UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

Multidisciplinary approach to genetic variability, inbreeding and fertility in the endangered Sorraia horse breed

Doutoramento em Biologia
Biologia da Conservação

Helena Josefina Kjöllerström

Tese orientada por:
Profª Doutora Maria do Mar Jácome Félix Oom
Prof. Doutor Bhanu Pratap Chowdhary

Documento especialmente elaborado para a obtenção do grau de doutor

2016

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

Multidisciplinary approach to genetic variability, inbreeding and fertility in the endangered Sorraia horse breed

Doutoramento em Biologia
Biologia da Conservação

Helena Josefina Kjöllerström

Tese orientada por:
Profª Doutora Maria do Mar Jácome Félix Oom
Prof. Doutor Bhanu Pratap Chowdhary

Júri:
Presidente:
- ●Doutora Maria Manuela Gomes Coelho de Noronha Trancoso

Vogais:
- ● Doutora Terje Raudsepp
- ● Doutora Raquel Maria Garcia dos Santos Chaves
- ● Doutor José António dos Santos Pereira de Matos
- ● Doutor Luís Lavadinho Telo da Gama
- ● Doutora Maria Manuela Gomes Coelho de Noronha Trancoso
- ● Doutora Maria do Mar Jácome Félix Oom

Documento especialmente elaborado para a obtenção do grau de doutor

SFRH/BD/81502/2011, cE3c UID/BIA/00329/2013, PRODER (57882)(EU)/Action 2.2.3 (2011-2013), LINK Equine Research Endowment, American Quarter Horse Foundation and USDA (#2012-67015-19632)

2016

This work was supported by FCT/MEC (Fundação para a Ciência e Tecnologia) PhD Fellowship (SFRH/BD/81502/2011), cE3c FCT Unit funding (Ref. UID/BIA/00329/2013), PRODER (Contract number 57882)(EU)/Action 2.2.3 (2011-2013) [Conservation and improvement of genetic resources, Sub-action 2.2.3.2. Animal component (Sorraia breed)], LINK Equine Research Endowment, American Quarter Horse Foundation and USDA (grant #2012-67015-19632).

This dissertation should be cited as:

Kjöllerström, H.J. (2016) Multidisciplinary approach to genetic variability, inbreeding and fertility in the endangered Sorraia horse breed. PhD Thesis, University of Lisbon, Portugal

To my Family and Friends

The wind of heaven is that which blows between a horse's ears.

~Arabian Proverb

Nota prévia

Na elaboração da presente dissertação e, nos termos do nº 2, alínea a) do Artigo 25 do regulamento de Estudos de Pós-Graduação da Universidade de Lisboa, publicado em Diário da República nº 57, 2ª Série de 23 de Março de 2015, esclarece-se que apenas foram considerados integralmente artigos científicos originais já publicados (4), submetidos (1) ou em preparação (1), em revistas indexadas de circulação internacional, os quais integram os capítulos da presente tese. A candidata declara que participou integralmente no planeamento, elaboração, análise e discussão dos resultados de todos os trabalhos com primeira autoria. No caso dos trabalhos de colaboração, nomeadamente o intitulado *Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis* (Capítulo 2, Paper 1), a autora participou na análise, interpretação e discussão de resultados, assim como na preparação do manuscrito.

A presente dissertação, por ser uma compilação de artigos científicos internacionais, foi redigida em língua Inglesa. No final de cada capítulo é apresentada uma lista de referências em vez de uma única lista no final da tese, razão pela qual poderá ocorrer duplicação das mesmas nos diferentes capítulos. Cada capítulo tem, também, associada a respectiva informação de suporte, quando necessária. Alguns capítulos têm formato diferente devido aos diferentes requisitos das revistas científicas nos quais os mesmos foram publicados, submetidos ou virão a ser submetidos.

Preliminary note

In the preparation of this dissertation, in accordance with paragraph 2, a) of Article 25 of the Post-Graduate Studies Regulation of the University of Lisbon (*Diário da República nº 57, 2ª Série de 23 de Março de 2015*), only original scientific articles already published (4), submitted (1) or in preparation (1), in international indexed journals were considered, integrating the chapters of this thesis. The candidate fully participated in the planning, preparation, result analysis, discussion and writing of the manuscripts presented as first author. In case of collaborative work, such as *Genetic diversity and demographic structure of the endangered Sorraia horse breed through pedigree analysis* (Chapter 2, Paper 1), the author participated in the analysis, interpretation and discussion of results, as well in the manuscript preparation.

In addition, because this dissertation is composed of a series of international publications, it has been written in English. A reference list at the end of each chapter, rather than in the end of the dissertation, is presented. This may lead to references duplication between chapters. Each chapter has also their own supplementary information, when necessary. The different format between chapters is a direct result of the different requirements from the scientific journals in which they were published, or will be submitted.

TABLE OF CONTENTS

ACKNOWLEDGMENTS III

RESUMO / PALAVRAS-CHAVE VI

ABSTRACT / KEYWORDS XI

CHAPTER 1: INTRODUCTION 1

1.1 - The Sorraia horse breed 2

1.2 - Genetic variability 9

1.2.1 - Pedigree analysis 9

1.2.2 - Molecular analysis 11

1.3- Reproductive fitness and inbreeding 16

1.4- Specific aims 20

1.5- Thesis Structure 21

REFERENCES 24

CHAPTER 2: BREED CHARACTERIZATION 33

PAPER 1 - Pinheiro M., Kjöllerström H.J. & Oom M.M. (2013) **Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis.** Livestock Science 152, 1-10 34

PAPER 2 - Kjöllerström H.J. & Oom M.M. **Morphological characterization and inbreeding effects in the endangered Sorraia Horse breed** (*submitted to the Journal of Animal Science*) 45

PAPER 3 - Kjöllerström, H.J., Raudsepp, T., Khanshour, A., Gosh, S., Oom, M.M., Chowdhary, B.P. **Genome-wide variability in the Sorraia horse** (*In preparation for PLoS ONE*) 70

CHAPTER 3: REPRODUCTIVE FITNESS 115

PAPER 4 - Kjöllerström H.J., Gama L.T. & Oom M.M. (2015) **Impact of inbreeding on fitness-related traits in the highly threatened Sorraia horse breed.** Livestock Science 180, 84-9 116

PAPER 5 - Kjöllerström H.J., Collares-Pereira M.J. & Oom M.M. (2011) **First evidence of sex chromosome mosaicism in the endangered Sorraia Horse breed.** Livestock Science 136, 273-6 123

PAPER 6 - Kjöllerström H.J., Oom M.M., Chowdhary, B.P & Raudsepp, T. (2016) **Fertility assessment in Sorraia stallions by sperm-FISH and *FKBP6* genotyping.** Reproduction in Domestic Animals 51, 351-9 128

CHAPTER 4: DISCUSSION 143

4.1 - Breed Characterization 144

4.2 - Reproductive fitness 152

CHAPTER 5: FINAL REMARKS 157

REFERENCES 160

GLOSSARY 166

Acknowledgements

First and foremost, I wish to thank my supervisor, Prof. Maria do Mar Oom: without your help, encouragement, support and faith in me, none of this would have been possible. Thank you for pushing me through all the hardship and chaos that happened a little bit throughout these years.

A special thank you to my co-supervisor, Dr. Bhanu P. Chowdhary: you gave my PhD a new route and I learned so much because of that. Thank you for the opportunity you gave me to join your lab at Texas A&M. My work and knowledge greatly improved because of that.

To Dr. Terje Raudsepp, a big thank you for all you taught me while I was at your lab. Thank you for all the discussions we had and how much you helped me when I was lost in what sense to make of my results. Thank you for making feel part of your family and for always welcoming me in your house and lab.

To Prof. Maria João Collares-Pereira, for all the help and wisdom shared regarding cytogenetics.

To my dad, a special thank you - your “paitrocínio” made my PhD happen! It made field work possible, travelling to Texas and also supported me for half of my PhD. Thank you for always supporting me and “forcing” me to do this because “the family business was not my future”, as you like to put it, except when it comes to meetings in Spain, ahahah.

To my mom for putting up with my meltdowns when push came to shove, and for always having wise words of encouragement to share with me.

To my sisters, publish or perish forever!! I love you, thank you for your support and faith in me, even in times when I didn’t believe in myself.

To my nephews and niece, you always find a way to brighten up my days. Thank you for all the laughter, smiles and unconditional love.

To Raquel because putting up with me for 6 years, through all the hard times I’ve been through, can’t have been easy. Thank you for never giving up on me and always being there for me, for whatever I needed at whatever time. I couldn’t have done it without you (literally!). Thank you for all the help: you’ve taught me to be a better person.

To Carla, for the companionship and scientific discussions. Thank you for all you’ve taught me and for being there for me when I needed.

To Ana Catarina, because Texas (and clearly the Universe) decided we needed to meet and help each other. You know you have a very special place in my life and heart; you’ve helped me through this journey. I’ve grown to be a different and better person because of you and your teachings. Thank you!

To my Portuguese friends, because I know I’m a pain and have made your lives harder at points because of my meltdowns and doubts.

To my Texas family, I cannot imagine my life without you. You have made my life so much better. At the same time I’ve discovered so many new things about myself. I’ve become a better person from meeting you.

To Kathy, thank you for making me a part of your family, for all the experiences shared and for all you’ve done for me throughout the years. Thank you!

To Doug, Lori and Clementine, thank you for everything and for making me part of your family. I’m very thankful to have met y’all.

To Dr. Gus Cothran, for the help with my work.

To Daisy, Ingrid, Vanessa, Sharmila, Samantha, Pranab, Rytis, Felipe, Fahad and Jan, you've made Lab work so much better and we've learned a lot from each other. Thank you for all the pep talks and shoulders to cry on when the "Lab midget" was making sure nothing worked!

To the "K" sisters, thank you for accepting me as one of you and always being there for me. You are truly special and make Texas feel like home every time. You are family to me and I can't imagine life without Sunday family dinners anymore.

To Odair, Davi, Ahmed, Jeann, Mark, Steven and all the guys in the gang, thank you for all the awesome times and companionship. You've made my time in College Station that more awesome as you also did the nightlife, salsa and country dancing.

