carried domestic alleles in SE Europe and North of Italy, respectively) is puzzling. It might be derived by the combination of different events (that could be also recent) and the subsequent accumulation of potentially advantageous or neutral mutations in the wild or by recurrent introgression from restocking actions. It is also interesting to note that E^{D2} was the most frequent domestic allele in SE European populations whereas E^{P} was the most frequent allele of non-wild origin in North of Italy. E^{D2} is commonly supposed to be present in local black pigs (even if there is no information in the literature about the *MC1R* genotypes of all local Balkan pig breeds). Allele E^{P} is present in white and spotted pig breeds, that are mainly cosmopolitan and commercial breeds (e.g. Large White, Landrace, Pietrain; Fontanesi et al. 2010) and that could have been used to improve performances of the wild restocked animals. Domestic alleles might produce phenotypic effects if present in homozygous conditions but also if in heterozygous conditions, depending on the genetic background of the pigs. Of the three wild boars that did not carry any E^{+} allele we have a report on one of them that had white spots in the legs. For all other animals (homozygous or carriers of domestic alleles) we did not receive any description that suggested that the animals did not have the usual wild boar phenotypes. Unfortunately, we did not have pictures of the animals and we could not take a more detailed look at their coat colour that could be eventually slightly different from the common wild type. Anyway, from this survey in wild boar populations it seems quite clear that E^{D2} and E^{P} should not be dominant (or completely dominant) over the E^{+} allele, providing a reason why these alleles segregate in wild boars without any substantial modification of coat colour (according to what was reported to us). The same can be applied to the recessive allele *e* even if it was observed in just few wild boars from North of Italy. It

is also not clear what could be the effect of the domestic alleles on coat colour of wild boar piglets for which this issue might be more relevant in terms of fitness than in adult animals. Another interesting evidence derives by the heterogeneity present at the *Extension* locus in wild boars (three domestic alleles were detected) that might come from multiple introgression events from different domestic pigs having different *MC1R* genotypes. Only in the Sardinian population we did not identify introgression signals for this gene. Other twenty-one Sardinian wild boars were recently analysed at the *MC1R* gene by Canu et al. (2016) who reported that two animals had genotype E^p/E^p, one had genotype E^p/E^+ and all other Sardinian wild boars had genotype E^+/E^+ (that accounted for about 12% of *MC1R* domestic alleles in Sardinian wild boars) suggesting that a larger number of animals should be sampled in this region to have a more precise evaluation of domestic allele distribution.

Introgression demonstrated by genotyping *MC1R* polymorphisms was confirmed by the results obtained with the *NR6A1* g.299084751C>T SNP. For this marker, in all three wild boar populations allele T was always observed, including the Sardinian population that showed the highest frequency of the domestic allele (0.12). This is the same frequency of domestic alleles that was observed for the *MC1R* gene by Canu et al. (2016) who, on the other hand, reported that all their analysed Sardinian wild boars had the wild type genotype at the *NR6A1* gene. A previous study on Sardinian wild boars based on microsatellite markers reported a low level of introgression from domestic pigs (Scandura et al. 2011a), with signals that was estimated in about half the value that could be suggested by the results we obtained with the *NR6A1* polymorphism and confirmed by Canu et al. (2016) considering only the *MC1R* genotypes. This could be due to a bias in the sampling [our present work and Canu et al. (2016) analysed a lower number of wild boars than Scandura et al. (2011a)] or by the higher informativity of polymorphisms in the *MC1R* and *NR6A1* genes in terms of possibility to detect also past events.

The quite high frequency of domestic *NR6A1* alleles 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 increase 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). It would be interesting to evaluate if this marker, in additional with others, could be used to predict population expansive potential of wild boars considering also that litter size has been previously reported to vary in wild boars from different regions (Gethöffer et al. 2007). From a molecular ecology perspective, the putative advantage of the domestic *NR6A1* allele, if confirmed, could provide the opportunity to evaluate the effect of natural selection on artificially selected alleles that might have reached wild boar populations through interdemic gene flow (Fulgione et al. 2016) derived by the contact or deliberate crossbreeding with domestic pigs.