Resumo

O Cavalo do Sorraia é um cavalo primitivo, considerado um recurso genético equino universal. Este tipo equino é, provavelmente, o representante actual do tipo equino ancestral de várias raças de cavalos de sela Ibéricas e do Novo Mundo. A raça foi recuperada em 1937 pela família Andrade, a partir de 12 animais fundadores, sendo gerida como uma população fechada desde então. Existem actualmente 10 criadores de cavalos Sorraia em Portugal e oito na Alemanha. Com apenas ~ 150 éguas reprodutoras e uma população mundial de ~ 300 animais, o Sorraia encontra-se em *critical risk status*, enfatizamdo a necessidade de estabelecer um plano de conservação que vise atingir uma população auto-sustentável a longo prazo. O Sorraia é conhecido por ter variabilidade genética reduzida, níveis de consanguinidade extremamente elevados, esperma de má qualidade e, também, por sofrer de subfertilidade. Os valores de consanguinidade são muito superiores aos descritos em qualquer outra raça de cavalos, com uma média de 0.38 na população actual.

Para estudar a variabilidade genética, estrutura demográfica e estado actual da população foram usados todos os animais inscritos no Livro Genealógico (Studbook) da raça Sorraia (população total, constituída por 653 animais, nascidos entre 1937-2006), ou uma amostra mais reduzida (constituída pelos 206 animais vivos, à data da realização dos trabalhos). O intervalo de geração médio foi de 7.94 anos. A diversidade genética encontrada foi reduzida, estando de acordo com trabalhos anteriormente publicados nesta raça. A contribuição dos fundadores provou ser muito díspar sendo que alguns já não se encontram sequer representados, demonstrando a perda de diversidade genética elevada que ocorreu ao longo das gerações. Apenas 7.46 fundadores (*effective founders*) e 4 antepassados (*effective ancestors*) estão actualmente representados, sendo que apenas 15 animais explicam a variabilidade genética global actual. Considerando todos os animais, o coeficiente de consanguinidade (0.27) e de parentesco médio (0.46) são extremamente elevados. A taxa de consanguinidade por geração (5.2%) foi muito superior à definida como máxima pela FAO (1%). Aproximadamente 97% dos animais da população actual apresentaram valores de consanguinidade superiores a 0.25. Devido às estratégias de conservação implementadas pela Associação de Criadores do Cavalo Sorraia, alguns parâmetros têm apresentado melhorias recentemente: os nascimentos aumentaram e a consanguinidade e parentesco médios têm vindo a estabilizar e, em alguns anos, diminuído.

Descrevemos morfologicamente dois grupos de garanhões separados por um período de 20 anos (*VELHOS* e *NOVOS*) através de 26 medidas corporais. Procurámos consequências da consanguinidade e ocorrência de depressão consanguínea na conformação, avaliando ao mesmo tempo o efeito da pelagem e idade nas medidas morfométricas analisadas. Garanhões

pertencentes ao grupo *VELHOS* tiveram em média valores morfométricos mais elevados que os *NOVOS*. A maioria das medidas corporais apresentou coeficientes de variabilidade baixos, corroborando a homogeneidade fenotípica deste grupo de animais e da raça. Apesar da conformação corporal global ter-se mantido durante o período analisado, verificamos que os garanhões Sorraia se tornam menores, excepto no que diz respeito ao tamanho da cabeça. Animais *NOVOS* apresentaram valores mais elevados de consanguinidade e idade. A consanguinidade teve impacto significativo e negativo sobre o diâmetro bicostal e perímetro torácico no grupo *NOVOS*. Teve também um efeito positivo significativo na altura do curvilhão no grupo *NOVOS* e na altura ao codilho no grupo *VELHOS*. A idade correlacionou-se na maioria com medidas que não possuem suporte ósseo na sua definição. Os garanhões rato escuro são mais baixos do que as outras pelagens, embora a maioria dessas diferenças não seja estatisticamente significativa. Seria importante usar esta informação como ferramenta complementar no actual programa de gestão da raça para garantir que no futuro os animais continuarão a seguir o padrão fenotípico da raça descrito no Studbook.

A caracterização molecular de uma população é importante para a sua conservação pois permite determinar a variabilidade genética existente e estabelecer metas de conservação. A estimativa de variabilidade genética é geralmente baseada na análise de marcadores polimórficos onde os mais utilizados são os microssatélites (ou *short tandem repeats* - STRs), *single nucleotide polymorphisms* (SNPs) e mais recentemente *copy number variations* (CNVs). Estes marcadores foram usados para determinar a variação genética existente no cavalo do Sorraia visando também contribuir para o melhoramento da gestão genética, bem como identificar marcadores eficientes para testes de paternidade. Como esperado, *loci* ligados ao cromossoma Y não apresentaram qualquer polimorfismo. Os resultados dos microssatélites autossómicos representam uma melhoria em relação aos previamente publicados na raça, produto do método de selecção dos marcadores, do maior número de STRs analisados e animais genotipados. O número médio de alelos (3.7), H_O (0.57), H_E (0.59) e heterozigotia média individual (0.57) foram baixos. A maioria dos parâmetros analisados apresentaram valores melhores na população Alemã que na Portuguesa, demonstrando que usar um garanhão por ano por égua é mais eficaz do que apenas um garanhão por ano por eguada. Alguns parâmetros como heterozigotia individual e d^2 médio melhoraram relativamente aos previamente publicados, evidenciando os resultados positivos do plano de gestão implementado. Tal como nos STRs, a consanguinidade e heterozigotia individual em SNPs foram melhores na Alemanha. Foram detectadas 47 *runs of homozygosity* (ROHs) em 11 cromossomas, com uma proporção média de 98.24% *sites* homozigóticos dentro das

regiões identificadas. Apesar do *software* Structure ter separado as nossas amostras (de 13 criadores diferentes) em três grupos, a análise factorial de correspondência (STRs) e análise de componentes principais (SNPs) separou-as em dois: Portugal e Alemanha. A troca de animais entre os dois países deve ser promovida para aumentar a variabilidade genética e reduzir a consanguinidade. Os CNVs provaram ser inadequados para avaliar a diversidade genética nesta raça altamente consanguínea. Ao descrever a variabilidade genética actualmente existente na raça, os dados aqui descritos podem ajudar a aumentar a heterozigotia permitindo escolher anualmente a melhor combinação de animais reprodutores.

Foi usada informação sobre todos os Sorraias de 1937-2010 para estudar os efeitos da consanguinidade na viabilidade ao nascimento e aos 6 meses de idade. Foi usada uma subamostra para analisar o impacto da consanguinidade na fertilidade de garanhões e éguas, intervalos entre partos e idade ao primeiro parto. O efeito da consanguinidade foi significativo apenas na fertilidade das éguas (P = 0.003) com 0.8% de redução na fertilidade com um aumento de 1% na consanguinidade. A idade da égua teve um efeito quadrático, reduzindo a mortalidade dos poldros (tanto ao nascimento como aos seis meses de idade) e aumentando o intervalo entre partos com o aumento da idade. A idade ao primeiro parto foi significativamente influenciada pela coudelaria.

Garantir taxas de fertilidade altas é de suma importância para a sobrevivência a longo prazo desta raça. Anomalias cromossómicas, particularmente de cromossomas sexuais, têm sido associadas a fertilidade reduzida e infertilidade em cavalos. Antes do início deste projecto, não havia estudos de avaliação de anomalias cromossómicas em animais com fertilidade reduzida nesta raça. O primeiro caso é por nós descrito neste trabalho, uma égua com fertilidade reduzida e comportamento de garanhão, com cariótipo mosaico 63, X0 / 64, XX (com frequências de 10.45% e 89.55%, respectivamente). A avaliação citogenética de animais com fertilidade baixa e/ou fenótipos sexuais ambíguos é necessária de modo a determinar a prevalência de anomalias cromossómicas nesta raça. A detecção precoce de animais sub-férteis é de suma importância para a gestão e conservação desta raça pois permite poupar aos criadores tempo, esforço e dinheiro.

Os garanhões desta raça são conhecidos por terem esperma de má qualidade. Neste estudo foram avaliados 11 garanhões fenotipicamente normais mas sub-férteis por cariotipagem, hibridação *in-situ* fluorescente em sémen de cromossomas sexuais e genotipagem do gene *FKBP6* - um *locus* de susceptibilidade para a reacção acrossómica reduzida (IAR). Todos os garanhões tiveram concentração espermática normal apesar da motilidade progressiva e formas morfologicamente normais serem baixas. Todos os animais apresentaram cariótipos

normais 64, XY. A maioria do esperma teve conteúdo cromossómico sexual haplóide normal, enquanto 11% exibiu vários tipos de aneuploidias. Não foi encontrada correlação entre a percentagem de anomalias cromossómicas sexuais no esperma e consanguinidade, formas morfologicamente normais ou idade. A sequenciação dos SNPs g.11040315G>A e g.11040379C>A do exão 4 do gene *FKBP6* mostrou que nenhum Sorraia apresenta o genótipo de susceptibilidade IAR (A/A-A/A). Todos os animais têm o mesmo genótipo G/G-A/A, corroborando a baixa variabilidade genética da raça. As nossas descobertas descartaram as anomalias cromossómicas e predisposição genética para IAR como factores que contribuem para a sub-fertilidade dos garanhões Sorraia analisados. No entanto, a sua baixa fertilidade pode ser parcialmente atribuída à maior percentagem de aneuploidias de cromossomas sexuais presentes no esperma destes animais.

Apesar de a consanguinidade não ter efeitos nefastos na maioria dos parâmetros analisados até à data, os níveis de consanguinidade devem continuar a ser monitorizados e controlados e a realização de estudos sobre depressão consanguínea permanecem pertinentes e incontornáveis para a conservação e sustentabilidade desta raça. É também importante que as medidas de gestão adequadas continuem a ser implementadas e melhoradas para preservar a diversidade genética existente.

Esperamos que os resultados obtidos neste trabalho ajudem no melhoramento genético e preservação da variabilidade genética desta raça extremamente rara e primitiva. Esperamos que este contributo ajude a prevenir a extinção deste recurso genético animal tão importante e icónico.

Palavras-chave: Cavalo Sorraia; Conservação; Variabilidade genética; Depressão consanguínea; Fertilidade; Citogenética

Abstract

The Sorraia horse is a primitive breed and regarded as a universal equine genetic resource. Recovered in 1937 in Font'Alva from 12 founder animals it has been managed as a closed population since. There are presently 10 Portuguese and eight German breeders. With only ~150 breeding mares and ~350 animals worldwide, the Sorraia horse population is in critical risk status. The Sorraia is known for low genetic variability, bad sperm quality, and high inbreeding (mean=0.38) that far exceeds those reported in other horse breeds.