As all animals investigated in our study were shot in the wild by hunters and were described to be phenotypically wild boars (without any other information available on their origin; potentially derived also from recent crossbreeding events), we can consider the obtained results as a picture describing wild genetic resources in the sampled regions. Combining information derived from the *MC1R* and *NR6A1* genes (i.e. at least one signal per animal, that means carriers of at least one domestic allele; applicable for the two populations in which for both genes we detected domestic alleles) the percentage of wild boars with signal of introgression from domestic genes further increased at 34.44% and 15.58% in SE European and North of Italy populations, respectively. These values further extend the level of "genetic contamination" in wild boar populations compared to what was estimated in other studies that investigated randomly sampled wild boars (about 5 to 11%), reaching values close to what was observed in recognized captive breeding wild boar stocks

and far from what would be expected from the natural gene flow between domestic pigs and wild boars (Koutsogiannouli et al. 2010; Scandura et al. 2011a; Frantz et al. 2012, 2013; Goedbloed et al. 2013a, 2013b; Canu et al. 2014, 2016). These data might raise concerns on the genetic integrity of European wild boar populations due to excessive introgression of domestic alleles that might have been re-shaping the European wild boar genome, and in turn the biology of these wild animals. It will be interesting to enlarge the sampling in other European regions (and also extending it to have a larger coverage in Italy and Sardinia) to obtain a more precise evaluation of this gene flow process in continental and insular wild boar populations.

Conclusions

In this study, we reported allele frequencies at the *MC1R* and *NR6A1* genes in wild boars from three European regions, i.e SE Europe, North of Italy and Sardinia. These genes were investigated as they have been directly targeted by the domestication processes in pigs. The results showed a high level of introgressed domestic alleles in European wild boar populations, that might be derived by independent gene flow, as suggested by different domestic *MC1R* alleles being the most frequent in different regions. These introgressed domestic alleles may contribute to increase genetic variability in wild boar populations that could be maintained in the wild only if they are neutral or could increase fitness. Additional studies are needed to evaluate these issues in wild boars, including the use of high density SNP chips to cover other genomic regions.

Acknowledgements

We would like to thank collaborators and hunters for providing wild boar samples: Dr. Maurizio Musolesi (Polizia Provinciale of Bologna); Savigni Silvano; the self-managed group of

“Compagnia Lecca di San Basilio”; the President and the members of “Torre Falcone di Stintino”; Mr. Giuliano Pesce, President of “Campu Chervaggiu” group; Mladen Gagrčin, Zoran Rašić, Dragan Rogožarski, the public company ‘Vojvodinašume’ Novi Sad (Serbia); Ranko Savić, Darko Cvijović, Veterinary Ambulance Nikšić (Montenegro); Vladimir Maletić (FYRO Macedonia); Matija Stregar (Slovenia); Albert Marinculić (Croatia); Ivona Panić (Bosnia and Herzegovina). Erasmus Mundus JoinEU-SEE Penta Staff mobility scholarships were granted to MD and NV providing financial support for conducting this research. This study was supported by University of Bologna RFO funds and MiPAAF Innovagen project.

References

Alves CP, Pinheiro Í, Godinho R, Vicente J, Gortázar C, Scandura M. 2010. Genetic diversity of wild boar populations and domestic pig breeds (*Sus scrofa*) in South-western Europe. Biol J Linnean Soc. 101:797-822.

Borchers N, Reinsch N, Kalm E. 2004. The number of ribs and vertebrae in a Piétrain cross: variation, heritability and effects on performance traits. J Anim Breed Genet. 121:392-403.

Canu A, Costa S, Iacolina L, Piatti P, Apollonio M, Scandura M. 2014. Are captive wild boar more introgressed than free-ranging wild boar? Two case studies in Italy. Eur J Wildlife Res. 60:459-467.