The information available in the Sorraia Studbook was used to study genetic variability, demographic structure and status of the extant population. Only 7.46 effective founders and 4 effective ancestors characterize the living population, with only 15 animals explaining its overall genetic variability. Inbreeding and average relatedness were extremely high. Inbreeding by generation rate (5.2%) far exceeded that defined by FAO as maximum (1%). In the living population, approximately 97% of the animals had inbreeding values above 0.25. Recently, conservation strategies from the Sorraia Breeders Association have improved some of these parameters: births have increased and average inbreeding and relatedness have stabilized and, in some years, even decreased.

We morphologically described two 20-years-apart Sorraia stallions groups (OLD & NEW) based on 26 body measurements. Most body measurements showed low variability coefficients, attesting to the breed's (and this group's) phenotypic homogeneity. Although overall body conformation ratio has not changed, we found that Sorraia stallions have become smaller, except for head size. Inbreeding had a significant negative impact on thoracic width and chest circumference in NEW. Yellow dun horses were taller and longer than mouse dun ones. Dark mouse dun horses appeared shorter than the other coat colours. This data should be used as a complementary tool in the current management breeding program.

We used STRs, SNPs and CNVs to evaluate the genetic variation in the Sorraia horse and improve genetic management, and looked for efficient markers for parentage testing. Y-linked *loci* had no polymorphism. Average number of alleles, H_O, H_E and mean individual heterozygosity were low. Most analysed parameters were better in the German rather than Portuguese population. There were 47 runs of homozygosity regions. Although Structure analysis separated our samples from 13 different breeders in three clusters, FCA (with STRs) and PCA (with SNPs) analysis separated them in two groups: Portugal and Germany. Interchange between countries should be promoted to increase genetic variability and reduce inbreeding. Copy number variations proved to be unsuitable to assess genetic diversity in this highly inbred breed. The genome-wide data described herein can help increase heterozygosity by choosing yearly the right combination of sires and mares for reproduction.

Data on all Sorraia horses from 1937 until the end of 2010 were used to study inbreeding effects on offspring's viability at birth and at 6 months of age. A sub-sample was used to analyse inbreeding impact on stallion and mare fertility rates, foaling intervals and age at first parturition. Inbreeding effect on analysed traits was only significant for mare fertility (P=0.003), with a 0.8% decrease fertility by 1% increase in inbreeding. Age of the mare had a quadratic effect, reducing foal mortality (both at birth and at six months of age) and increasing foaling interval with increasing age. Age at first parturition was significantly influenced by stud farm.

Chromosomal abnormalities, particularly of sex chromosomes, have been associated with reduced fertility and infertility in horses. We describe the first case of chromosomal abnormalities in this breed: a sub fertile mare with reduced fertility and stallion-like behaviour with mosaic 63,X0/64,XX karyotype (10.45% and 89.55% frequency, respectively). Early detection of sub fertile animals is of paramount importance for the management and conservation of this breed, saving breeder's time, efforts and money.

Eleven phenotypically normal but sub-fertile stallions were studied by karyotyping, sex chromosome sperm-FISH and *FKBP6* analysis – a susceptibility *locus* for impaired acrosome reaction (IAR). All karyotypes were normal 64,XY. Most sperm had normal haploid sex chromosome content, while 11% carried various sex chromosome aneuploidies. There was no correlation between the percentage of sperm sex chromosome abnormalities and inbreeding, sperm morphology or age. All animals had G/G-A/A genotype, corroborating this breed's low genetic variability. Our findings ruled out chromosomal abnormalities and genetic predisposition for IAR as contributing factors for subfertility in the analysed Sorraia stallions. However, their low fertility could be partially attributed to higher rate of sperm sex chromosome aneuploidies.

Even though inbreeding hasn't had alarming effects on the majority of traits analysed, inbreeding should continue to be controlled and inbreeding depression studies remain vital for the long-term conservation and sustainability of this breed. These results will help improve genetic health and preserve the genetic variability of this extremely rare and primitive horse breed. With this work we hope to help prevent the permanent loss of this iconic and important animal genetic resource.

Keywords: Sorraia horse; Conservation; Genetic variability; Inbreeding Depression; Fertility; Cytogenetics

CHAPTER 1 - INTRODUCTION

1.1 The Sorraia horse breed

The bond between horses and humans goes back in time longer than there have been records. Humans have always been amazed by these remarkable creatures and their success greatly benefited from this bond.

The *Equus* lineage from which all modern horses, zebras and donkeys originated is said to have appeared 4.0-4.5 million years before present (Orlando *et al.* 2013). Domestication of horses in Europe is thought to have taken place during the Holocene in the Eastern steppes and the Iberian Peninsula (Warmuth *et al.* 2011). The Sorraia-type horse was surely one of the earliest domesticated horses, since it has been present in the Iberian Peninsula since early Pleistocene, and has been known to be used for travel, labour and companionship (Oom *et al.* 2004; Oom 2006).

The Sorraia horse (Figure 1), as well as all other domestic horse breeds, belongs to the species *Equus caballus* Linnaeus, 1758 (e.g. Bennett and Hoffmann (1999)). The breed finds its origins in a primitive equine type and has been described as one of the four ancestral horse types of the current domestic breeds, with a continuous presence in the Iberian Peninsula since early Pleistocene (Gonzaga 2004). It can be considered as a representative of the ancestor of Iberian saddle horses (Andrade 1945; Oom *et al.* 2004) - the Lusitano and Andalusian - and has also been involved in the foundation of several New World horse breeds (Andrade 1945; Bouman 1989; Luis *et al.* 2006).

While hunting in the Sorraia river valley Dr. Ruy d'Andrade, the breed's founder, found a herd of grazing horses and he realized that they resembled the ones portrayed in cave paintings from Southern Iberian Peninsula (e.g. Altamira) and France (e.g. Lascaux) dating back to the Palaeolithic (Andrade 1926, 1937, 1945, 1954) (Figure 1). A few years later, in 1937, Dr. Ruy d'Andrade bought a few horses with this primitive phenotype and started the Sorraia Horse breed, at Agolada Farm (Ribatejo), named after the river basin in which he had seen this equine type for the first time in Portugal. The breed was established from 7 mares (one pregnant with a colt) and 3 stallions, plus an additional stallion brought from Argentina in 1948, and has been managed as a closed population since its foundation (Oom *et al.* 2004).

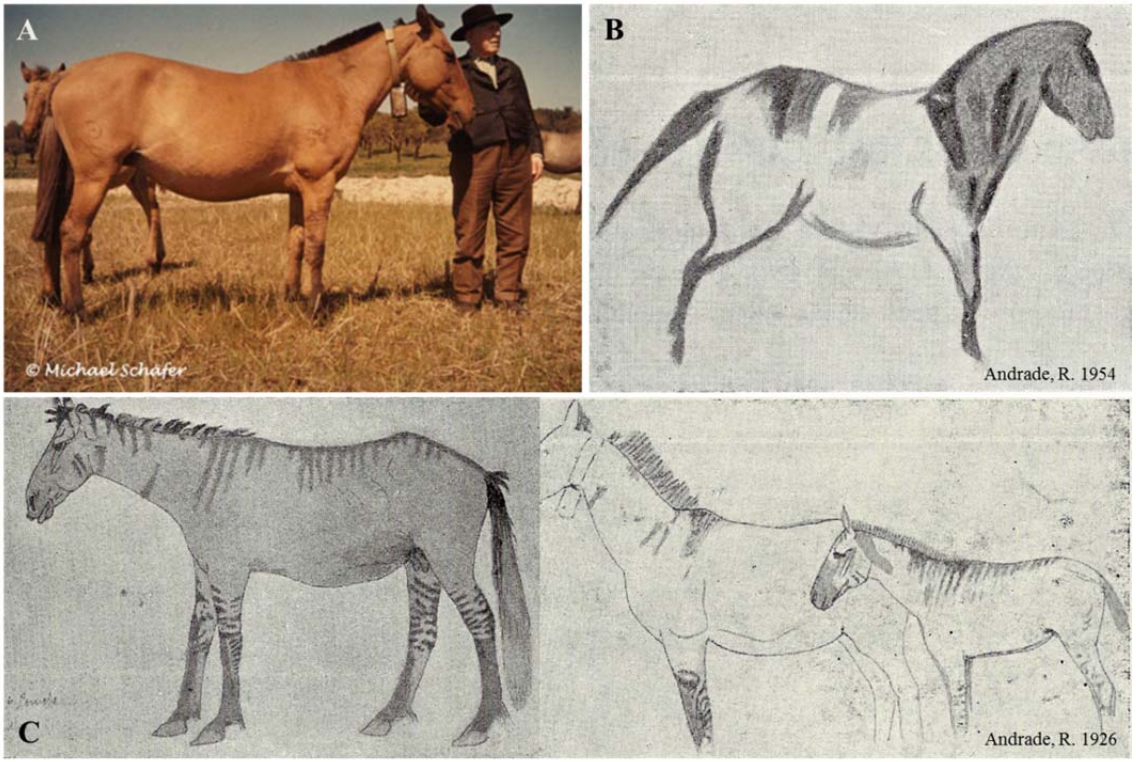


Figure 1 - (A) Dr. Ruy d' Andrade with a Sorraia mare in 1960; (B) cave paintings in Spain (Andrade 1954); (C) Drawings of Sorraia mare and Sorraia mare with foal (Andrade 1926).

As a primitive breed, these horses are built to endure and survive in harsh conditions (Oom *et al.* 2004). They also possess primitive traits such as yellow or mouse dun coat colour (Figure 2A), black dorsal and axial stripe (the latter less frequent) (Figure 2B, C), leg barring (zebra markings on front and hind legs) (Figure 2D, E), face masks and dark ear tips (Figure 2F), and mane and tail frosting (Figure 2B, H) (Andrade 1945; Oom *et al.* 2004).

Due to Portugal's political circumstances at the time, the Andrade herd was turned over to the National Stud Farm in 1975, in Alter do Chão (Alentejo), under the supervision of the Portuguese Government. The herd was later returned to the Andrade family, at Font'Alva Farm (Alentejo), which offered two colts and two fillies to the National Stud Farm starting a new subpopulation there, isolated until 2001 (Oom *et al.* 2004). In 1976, three mares and three stallions were sold to Germany, with no further addition until recently (Oom *et al.* 2004; Luis *et al.* 2007a), starting an important Sorraia breeding farm abroad. The third Portuguese Sorraia nucleus was established in 1985 with four mares sold to Quinta da Boavista (Ribatejo). After the death of Engº Fernando Sommer d'Andrade (son of Dr. Ruy d'Andrade) the Font'Alva nucleus was

subdivided between heirs, with the biggest herd remaining in the founder stud farm (Font'Alva) (Oom *et al.* 2004; Oom 2006) (Figure 3).