Canu A, Vilaca ST, Iacolina L, Apollonio M, Bertorelle G, Scandura M. 2016. Lack of polymorphism at MC1R wild-type allele and evidence of domestic allele introgression across European wild boar populations. Mammal Biol. http://dx.doi.org/10.1016/j.mambio.2016.01.003

Duijvesteijn N, Veltmaat JM, Knol EF, Harlizius B. 2014. High-resolution association mapping of number of teats in pigs reveals regions controlling vertebral development. BMC Genomics. 15:542.

Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resours. 10:564-567.

Fajardo V, González I, Martín I, Rojas MA, Hernández PE, García AT, Martín R. 2008. Differentiation of European wild boar (*Sus scrofa scrofa*) and domestic swine (*Sus scrofa domestica*) meats by PCR analysis targeting mitochondrial D-loop and the nuclear melanocortin receptor 1 (*MC1R*) genes. Meat Sci. 78:314-322.

Fang MY, Larson G, Ribeiro HS, Li N, Andersson L. 2009. Contrasting mode of evolution at a coat color locus in wild and domestic pigs. PLoS Genet. 5:e1000341.

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.

Fontanesi L, Ribani A, Scotti E, Utzeri VJ, 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.

Frantz AC, Massei G, Burke T. 2012. Genetic evidence for past hybridisation between domestic pigs and English wild boars. Conserv Genet. 13:1355-1364.

Frantz AC, Zachos FE, Kirschning J, Cellina S, Bertouille S, Mamuris Z, Koutsogiannouli EA, 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 Linnean Soc. 110:104-115.

Frantz LA, Schraiber JG, Madsen O, Megens HJ, Cagan A, Bosse M, Paudel Y, Crooijmans RP, Larson G, Groenen MA. 2015. Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. Nat Genet. 47:1141-1148.

Freeman VA. 1939. Variations in the number of vertebrae of swine. J Hered. 30: 61-64.

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.

Gethöffer F, Sodeika, G, Pohlmeyer K. 2007. Reproductive parameters of wild boar (*Sus scrofa*) in three different parts of Germany. Eur J Wildlife Res. 53:287-297.

Giuffra E, Kijas JM, Amarger V, Carlborg O, Jeon JT, Andersson L. 2000. The origin of the domestic pig: independent domestication and subsequent introgression. Genetics. 154:1785-1791.

Goedbloed DJ, Megens HJ, Van Hooft P, Herrero-Medrano JM, Lutz W, Alexandri P, Crooijmans RP, Groenen M, Van Wieren SE, Ydenberg RC, Prins HH. 2013a. Genome-wide single nucleotide polymorphism analysis reveals recent genetic introgression from domestic pigs into Northwest European wild boar populations. Mol Ecol. 22:856-866.

Goedbloed DJ, van Hooft P, Megens HJ, Langenbeck K, Lutz W, Crooijmans RPMA, van Wieren SE, Ydenberg RC, Prins HHT. 2013b. Reintroductions and genetic introgression from domestic pigs have shaped the genetic population structure of Northwest European wild boar. BMC Genet. 14:43.

Gongora J, Peltoniemi OAT, Tammen I, Raadsma H, Moran C (2003) Analyses of possible domestic pig contribution in two populations of Finnish farmed wild boar. Acta Agric Scand (A). 53:161-165

Hubisz M, Falush D, Stephens M, Pritchard JK. 2009. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour. 9:1322-1332.

Kijas JHM, Wales R, Tornsten A, Chardon P, Mollet M, Andersson L. 1998. Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. Genetics. 150:1177-1185.

Kijas JM, 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. 158:779-785.

Koutsogiannouli EA, Moutou KA, Sarafidou T, Stamatis C, Mamuris Z. 2010. Detection of hybrids between wild boars (*Sus scrofa scrofa*) and domestic pigs (*Sus scrofa domestica*) in Greece using the PCR-RFLP method on Melanocortin-1 Receptor (MC1R) mutations. Mamm Biol. 75:69-73.

Mikawa S, Morozumi T, Shimanuki SI, 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, germcell nuclear factor (NR6A1). Genome Res. 17:586-593.

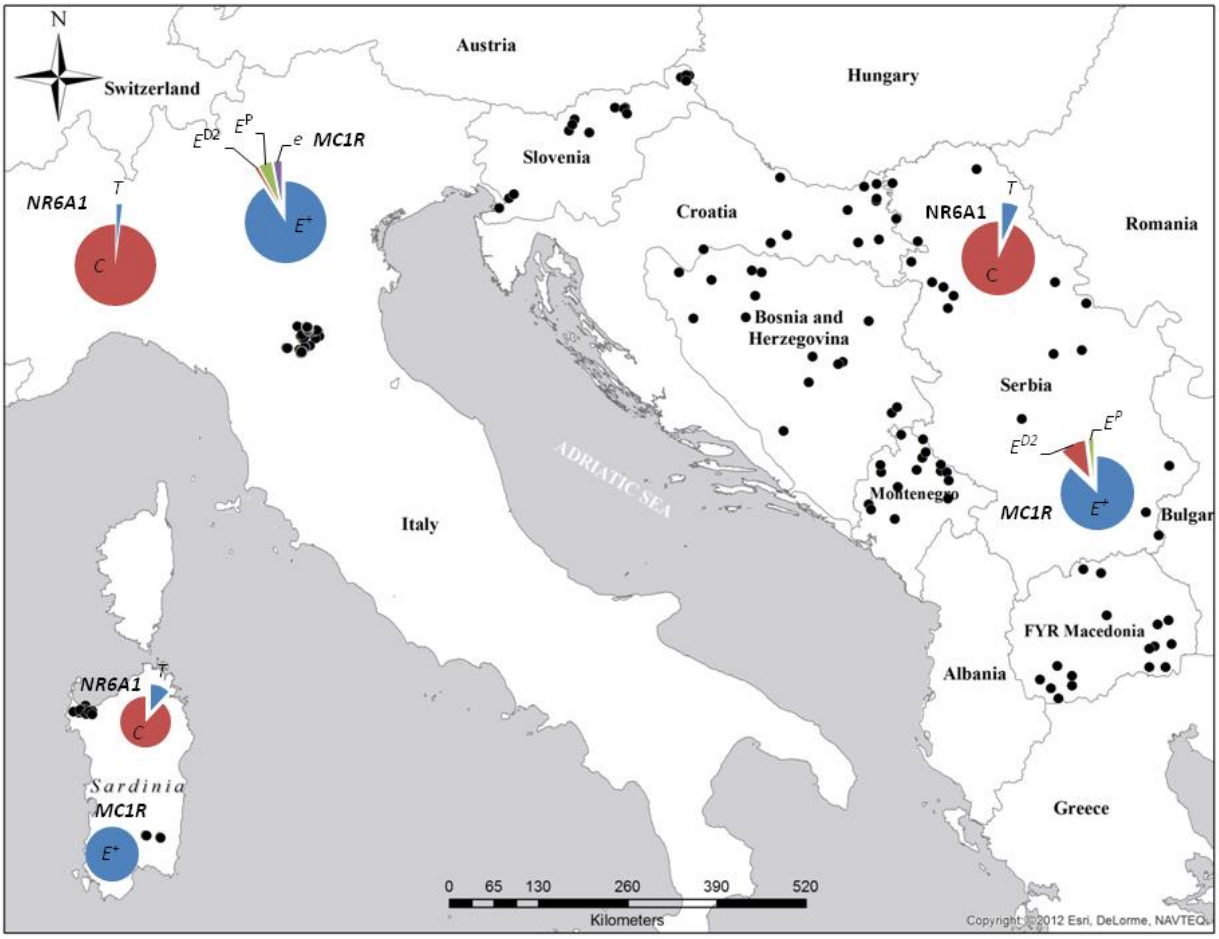
Rubin CJ, Megens HJ, Martinez Barrio A, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg Ö, Jern P, Jørgensen CB, Archibald AL, Fredholm M, Groenen MA, Andersson L. 2012. Strong signatures of selection in the domestic pig genome. Proc Natl Acad Sci USA. 109:19529-19536.

Scandura M, Iacolina L, Cossu A, Apollonio M. 2011a. Effects of human perturbation on the genetic make-up of an island population: the case of the Sardinian wild boar. Heredity. 106:1012-1020.