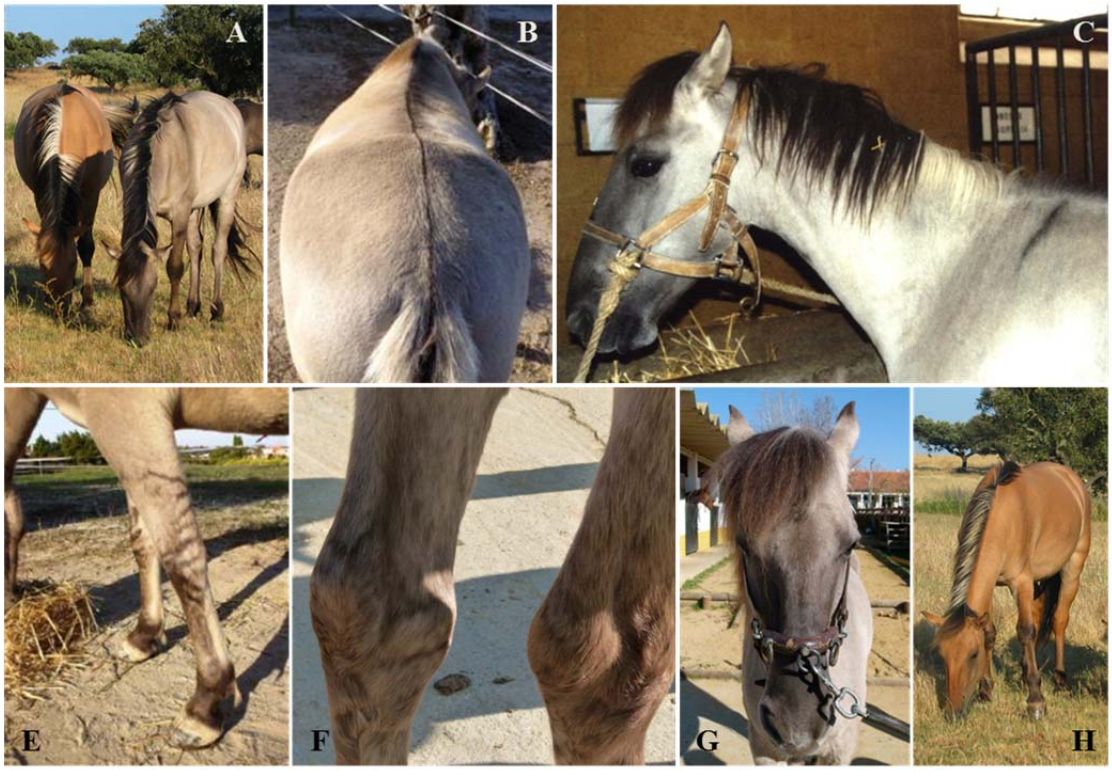


Figure 2 - Physical primitive characteristics of the Sorraia horse: yellow dun or mouse dun coat colour (A), black dorsal (B) and axial stripe (C), leg barring on front (E) and hind legs(F), face masks and dark ear tips (G), and mane and tail frosting (H).

There are currently ten breeders in Portugal and eight in Germany (Figure 3). There are also a few Sorraia horses in Spain, France, Belgium, Denmark, Brazil, Mexico, Canada and USA, though only a few have produced Sorraia offspring so far. In fact, for example, in 2000-2001 two stallions were sold to the USA to two separate owners. In Nature's Baroques Farm (Eagle Creek - Oregon) one of the stallions was bred to local mares and has had offspring registered in several studbooks since 2003 (Oom 2006). In 2006 a pregnant mare was bought by a Brazilian breeder (Coudelaria Verde Prado, Monte Alegre do Sul - São Paulo) and has since given birth to one foal. Artificial insemination with shipped semen from Portugal has been attempted but so far unsuccessfully.

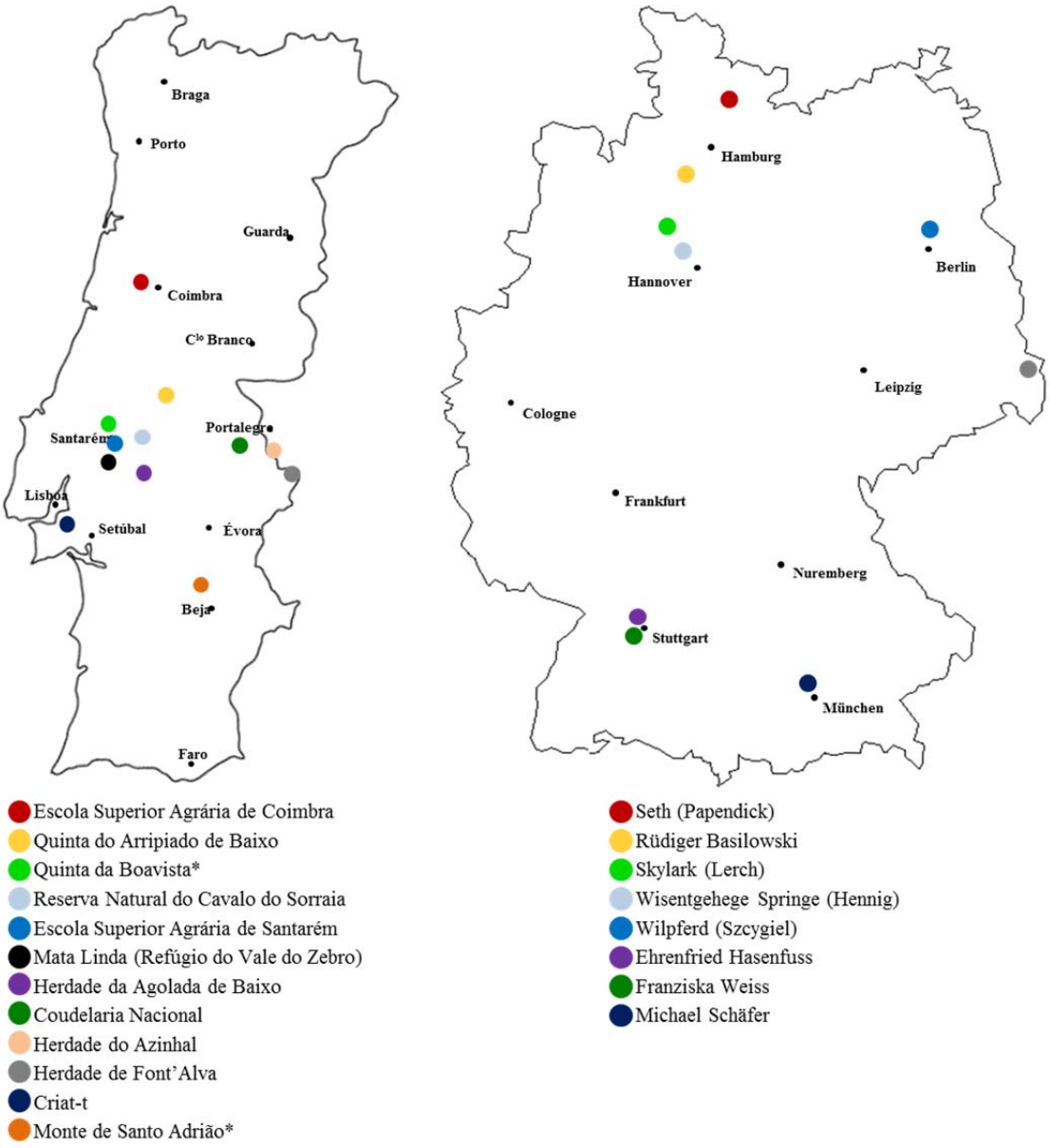


Figure 3 - Geographic distribution of Portuguese (left) and German (right) Sorraia horse breeders. Breeder legends are underneath the respective map. *Stud farms that no longer breed Sorraia horses.

A breed's phenotypic characterization is one of the most important and basic information in Animal Genetic Resources (AnGR) Conservation Management Programs (Melo *et al.* 2011), complementing historical and genetic information. By establishing the basis for which to start management breeding programs (FAO 2012), this characterization is fundamental for the establishment of AnGR national inventories and effective monitoring of these populations. Morphometric measurements also establish realistic breed morphological status and define their deviation from the breed standard

throughout generations. Sorraia morphology was studied by Oom (1992) and compared to that of the Lusitano and the Garrano. With stallion average withers height of 148cm (Oom 1992; Oom *et al.* 2004) the Sorraia is shorter than the Lusitano, with an average of 160cm for that morphometric characteristic (Oom & Ferreira 1987; Oom 1992; Solé *et al.* 2013; Vicente *et al.* 2014a) and taller than both the Garrano, with 130cm (Oom 1992; Santos & Ferreira 2012), and the Terceira Pony, with 128cm (Lopes *et al.* 2015), all of them autochthonous horse breeds from Portugal. FEI regulations (FEI 2014) regard Sorraias as ponies, but the Sorraia Breeders Association consider them as "small horses" (Oom *et al.* 2004), a fact that was demonstrated by molecular genetics as they clustered with standard-sized horses (Andalusian and Lusitano) instead of Iberian pony breeds, like the Garrano (Luis *et al.* 2007b).

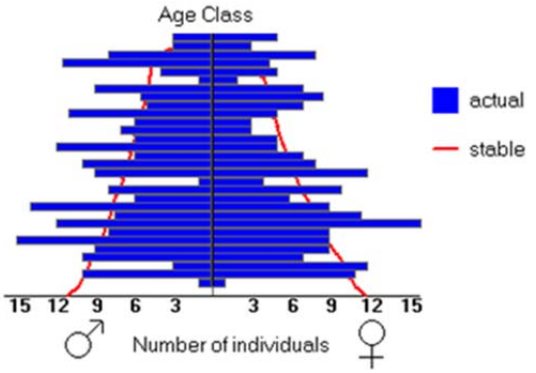


Figure 4- Age pyramid of the current Sorraia horse population (n=434).