Scandura M, Iacolina L, Apollonio M. 2011b. Genetic diversity in the European wild boar *Sus scrofa*: phylogeography, population structure and wild x domestic hybridization. Mamm Rev. 41:125-137.

Verhoeven KJ, Macel M, Wolfe LM, Biere A. 2011. Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. Proc Royal Soc London B: Biol Sci. 278:2-8.

Figure 1. Sampling locations of wild boars (black dots) and *MC1R* and *NR6A1* allele frequencies (pie charts) in the three investigated regions: North of Italy, South-East Europe (Former Yugoslavia countries) and Sardinia.



3 FINAL CONCLUSIONS

Considering that coat colour has been one of the main target of both natural and artificial selection and in particular one of the first of domestication process, this Thesis has been focused on the detection of genetic variability of new variants in coat colour genes in different species, especially those associated with particular coat colour phenotypes and introgression of domesticated alleles in wild populations. Moreover, among these variants, the same coat colour phenotypes detected in animals are shown also in some pathological conditions in humans (e.g. albinism), suggesting that rabbits and donkeys could be considered as animal models for albinism because these variants have never been detected in humans.

Another aspect is related to the presence of different pigmentations in wild populations which are not commonly detected due to the strong natural selection pressure acting on coat colour. In fact coat colour is correlated to the fitness of several wild species mainly in terms of sexual selection, UV protection, mimicry and it is unwilling to change.

The lack of pigmentation, different from the *Brown locus*, is an extreme and very rare phenotype in the wild that lead to a white colour, because as reported above it is not an advantageous character; however the Asinara White donkey population show that, in particular conditions, it can be tolerated. Although these donkeys may have UV radiation damages in absence of melanins, they can normally survive and mate in a Mediterranean sunny envinroment thanks to the presence of uninhabited ruins that they use as shelters together with a low activity during the sunniest hours of the day and the absence of natural predators.

The two causative mutations detected in TYRP1 and TYR genes in rabbit and donkey respectively can be considered two relevant signatures of domestication in these species.

Another field explored in this Thesis has been the use of coat colour genes as introgression markers. As already mentioned before, wild species/populations don’t tend to show coat colour variation and appear generally uniform for this trait. This led several authors to use variants in coat colour genes, among which MC1R, as introgression markers. Although MC1R has been investigated by several authors because contains domestic haplotypes reated to different coat colours, introgression phenomena from wild to domestic populations have decreased its differentiation power between domestic and wild populations. This is the reason why in this study another domestic marker (NR6A1 gene) has been used to detect domestic gene flow towards wild boar populations in different European regions; NR6A1, in fact, is strongly associated with the number of vertebrae in

pigs and it depends on the domestication process that led to an increased body size and growth rate in pigs.

The results suggest that accurate monitoring and management of rural stock populations in order to avoid backcrossing events are needed both for the conservation of the ecologic natural equilibrium of wild environments and to make wild species less invasive for human activities.

A relatively recent application of introgression genetic markers is to use them to differentiate meat products originating from wild boars, considered an upper market niche, against pig meat products with a less economic value. This approach to detect fraudolent attitudes in foodstuffs is promising as demonstrated in the paper published in Meat Science (2014).

In this contest, future applications will involve the detection of more and new markers of domestication in different species in order to be able to increase the accurancy of these analyses and the authentication of more of foods types, as well as the identification of more food frauds.

4 APPENDIX

The Appendix contains the list of all posters presented in several Scientific Congresses attended during my PhD period, after them are listed Oral Communications and other published papers not mentioned in the Thesis.

Posters in Scientific Congresses

Utzeri V.J., Ribani A., Schiavo G., Bertolini F., Geraci C., Bovo S., Fontanesi L. *Food metagenomics against frauds: applications of next generation semiconductor based sequencing on meat and dairy products and honey* – FoodInnova 2017 – 31 January – 3 February 2017 – Cesena (Italy)

Ribani A., Utzeri V.J., Geraci C., Dall'Olio S., Nanni Costa L., Fontanesi L. Local pig breeds are "less domesticated" than commercial populations: evidences from variability in the MC1R and NR6A1 genes in Italian autochthonous breeds – 9th International Symposium on Mediterranean Pig – Portalegre (Portugal) – 3-5 November 2016

Ribani A., Bertolini F., Schiavo G., Scotti E., Utzeri V.J., Dall'Olio S., Trevisi P., Bosi P., Fontanesi L. Next generation semiconductor based sequencing of bitter taste receptor genes in different pig populations and association study using a selective DNA pool-seq approach – 67th EAAP Annual Meeting (European Federation of Animal Science) – Belfast (UK) – 29 August – 2 September 2016

Ribani A., Bertolini F., Schiavo G., Scotti E., Utzeri V.J., Dall'Olio S., Trevisi P., Bosi P., Fontanesi L. A next generation semiconductor based target re-sequencing DNA pool-seq approach for the identification of SNPs and association studies: application to bitter taste receptor genes in

different pig populations - 35th ISAG (International Society for Animal Genetics) Conference – Salt Lake City (USA) 23-27 July 2016

Schiavo G., Hoffmann O.I., Ribani A., Utzeri V.J., Ghionda M.C., Bovo S., Fontanesi L. A genomic landscape of mitochondrial DNA insertions in the nuclear pig genome - 35th ISAG (International Society for Animal Genetics) Conference – Salt Lake City (USA) 23-27 July 2016

Fontanesi L., Bovo S., Schiavo G., Mazzoni G., Ribani A., Utzeri V.J., Dall'Olio S., Bertolini F., Fanelli F., Mazzullo M., Galimberti G., Calò D.G., Trevisi P., Pagotto U., Bosi P. Deconstructing the pig genome-metabolome functional interactions – 35th ISAG (International Society for Animal Genetics) Conference – Salt Lake City (USA) 23-27 July 2016

Demars J., Iannuccelli N., Utzeri V.J., Fontanesi L., Allain D. The *MLPH* expression is decreased in rabbits owning a dilution of coat colour - Proceedings of the 11th World Rabbit Congress, Qingdao, China, 15-18 June 2016.

Ribani A., Utzeri V.J., Schiavo G., Bovo S., Geraci C., Fontanesi L. Food genomics: application of innovative DNA analysis technologies for authentication of food products. - TRADEIT Entrepreneurship Summer Academy – Postdam (Germany) 6-10 June 2016

Utzeri V.J., Ribani A., Dall'Olio S., Scotti E., Veličković N., Fontanesi L. Evidence of introgression of domesticated genes in several European wild boar populations XXI ASPA Congress (Animal Science and Production Association) – Milan (Italy) 9-12 June 2015

Utzeri V.J., Bertolini F., Ribani A., Dall'Olio S., Fontanesi L. Identification of the causative mutation of the albinism determining the white coat colour of the Asinara donkey breed - XXI ASPA Congress (Animal Science and Production Association) – Milan (Italy) 9-12 June 2015

Fontanesi L., Sparacino G., Utzeri V.J., Scotti E., Fornasini D., Dall'Olio S., Frabetti A. Identification of polymorphisms in the GHR gene and association with finishing weight in a commercial meat rabbit line - XXI ASPA Congress (Ani mal Science and Production Association) – Milan (Italy) 9-12 June 2015

Trevisi P., Ribani A., Colombo M., Utzeri V.J., Bosi P., Fontanesi L. A first nutrigenomic trial in pigs identifies a DNA polymorphism affecting kynurenine metabolites after tryptophan addition and *E. coli* challenge, 13th Digestive Physiology of Pigs – Kliczkow (Poland) 19-21 May 2015

Ribani A., Utzeri V.J., Scotti E., Bertolini F., Dall'Olio S., Fontanesi L. Sequence analysis of the tyrosinase gene (*TYR*, Albino *locus*) in wild and domesticated rabbits (*Oryctolagus cuniculus*) and in other wild Lagomorph species, XX ASPA Congress (Animal Science and Production Association) - Bologna 11 - 13 Giugno 2013.