Only about 400 Sorraia horses exist worldwide and less than 150 of them are breeding females (Figure 4). Comparing to the other Portuguese breeds, this value is much lower than the Lusitano, with ~5000 (Vicente *et al.* 2012), and the Garrano, with ~1400 (Santos & Ferreira 2012), but higher than the Terceira pony, with roughly only 100 animals in total (Lopes *et al.* 2015). These low numbers correspond to a critical maintained risk status population according to FAO criteria (FAO 2015). In addition to this, it is also listed as "in danger of extinction" by the Portuguese Government (PDR2020 2014), requiring a conservation-breeding plan for the establishment of a long-term self-sustaining population. This management has been done by the Sorraia Breeders Association, advising the breeders in the management of their individual herds, e.g. by choosing the most suitable stallion, from a genetic point of view, for their breeding mares and by promoting the rotation of stallions between stud farms.

Very often, Sorraias have shown lower breeding performance with in-hand mating. In Portugal, mares are managed extensively in most of the stud farms, and only one stallion is turned into the pasture during a breeding season (usually February to June) per year, whereas in Germany frequently one stallion is chosen per mare, per year (Oom *et al.* 2004). This strategy, whenever possible, is being followed by Portuguese breeders, namely those with only a few breeding mares and with in-hand mating practice. Mares first reproduce at 4 years and 11 months and stallions at 4 years and 6 months, with wider fertility peaks in the former than the latter (4 to 16 years versus 10 to 14 years, respectively) (Kjöllerström 2005), and also wider than in other breeds (4.2 in Lusitano mares and stallions and 5.2 years for Pura Raza Española mares) (Valera *et al.* 2000).

Inbreeding and mean kinship coefficients were calculated in SPARKS (ISIS 2011) and PM2000 (Ballou *et al.* 2002) by the additive relationship matrix (Ballou 1983) with the available complete pedigrees. Minimizing mean kinship (mK) is routinely used in zoo populations as it is the most effective way to establish an effective genetic management plan (Frankham *et al.* 2004). A description of analysed parameters is given in the end of this dissertation as a Glossary.

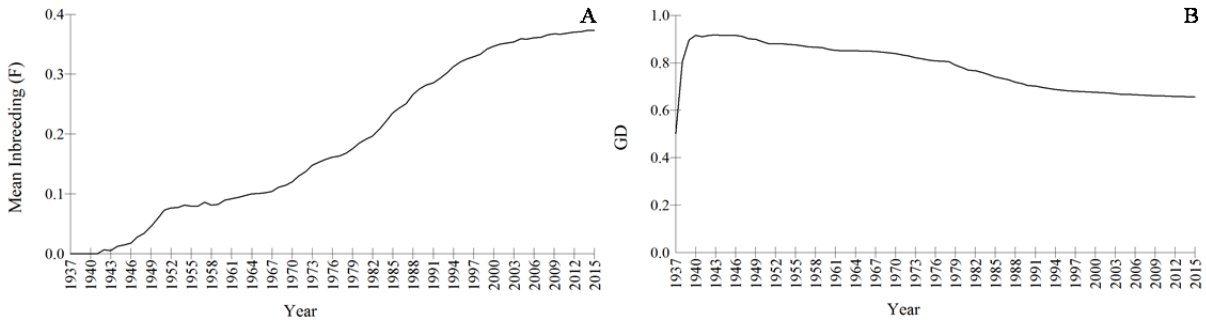


Figure 5 - Evolution of mean inbreeding coefficient, F (A) and genetic diversity (B) since the breed's foundation (1937) (n=837).

Inbreeding coefficient (F) is defined as the probability that two alleles at a *locus* are identical by descent, resulting from the mating of kin (Lacy 2000; Frankham *et al.* 2002). Inbreeding is unavoidable in small closed populations and with time all individuals become related by descent, as in the Sorraia horse breed. Due to the small number of founders, reduced effective population size and complete genetic isolation, inbreeding has increased to extremely high values (Kjöllerström 2005; Luis *et al.* 2007a) (Figure 5A) and genetic diversity has decreased (Figure 5B). Average

inbreeding coefficient, calculated from complete pedigrees back to the founders (N=434), is 0.38, ranging from 0.22 to 0.60 (Figure 5A and 6), genetic diversity is 0.65, 0.35 mean kinship, 1.43 founder genome equivalents (with a potential of reaching 2.60) and only 10 founders are currently represented, which means the contribution from two founders has been lost. The effective population size (Ne) is 11.62 over the past 8.41 generations, current Ne is 136.31 and Ne/N ratio is 0.28.

Sorraia inbreeding values (0.38) greatly exceed that of other horse breeds: 0.992 in Lusitano (Vicente *et al.* 2014b), 0.0065 in Garrano (Cipriano 2007), 0.082 in Andalusian (Gómez *et al.* 2009), 0.10 in Lipizzaner (Curik 2003), 0.13 in Thoroughbred (Cunningham *et al.* 2001), 0.10 in Finnish Standardbred Trotter and 0.004 in Finnhorse (Sairanen *et al.* 2009). The endangered Przewalski horse, recovered from only 13 founders, has an average inbreeding coefficient of 0.14 (Der Sarkissian *et al.* 2015), much lower than the Sorraia.

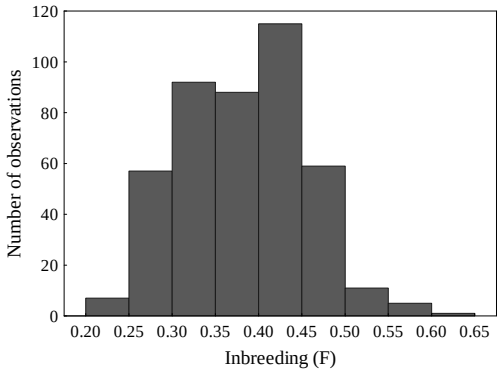


Figure 6 - Histogram of inbreeding coefficients (F) of the Sorraia horse population (n=434).

All Sorraia horses are registered in the Sorraia Horse Studbook, officially recognized by the Portuguese Minister of Agriculture, Fisheries and Forestry on January 12th 2005, who delivered to the Sorraia Breeders Association its maintenance and management (Oom *et al.* 2004). The Studbook registries are kept up-to-date in the official Equine National Registry, with new information regarding births, deaths and transfers. All information is also entered in SPARKS v1.5 software (ISIS 2011), specially designed for compiling and analysing Studbook data from zoo populations. It is important to mention that all animals undergo a mandatory parentage testing analysis with an extended microsatellite panel (due to reduced genetic variation) before their registration in the Sorraia Studbook (Oom *et al.* 2004). The general characteristics of the animals registered in the Studbook also need to be in agreement with the breed's official standard (Oom *et al.* 2004).

1.2 Genetic variability

1.2.1 Pedigree analysis

As mentioned before, all Sorraia horses that pass parentage testing confirming declared genealogy are registered in the Studbook (Oom *et al.* 2004). All Studbook data was entered, in 2004, in a software specially designed for the management of endangered captive breeding populations: SPARKS v1.5 (ISIS 2011). This allowed the genetic and demographic analysis of the population in a way not possible before (Kjöllerström 2005). SPARKS is a very powerful tool for Studbook keeping that allows basic genetic and demographic analysis (GENES and DEMOG), as well as for genetic and demographic data export to other software for more detailed studies: PM2000 (Ballou *et al.* 2002). For pedigree analysis to be as accurate and reliable as possible it is important that the most complete information on each individual is included.

Pedigree completeness affects the outcome of inbreeding depression (Smith *et al.* 1998), by underestimating inbreeding coefficient values, so that, in order for conservation programs to be successful, pedigrees must be as complete as possible and not have missing data of ancestors (Oliehoek & Bijma 2009). Pedigree registration, such as in herd books, should record the status of animals with missing parents (e.g. founder or non-founder) (Oliehoek & Bijma 2009). For these reasons the availability of complete pedigrees tracing back to the founders in the Sorraia horse, some of them with 12 complete generations (Oom *et al.* 2004), is an asset and makes this breed a perfect candidate for studies regarding demographic, genetic and inbreeding evolution, as well as for the evaluation of management breeding strategies.

The following data was entered for each animal: Studbook number, sire name, sire Studbook number, dam name, dam Studbook number, place of birth, date of birth, name, breeder, coat colour, sex and other information considered relevant for management (e.g. if the animal was castrated or if it is part of the living population to be managed) (Oom *et al.* 2004; Kjöllerström 2005) (Figure 7). Death and transfer information data were inserted whenever available. These data were complemented with

information gathered throughout the years by Prof. Maria do Mar Oom, namely regarding still births and abortions, data that were not usually reported by the breeders.

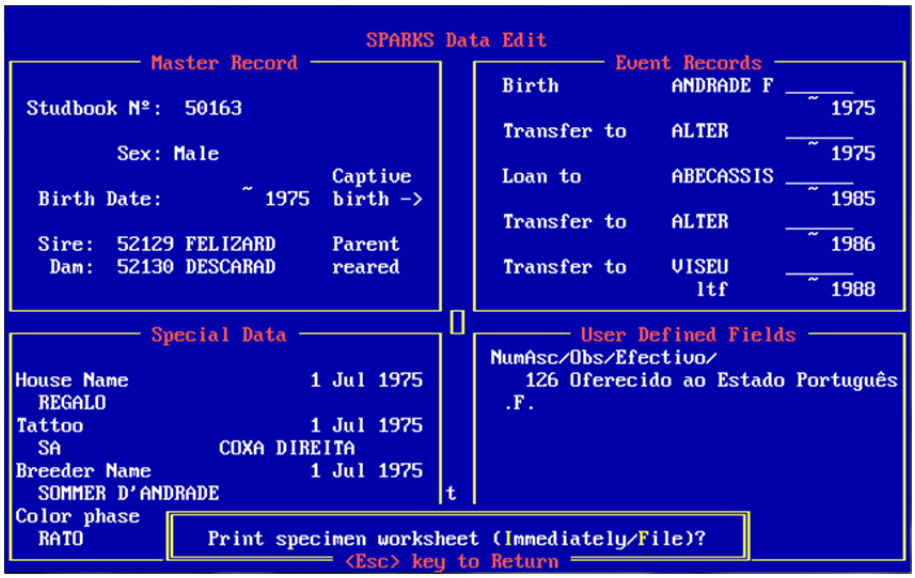


Figure 7 - Example of a Sorraia Studbook entry screen in SPARKS software.