Oral presentations in Scientific Congresses

A premature stop codon in the TYRP1 gene is associated with brown coat colour in the European rabbit - RGB-Net Seminars and Meetings – Zagreb (Croatia) 7-8 May 2014

Other publications

Schiavo, G., Ivett Hoffmann, O., Ribani, A., Utzeri, V.J., Ghionda, M.C., Bertolini, F., Geraci, C., Bovo, S., and Fontanesi, L. (2017). A genomic landscape of mitochondrial DNA insertions in the pig nuclear genome provides evolutionary signatures of interspecies admixture. Accepted by *DNA Research*.

Abstract

Nuclear DNA sequences of mitochondrial origin (*numts*) are derived by insertion of mitochondrial DNA (mtDNA), into the nuclear genome. In this study, we provide, for the first time, a genome picture of *numts* inserted in the pig nuclear genome. The *Sus scrofa* reference nuclear genome (Sscrofa10.2) was aligned with circularized and consensus mtDNA sequences using LAST software. A total of 430 *numt* sequences that may represent 246 different *numt* integration events (57 *numt* regions determined by at least two *numt* sequences and 189 singletons) were identified, covering about 0.0078% of the nuclear genome. *Numt* integration events were correlated (0.99) to the chromosome length. The longest *numt* sequence (about 11 kbp) was located on SSC2. Six *numts* were sequenced and PCR amplified in pigs of European commercial and local pig breeds, of the Chinese Meishan breed and in European wild boars. Three of them were polymorphic for the presence or absence of the insertion. Surprisingly, the estimated age of insertion of two of the three polymorphic *numts* was more ancient than that of the speciation time of the *Sus scrofa*, supporting that these polymorphic sites were originated from interspecies admixture that contributed to shape the pig genome.

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

Abstract

Taste perception in animals affects feed intake and may influence production traits. In particular, bitter is sensed by receptors encoded by the family of TAS2R genes. In this research, using a DNA pool-seq approach coupled with next generation semiconductor based target resequencing, we analysed nine porcine TAS2R genes (*TAS2R1, TAS2R3, TAS2R4, TAS2R7, TAS2R9, TAS2R10,*

TAS2R16, TAS2R38 and *TAS2R39*) to identify variability and, at the same time, estimate single nucleotide polymorphism (SNP) allele frequencies in several populations and testing differences in an association analysis. Equimolar DNA pools were prepared for five pig breeds (Italian Duroc, Italian Landrace, Pietrain, Meishan and Casertana) and wild boars (5–10 individuals each) and for two groups of Italian Large White pigs with extreme and divergent back fat thickness (50 + 50 pigs). About 1.8 million reads were obtained by sequencing amplicons generated from these pools. A total of 125 SNPs were identified, of which 37 were missense mutations. Three of them (p.Ile53Phe and p.Trp85Leu in *TAS2R4*; p.Leu37Ser in *TAS2R39*) could have important effects on the function of these bitter taste receptors, based on *in silico* predictions. Variability in wild boars seems lower than that in domestic breeds potentially as a result of selective pressure in the wild towards defensive bitter taste perception. Three SNPs in *TAS2R38* and *TAS2R39* were significantly associated with back fat thickness. These results may be important to understand the complexity of taste perception and their associated effects that could be useful to develop nutrigenetic approaches in pig breeding and nutrition.

Fontanesi L., Di Palma F., Flicek P., Smith A. T., Thulin C. G., Alves P. C. and the Lagomorph Genomics Consortium. LaGomiCs—Lagomorph Genomics Consortium: An International Collaborative Effort for Sequencing the Genomes of an Entire Mammalian Order. Journal of Heredity 2016, esw010.