Pedigree information has been used to assess genetic variability in the Lusitano (Vicente *et al.* 2012), Andalusian (Valera *et al.* 2005), Spanish Arab (Cervantes *et al.* 2008), Lipizzan (Zechner *et al.* 2002), German Paint (Siderits *et al.* 2013), Holstein (Roos *et al.* 2015), Old Kladruber (Vostrá-Vydrová *et al.* 2016) and in the Assateague Island wild horses (Eggert *et al.* 2010). In some of these cases, population structure was also analysed. In the Sorraia, pedigree analysis was previously used to determine current genetic and demographic status in 2005 (Kjöllerström 2005). Pedigree data has also been used to analyse the effects of inbreeding depression on fitness related traits (Kalinowski *et al.* 2000; Kalinowski & Hedrick 2001; Carolino & Gama 2008; Santana *et al.* 2012).Pedigree analyses are very important in conservation programs. Outlined goals are typically to preserve 90% of the genetic variability over 100 years, reduce drift effects, minimise founder effects, and minimise inbreeding by breeding genetically significant and/or remotely related individuals with the lowest mean kinship values (Frankham *et al.* 2002; Guerier *et al.* 2012). This analysis allowed us to calculate the following parameters: average relatedness coefficient (AR), effective number of ancestors (fa), inbreeding (F), increase in inbreeding per generation, individual increase in inbreeding (ΔF_i), founder genome equivalents (fge) and mean kinship (mk) (see Glossary and Paper 1 for detailed descriptions).

1.2.3 Molecular analysis

The morphological characterization of a breed is very important to establish its standard, but molecular characterization is also of paramount importance in conservation programmes as it allows for the quantification of extant genetic variability and conservation goals to maintain intra and inter-breed variability (Frankham *et al.* 2002; Toro *et al.* 2009). This is also true for the Sorraia horse population and was one of the main goals of this study.

The nuclear genome of horses is roughly 2.7 Mb in size, packaged into 31 pairs of autosomes and a single pair of sex chromosomes (2N=64) (Chowdhary 2013). Several molecular markers are available (Frankham *et al.* 2002; Reece 2004) and have been used in horse research (Bowling & Ruvinsky 2000; Swinburne & Lindgren 2013) such as allozymes, mitochondrial DNA (mtDNA), minisatellites, microsatellites (short tandem repeats, STRs, or simple sequence repeats, SSRs), variable number tandem repeats (VNTRs), random amplified polymorphic DNA (RAPDs), amplified fragment length polymorphism (AFLPs), restriction fragment length polymorphism (RFLPs), single nucleotide polymorphisms (SNPs), copy number variants (CNVs) and, most recently, DNA sequencing. These markers have been of paramount importance in disentangling the genetic background of simple and complex inherited diseases as well as multigenetic traits, all of which relevant to the health and welfare of horses (Bannasch 2008; Chowdhary & Raudsepp 2008).

Previous studies have revealed the low genetic variation of the Sorraia breed by analysing microsatellites (Luís *et al.* 2002b; Luis *et al.* 2007a; Luis *et al.* 2007b), proteins (Luis *et al.* 2007b), major histocompatibility complex (Luis *et al.* 2005), mitochondrial DNA (Luís *et al.* 2002a; Luis *et al.* 2006) and pedigree analysis (Kjöllerström 2005).

Microsatellites are short tandem repeat (STRs) sequences of a 1-6 bp motif (normally 10 to 30 repeats), randomly distributed throughout the genome of most species and whose function is yet unknown (Frankham *et al.* 2002; Reece 2004; Allendorf *et al.* 2012). They are highly polymorphic and highly informative, co-dominantly inherited and mostly located in non-coding regions, thus selectively neutral (Frankham *et al.* 2002; Reece 2004; Schaid *et al.* 2004; Allendorf *et al.* 2012). High polymorphism is

mainly due to high mutation rates (about 10^{-4} per generation), resulting in higher heterozygosity and allelic diversity than in other types of markers such as allozymes (Allendorf *et al.* 2012), making them suitable for intra and inter-breed genetic diversity studies (Toro *et al.* 2009). Sampling can be non-invasive and amplification can be done by PCR reaction from small amounts of DNA from different tissues (blood, hair, scat, skin, etc.) making STRs particularly useful in conservation genetic studies of endangered species (e.g. Frankham *et al.* (2002); Toro *et al.* (2009)). Adding to previous advantages, STRs can be multiplexed allowing for several *loci* to be genotyped simultaneously for many individuals, making them extremely cost-efficient and thus markers of choice in many studies (Allendorf *et al.* 2012; Swinburne & Lindgren 2013). There are some disadvantages in using STRs such as null alleles resulting from nucleotide substitutions on primer binding sites, leading to PCR amplification errors, homoplasy due to high mutation rates and primers species specificity (Allendorf *et al.* 2012).

These markers have been extensively used in parentage testing as well as in genetic diversity, population structure, conservation, evolutionary and domestication studies in several species (Takezaki & Nei 1996; Frankham *et al.* 2002; Guerier *et al.* 2012; Putman & Carbone 2014; Bruford *et al.* 2015) namely in horses (Bowling 2001; Bodó *et al.* 2005; Luis *et al.* 2007b; Leroy *et al.* 2009; Warmuth *et al.* 2011; Penedo & Raudsepp 2013; Warmuth *et al.* 2013). Ellegren *et al.* (1992) identified the first microsatellites in the horse genome and the first STR parentage testing panel was established by Bowling *et al.* (1997).

Single nucleotide polymorphisms (SNPs) are the most common form of genetic variation between individuals, mostly bi-allelic, located in both coding and non-coding regions of the genome (Reece 2004; Toro *et al.* 2009; Allendorf *et al.* 2012). Mutation rate is lower than in STRs, usually of about 10^{-3} nucleotide change per generation (Allendorf *et al.* 2012). They are useful for genetic variation estimation, parentage testing, and population structure studies (Höglund 2009; Allendorf *et al.* 2012).

SNP genotyping Beadchips are available for several species (Geneseek®), including horses (McCue & Mickelson 2013). The Ilumina 70k-Equine SNP chip used in this study was derived from the EquCab2.0 SNP collection, covering the entire genome with over 65,000 evenly distributed SNPs (Geneseek®).The availability of these SNP chips enables the identification of genes and polymorphisms potentially linked to traits of interest. Although the Ilumina 70k-Equine SNP and Affymetrix Axiom® Equine 670K

Genotyping Array (Axiom MNEC670) are commercially available, most published studies have been done using the previous 50K-Equine SNP chip (McCue & Mickelson 2013). Some of these studies involved genome-wide association mapping of simple and complex traits (Orr *et al.* 2010; Raudsepp *et al.* 2012; McCue & Mickelson 2013), resolving the relationship between equine breeds and studying their genetic diversity (McCue *et al.* 2012; Petersen *et al.* 2013), or screening for chromosomal abnormalities (Holl *et al.* 2013).

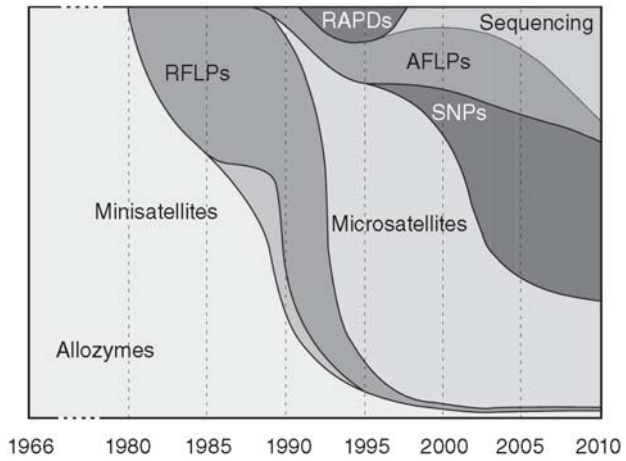


Figure 8 - Relative popularity of molecular markers in conservation genetics. Each time point (horizontal axis) has the corresponding proportion of molecular markers used (vertical axis) (Allendorf *et al.* 2012).

In recent years, SNPs have begun to replace STRs (Figure 8) due to fast and efficient genotyping, lower error rate, ease of standardization between laboratories, and short size so that they can be amplified from fragments less than 50bp long (very useful in studies of degraded DNA such as ancient DNA or forensics) (Allendorf *et al.* 2012; Swinburne & Lindgren 2013). Nevertheless, as SNPs are mostly bi-allelic and less informative than STRs, significantly more SNPs are needed to achieve similar levels of information of STRs (Schaid *et al.* 2004). Another drawback is that SNPs demand high informatics capabilities (Schaid *et al.* 2004).

Copy number variants (CNVs) are common attributes of vertebrate genomes and significant sources of genetic variability (Fontanesi *et al.* 2010; Gurgul *et al.* 2014). Consisting of insertions, deletions, duplications and other complex rearrangements such as translocations or inversions, they result in structural differences between genomes (Alkan *et al.* 2011; Cantsilieris & White 2013; Ghosh *et al.* 2014b). Located on coding

or non-coding regions of the genome, they can potentially be non-neutral markers (Gurgul *et al.* 2014). Typically varying in size between 50 base-pairs to several megabase-pairs, they contribute to phenotypic diversity but can also be disease-associated (Fontanesi *et al.* 2010; Ghosh *et al.* 2014b). Copy number variants studies are based on whole-genome genotyping (using different next generation sequencing (NGS) technologies)) or on specially designed microarrays (Gurgul *et al.* 2014). In 2011, a horse whole genome oligonucleotide tiling array became available for the horse (Qu *et al.* 2011), being the first high resolution tool for CNV discovery (Raudsepp & Chowdhary 2013). It comprises over 400,000 probes, of which 82,354 are located on exons, 519 on the Y chromosome, and 26,790 probes on subtelomeric regions (Raudsepp & Chowdhary 2013). In livestock, CNVs have been used to research disease susceptibility, developmental disorders and morphological traits (Gurgul *et al.* 2014). In horses, a few studies described horse-specific CNVs (Doan *et al.* 2012a; Doan *et al.* 2012b; Ghosh *et al.* 2014b; Wang *et al.* 2014) and one linked CNVs to horse body size (Metzger *et al.* 2013). Other studies in horses looked into the genetic background of diseases with economic relevance to the equine industry such as: melanomas (Sundstrom *et al.* 2012); recurrent laryngeal neuropathy (Dupuis *et al.* 2013); *Rhodococcus equi* susceptibility (McQueen *et al.* 2014), cryptorchidism (Ghosh *et al.* 2014a); recurrent airway obstruction (Ghosh *et al.* 2016). In our study, we used this tool for genetic variability analysis and not for complex traits or genetic disorders association.

Low resolution of detection platforms hinders the detection of small CNVs (Gurgul *et al.* 2014). In addition, the use of different platforms makes it difficult to compare results between studies making alternative confirmation methods essential (Cantsilieris & White 2013). In order to increase confidence in copy number data, bias should be identified and minimized as copy number variations results can be affected, for example, by sample quality and assay choice (Cantsilieris & White 2013). Also, if no platform or genome assembly are available, cross-species analysis will be necessary, potentially compromising CNV detection and sensitivity (Fontanesi *et al.* 2010). Due to lesser bias, the ability to detect higher number of CNVs per experiment and appropriateness to any species, high-throughput NGS methods, some with longer reads, can counter the above mentioned drawbacks (Gurgul *et al.* 2014). However, these methods call for high computational resources and technical issues can also negatively impact the results (Alkan *et al.* 2011).

Single nucleotide polymorphisms (SNPs) and copy number variants (CNVs) are useful in disentangling the effects of history and demography from natural selection in a genome-wide fashion, but are also valuable in determining the genetic variation within and between populations (Allendorf *et al.* 2012). Both SNPs and CNVs have been described in the horse (Doan *et al.* 2012a; McCue *et al.* 2012; McCue & Mickelson 2013; Petersen *et al.* 2013; Ghosh *et al.* 2014b). Both can be located in coding regions of the genome and are thus non-neutral markers (Reece 2004; Allendorf *et al.* 2012). Mutations in these markers may change the amino-acid sequence or change gene expression, impairing protein function and reducing fitness (Reece 2004; Allendorf *et al.* 2012).

Although molecular variability has been described in the Sorraia horse before, we propose to combine three different methodologies, as well as using SNPs and CNVs for the first time. We hope to provide a better understanding of the remaining genome-wide genetic variability of this extremely endangered breed, while simultaneously providing new and more variable markers than those currently used in parentage testing and that are fixed in this population. We expect these results will aid the management breeding programme implemented by the Sorraia Breeders Association and help preserve this critically endangered and iconic animal genetic resource.

Using microsatellites, single nucleotide polymorphisms and copy number variants gave us an understanding of the extant genome-wide variability of this extremely endangered breed. All of these markers have been previously described in horses in studies of genetic variability, diseases, and evolution.

1.3 Reproductive fitness and inbreeding

By increasing homozygosity from the mating of kin, inbreeding exposes recessive deleterious alleles and might cause reduced fitness, also known as inbreeding depression (Lacy 2000; Frankham *et al.* 2002; Charpentier *et al.* 2007; Charlesworth & Willis 2009; Leroy 2014). Some studies reported no evidence of these recessive deleterious alleles, suggesting their purging from the genetic pool, although this purging could still have an impact in the viability of small inbred populations through inbreeding depression (Charpentier *et al.* 2007). Inbreeding depression affects different aspects of biological systems and is of paramount importance in the conservation of small isolated natural or captive populations as the Sorraia horse.

Domestic or wild animals have been subjected to inbreeding depression for a long time, particularly evident on fitness related traits such as growth, birth weight and juvenile survival (Swiger *et al.* 1961; Ralls *et al.* 1979; Mc Parland *et al.* 2007; Santana *et al.* 2012). According to Ralls *et al.* (1979), inbreeding lowered survival on 41 out of 44 species of ungulates, including horses. If animals survive until adulthood, inbreeding depression might reduce survival, growth rate, fertility, ability to mate, fecundity and lead to insufficient parental care (Ryan *et al.* 2002; Charlesworth & Willis 2009; Collins *et al.* 2012). Santana *et al.* (2012) showed that animals from Brazilian cattle breeds could support about 20% of inbreeding before showing inbreeding depression effects on some traits, although in other traits even a very small inbreeding level (1.9%) had negative impacts. Hip height in Nellore cattle is negatively influenced by inbreeding (Santana *et al.* 2010) and, in general, highly inbred animals have been shown to be smaller and narrower than non-inbred ones (Smith *et al.* 1998; Croquet *et al.* 2006).

Inbreeding depression on morphological traits in horses has been reported in the Italian Haflinger (Gandini *et al.* 1992), Spanish Purebred (Gómez *et al.* 2009) and Lusitano (Oom 1992; Vicente *et al.* 2014b). Pedigree information has been used to assess the effects of inbreeding depression on fitness related traits (Kalinowski *et al.* 2000; Kalinowski & Hedrick 2001; Carolino & Gama 2008; Santana *et al.* 2012).

Despite extremely high inbreeding, little is known about inbreeding depression in Sorraia horses. It is relevant to determine the effect such a high inbreeding level might

have on the population's fitness, looking for the likely causes of the low fertility and viability observed, namely by investigating the relationships between offspring and parental inbreeding coefficients and the viability of new-borns, fertility, foaling intervals and age at first parturition. This should give a better understanding of the consequences of inbreeding in the Sorraia horse breed and provide a base for a management-breeding program that will establish a long-term self-sustaining population. Pariacote *et al.* (1998) suggested that selection through management breeding programs could mitigate the negative effects of inbreeding depression, so only by implementing an updated breeding strategy will the preservation of this extremely important and critically endangered animal genetic resource be possible.

As pedigree completeness influences the results and effects of inbreeding depression (Smith *et al.* 1998), and due to its extremely high inbreeding levels already discussed above, the Sorraia horse is an excellent model to study this subject as all animals have complete pedigree information tracing back to the founders, some with 12 complete generations (Oom *et al.* 2004).

As mentioned before, horses have 64 well characterized metacentric, sub-metacentric or acrocentric (ISCNH 1997) chromosomes (2N=64), grouped into 31 pairs of autosomes and a single pair of sex chromosomes (two X chromosomes in mares; one X and one Y in stallions) (Chowdhary 2013). Cytogenetic research in horses has shown that chromosomal abnormalities are frequently associated to infertility/subfertility, repeated early embryonic death, abortion and stillbirth, although occasional pregnancies may occur (Breen *et al.* 1997; Chowdhary & Raudsepp 2000; Bugno *et al.* 2008; Lear & Bailey 2008). Mares that only produce few offspring during their life-time should be considered potential chromosome aberrations carriers (Lear & Bailey 2008; Vanderwall 2008), as generating offspring does not exclude chromosomal abnormalities. Most mares with sex chromosome abnormalities have normal phenotypes, requiring karyotyping for definitive diagnosis. The most common chromosomal abnormality in horses is X monosomy (63,X0) (Payne *et al.* 1968), representing over 50% of all reported abnormalities (Bugno *et al.* 2001). The second most common in infertile mares is XY sex reversal, followed by XX sex reversal, mosaicism, chromosomal rearrangements, the rarest being XXX trisomy (Lear & Bailey 2008). Chromosomal mosaicism is the presence of two or more chromosomally distinct cell lines in an individual and accounts for about 30% of horse chromosomal abnormalities (Lear & Bailey 2008). However, chromosomal abnormalities are only a small part of the causes

of reduced fertility or infertility in horses and the underlying genetic causes remain unclear (Raudsepp *et al.* 2013). Cytogenetic analyses, even conventional ones, have never been performed in the Sorraia breed, despite several identified cases of infertility and subfertility.

Stallion fertility is a complex trait regulated by several factors such as: environmental, physiological, behavioural, epigenetic and genetic (Raudsepp *et al.* 2013). There is little knowledge regarding genetic factors and mainly limited to chromosomal aberrations. As for mares, stallions with balanced translocations can be sub fertile because of affected spermatogenesis that can reduce embryo viability, but still have an overall normal phenotype (Raudsepp *et al.* 2013; Raudsepp & Chowdhary 2016). Conventional karyotyping is thus one of the first steps in evaluating their genetic health (Durkin *et al.* 2011). Additionally, sperm's chromosome complement can be quickly assessed directly on decondensed sperm heads by fluorescence *in situ* hybridization (FISH). This sperm-FISH method was developed initially for humans (Wyrobek *et al.* 1990) but is now a cutting-edge technique for detecting meiotic chromosome defects in domestic species sperm (Raudsepp & Chowdhary 2016). It has been optimized for stallion sperm (Bugno-Poniewierska *et al.* 2009) and used for the study of sex chromosome aneuploidies in reproductively normal stallions (Bugno *et al.* 2010) and to evaluate the correlation between stallion age and rate of chromosome abnormalities (Bugno-Poniewierska *et al.* 2011; Bugno-Poniewierska *et al.* 2014). To the best of our knowledge there have been no sperm-FISH studies on sub fertile or infertile stallions so far, including the Sorraia horse.

Equine reproduction is a very complex field where little is known regarding the part that chromosomal structural rearrangements, CNVs, gene expression regulation, and epigenetic mechanisms have in in fertility (Raudsepp *et al.* 2013). With samples with well-defined phenotypes, the new genomic tools bring high hopes of detecting what affects stallion and mare reproduction (Raudsepp *et al.* 2013). Genome-wide association studies have recently revealed genes and genomic regions potentially associated with stallion fertility (Schrimpf *et al.* 2014; Schrimpf *et al.* 2015) and also conditions underlying subfertility (Raudsepp *et al.* 2012). One of these genes is *FK506-binding protein 6* (*FKBP6*), first found in humans (Crackower *et al.* 2003; Miyamato *et al.* 2006), proposed as a susceptibility *locus* for subfertility in Thoroughbred stallions by impaired acrosome reaction (IAR) (Raudsepp *et al.* 2012). Stallions with IAR have a double homozygous AA-AA genotype consisting of SNPs g.11040315G>A and

g.11040379C>A in *FKBP6* exon 4. An association between *FKBP6* and stallion fertility has also been recently reported in Hanoverian stallions (Schrimpf *et al.* 2015) were SNP g.11040379C>A was associated with higher conception rates in A/A stallions and lower in C/C stallions. However, the double homozygous A/A-A/A genotype was not associated with fertility measured as pregnancy rate per oestrus cycle (Schrimpf *et al.* 2015). Association between FKBP6 variants and stallion fertility appear to be different depending on breed tested and should thus be further analysed in order to have a better understanding the role of the genetic variants involved. Through the analysis of sub fertile Sorraia stallions by karyotyping, sperm-FISH and *FKBP6* exon 4 genotyping, we aimed at determining the prevalence of chromosome abnormalities, sex chromosomes meiotic segregation errors and frequency of IAR susceptibility genotype, respectively. This will hopefully aid in stallion selection for the management breeding program.