Abstract

The order Lagomorpha comprises about 90 living species, divided in 2 families: the pikas (Family Ochotonidae), and the rabbits, hares, and jackrabbits (Family Leporidae). Lagomorphs are important economically and scientifically as major human food resources, valued game species, pests of agricultural significance, model laboratory animals, and key elements in food webs. A quarter of the lagomorph species are listed as threatened. They are native to all continents except Antarctica, and occur up to 5000 m above sea level, from the equator to the Arctic, spanning a wide range of environmental conditions. The order has notable taxonomic problems presenting significant difficulties for defining a species due to broad phenotypic variation, overlap of morphological characteristics, and relatively recent speciation events. At present, only the genomes of 2 species, the European rabbit (Oryctolagus cuniculus) and American pika (Ochotona princeps) have been sequenced and assembled. Starting from a paucity of genome information, the main scientific aim of the Lagomorph Genomics Consortium (LaGomiCs), born from a cooperative initiative of the European COST Action "A Collaborative European Network on Rabbit Genome

Biology - RGB-Net" and the World Lagomorph Society (WLS), is to provide an international framework for the sequencing of the genome of all extant and selected extinct lagomorphs. Sequencing the genomes of an entire order will provide a large amount of information to address biological problems not only related to lagomorphs but also to all mammals. We present current and planned sequencing programs and outline the final objective of LaGomiCs possible through broad international collaboration.

Fontanesi L., Sparacino G., Utzeri V.J., Scotti E., Fornasini D., Frabetti A.. Identification of polymorphisms in the rabbit growth hormone receptor (GHR) gene and association with finishing weight in a commercial meat rabbit line. Animal biotechnology 2016, 27(2): 77-83

Abstract

A shortcut to identify DNA markers associated with economic traits is to use a candidate gene approach that is still useful in livestock species in which molecular tools and resources are not advanced or not well developed. Mutations in the growth hormone receptor (GHR) gene associated with production traits have been already described in several livestock species. For this reason GHR could be an interesting candidate gene in the rabbit. In this study we re-sequenced all exons and non-coding regions of the rabbit GHR gene in a panel of 10 different rabbits and identified 10 single nucleotide polymorphisms (SNPs). One of them (g.63453192C>G or c.106C>G), located in exon 3 was a missense mutation (p.L36V) substituting an amino acid in a highly conserved position across all mammals. This mutation was genotyped in 297 performance tested rabbits of a meat male line and association analysis showed that the investigated SNP was associated with weight at 70 days ($P < 0.05$). The most frequent genotype (GG) was in animals with higher weight at this age, suggesting that the high directional selection pressure toward this trait since the constitution of the genotyped line might have contributed to shape allele frequencies at this polymorphic site.

Bertolini F., Scimone C., Geraci C., Schiavo G., Utzeri V. J., Chiofalo V., Fontanesi L. Next generation semiconductor based sequencing of the donkey (Equus asinus) genome provided comparative sequence data against the horse genome and a few millions of single nuleotide polymorphisms. PloS one 2015, 10(7), e0131925

Abstract

Few studies investigated the donkey (Equus asinus) at the whole genome level so far. Here, we sequenced the genome of two male donkeys using a next generation semiconductor based sequencing platform (the Ion Proton sequencer) and compared obtained sequence information with the available donkey draft genome (and its Illumina reads from which it was originated) and with the EquCab2.0 assembly of the horse genome. Moreover, the Ion Torrent Personal Genome Analyzer was used to sequence reduced representation libraries (RRL) obtained from a DNA pool including donkeys of different breeds (Grigio Siciliano, Ragusano and Martina Franca). The number of next generation sequencing reads aligned with the EquCab2.0 horse genome was larger than those aligned with the draft donkey genome. This was due to the larger N50 for contigs and scaffolds of the horse genome. Nucleotide divergence between E. caballus and E. asinus was estimated to be ~ 0.52-0.57%. Regions with low nucleotide divergence were identified in several autosomal chromosomes and in the whole chromosome X. These regions might be evolutionally important in equids. Comparing Y-chromosome regions we identified variants that could be useful to track donkey paternal lineages. Moreover, about 4.8 million of single nucleotide polymorphisms (SNPs) in the donkey genome were identified and annotated combining sequencing data from Ion Proton (whole genome sequencing) and Ion Torrent (RRL) runs with Illumina reads. A higher density of SNPs was present in regions homologous to horse chromosome 12, in which several studies reported a high frequency of copy number variants. The SNPs we identified constitute a first resource useful to describe variability at the population genomic level in E. asinus and to establish monitoring systems for the conservation of donkey genetic resources.