1.4 - Specific aims

With the publication of the Sorraia horse Studbook in 2004 (Oom *et al.* 2004), records from all Sorraia horses became available. This provided an opportunity to investigate the evolution of this endangered breed using different tools and approaches (e.g. pedigree, genetic, demographic and morphometric analysis; genome-wide genetic variability assessment; inbreeding and inbreeding depression; reproductive fitness). A Sorraia DNA bank with samples from over 400 animals was established throughout the duration of this PhD. This allowed all succeeding analyses to be performed and will also allow the continuing study of this breed for years to come when new tools and projects arise.

Specifically, this project aimed at studying:

1) Genetic diversity and demographic structure evolution through time using pedigree analysis (N=653) - *Paper 1*
2) Morphometric evolution of stallions over the past 20 years and the effects of inbreeding depression on morphological traits (N=40) - *Paper 2*
3) Genome-wide genetic variability through microsatellites (N=190), single nucleotide polymorphisms (N=48) and copy number variations (N=14) - *Paper 3*
4) Inbreeding depression on fitness related traits through pedigree analysis (N=749) - *Paper 4*
5) Cytogenetic evaluation using classic and molecular cytogenetic methods (N=28) - *Paper 5*
6) Fertility analysis of stallions by sex chromosome content of sperm cells analysis (N=6) and sequencing/genotyping of fertility-related gene *FKBP6* (N=11) - *Paper 6*

1.5 - Thesis Structure

This thesis is comprised of a comprehensive introduction where breed history, current status, genetic variability, inbreeding and multidisciplinary approaches to achieve the outlined objectives is explained. This is followed by two chapters, each one comprised of three papers that intended to answer the questions raised in our above-mentioned aims. Next, we present a discussion where all topics are approached and results unified. This discussion is mainly focused on presenting the findings of this work. For more detailed results and breed comparisons, the respective papers should be consulted. In the end, the Final remarks chapter will cover what was achieved and what should be the future prospects for the Sorraia horse breed. Description of analysed parameters is provided in the end of this thesis as a Glossary.

Below we present a short explanation in how each paper answers the questions raised in the thesis' aims.

The genetic and genomic analysis herein would not have been possible without the accomplishment of the Sorraia DNA bank. It was a priority and was kept up-to-date. This aim allowed us to do the needed research to produce the results for aims #3 and #6.

Populations that are under conservation management breeding programmes, such as the Sorraia horse, should be routinely evaluated for their current status through the analysis of pedigree data, available in the Studbook (Oom *et al.* 2004; FAO 2015). This will contribute to a more thorough knowledge of the breeds' genetic evolution and demographic structure, allowing better management and a method to evaluate the success of the conservation strategies in place. With this in mind, we conducted a study that led to **Chapter 2 - Paper 1**, "*Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis*"; which answers these questions and served as a stepping stone to the rest of the PhD project. This paper was the result of a Master's Thesis by Márcia Pinheiro, explaining why first authorship is hers. For this paper, data entry was done by me as well part of the demographic analysis, manuscript preparation, submission and revision/review. A thorough morphological description, morphological evolution and impact of inbreeding depression, coat colour and age on morphological traits are of paramount importance for

breed characterization and can be added to the management breeding program currently in place. The objective was thus to morphologically describe the Sorraia horse breed while trying to assess mainly the impact of inbreeding depression, but also coat colour and age, on its conformation evolution over the last 20 years. These questions in aim #2 are answered by **Chapter 2 - Paper 2** "*Morphological characterization and inbreeding effects in the endangered Sorraia Horse breed*".

Aim #3, describing the genetic variability of the Sorraia horse using different molecular approaches (microsatellites, single nucleotide polymorphisms and copy number variations) to get a genome-wide understanding of the remaining variation of this extremely endangered breed, resulted in **Chapter 2 - Paper 3** "*Genome-wide variability in the Sorraia horse*".

Although Sorraia horses are known for their extremely high inbreeding coefficients, there is little information regarding the effects on fitness related traits. It was thus of utmost importance to determine its effect on fertility levels, observed viability and other life-history traits. This was done by studying the relationships between offspring and parental inbreeding and the viability of new-borns; between inbreeding and fertility in both sexes; inbreeding and foaling intervals and age at first parturition. This study resulted in **Chapter 3 - Paper 4** "*Impact of inbreeding on fitness-related traits in the highly threatened Sorraia horse breed*", answering the questions in aim #4 and providing new insights that can be added to the current management-breeding program.

Fertility is of utmost importance for the long-term self-sustaining of the extant population. It has been shown that chromosomal abnormalities, especially on sex chromosomes, are associated with infertility or subfertility in horses. To date, no cytogenetic studies have been performed in this breed to assess the prevalence of chromosome abnormalities in animals with low breeding performance and ambiguous sexual phenotypes. **Chapter 3 - Paper 5** "*First evidence of sex chromosome mosaicism in the endangered Sorraia Horse breed*" and other results that are mentioned in the Discussion section of this thesis answer the questions raised in aim #5.

Finally, aim #6 was answered by **Chapter 3 - Paper 6** "*Fertility assessment in Sorraia stallions by sperm-FISH and FKBP6 genotyping*", where karyotyping, sperm-FISH and *FKBP6* exon 4 genotyping where done in sub fertile Sorraia stallions to determine the prevalence of chromosome abnormalities, meiotic segregation errors of the sex chromosomes and the frequency of the IAR-susceptibility genotype, respectively.

We expect that all these results will aid the Sorraia Breeders Association genetic management program and to support the best advices to the breeders, improving the breeds' genetic health and also help preserve and prevent the permanent loss of this iconic and important animal genetic resource.

References

Alkan C., Coe B.P. & Eichler E.E. (2011) Genome structural variation discovery and genotyping. *Nature reviews. Genetics* **12**, 363-76.

Allendorf F.W., Luikart G.H. & Aitken S.N. (2012) *Conservation and the Genetics of Populations (2)*. Wiley-Blackwell, Somerset, GB.

Andrade R. (1926) *Apontamentos para um estudo sobre a origem e domesticação do cavalo na Península Ibérica. Aproximações.* Centro Tipográfico Colonial, Lisboa, Portugal.

Andrade R. (1937) Les chevaux du Sorraia. *Comptes Rendues du XIIe Congrès International de Zoologie* **1935**, 2368-70.

Andrade R. (1945) O Cavalo do Sorraia. *Boletim Pecuário* **XIII**, 1-13.

Andrade R. (1954) *Alrededor del Caballo Español*. Coleccion de Estudios, Lisboa, Portugal.

Ballou J. (1983) Calculating inbreeding coefficients from pedigrees. In: *Genetics and Conservation: A Reference for Managing Wild Animal and Plant Populations* (eds. by Schonewald-Cox CM, Chambers SM, MacBryde B & Thomas L), pp. 509-20, Menlo Park, California: Benjamin/Cummings.

Ballou J.D., Lacy R.C. & Pollak J.P. (2002) Population Management 2000. Chicago Zoological Society.

Bannasch D. (2008) Genetic Testing and the Future of Equine Genomics. *Journal of Equine Veterinary Science* **28**, 645-9.

Bennett D. & Hoffmann R.S. (1999) Equus caballus. *Mammalian Species*, 1-14.

Bodó I., Alderson L. & Langlois B. (2005) *Conservation Genetics of Endangered Horse Breeds*. Wageningen Academic Publishers.

Bouman I. (1989) Origins of the domestic horse. *Przewalski Horse*, 32-43.

Bowling A.T. (2001) Historical development and application of molecular genetic tests for horse identification and parentage control. *Livest Prod Sci* **72**, 111-6.

Bowling A.T., Eggleston-Stott M.L., Byrns G., Clark R.S., Dileanis S. & Wictum E. (1997) Validation of microsatellite markers for routine horse parentage testing. *Anim Genet* **28**, 247-52.

Bowling A.T. & Ruvinsky A. (2000) *The Genetics of the Horse*. CABI, Wallingford, UK.

Breen M., Langford C.F., Carter N.P., Fischer P.E., Marti E., Gerstenberg C., Allen W.R., Lear T.L. & Binns M.M. (1997) Detection of equine X chromosome abnormalities in equids using a horse X whole chromosome paint probe (WCPP). *Vet J* **153**, 235-8.

Bruford M.W., Ginja C., Hoffmann I., Joost S., Orozco-terWengel P., Alberto F.J., Amaral A.J., Barbato M., Biscarini F., Colli L., Costa M., Curik I., Duruz S., Ferencakovic M., Fischer D., Fitak R., Groeneveld L.F., Hall S.J., Hanotte O., Hassan F.U., Helsen P., Iacolina L., Kantanen J., Leempoel K., Lenstra J.A., Ajmone-Marsan P., Masembe C., Megens H.J., Miele M., Neuditschko M., Nicolazzi E.L., Pompanon F., Roosen J., Sevane N., Smetko A., Stambuk A., Streeter I., Stucki S., Supakorn C., Telo Da Gama L., Tixier-Boichard M., Wegmann D. & Zhan X. (2015) Prospects and challenges for the conservation of farm animal genomic resources, 2015-2025. *Front Genet* **6**, 314.

Bugno-Poniewierska M., Jablonska Z. & Slota E. (2009) Modification of equine sperm chromatin decondensation method to use fluorescence in situ hybridization (FISH). *Folia Histochem Cytobiol* **47**, 663-6.

Bugno-Poniewierska M., Kozub D., Pawlina K., Tischner M., Jr., Tischner M., Slota E. & Wnuk M. (2011) Determination of the correlation between stallion's age and number of sex chromosome aberrations in spermatozoa. *Reprod Domest Anim* **46**, 787-92.

Bugno-Poniewierska M., Pawlina K., Tischner M. & Tischner M. (2014) Age-related effects on sex chromosome aberrations in equine spermatozoa. *Journal of Equine Veterinary Science* **34**, 34.

Bugno M., Jablonska Z., Tischner M., Klukowska-Rotzler J., Pienkowska-Schelling A., Schelling C. & Slota E. (2010) Detection of sex chromosome aneuploidy in equine spermatozoa using fluorescence in situ hybridization. *Reprod Domest Anim* **45**, 1015-9.

Bugno M., Slota E. & Zabek T. (2001) Two cases of subfertile mares with 64,XX/63,X mosaic karyotype. *Ann. Anim. Sci.* **1**, 7–11.

Bugno M., Zabek T., Golonka P., Pienkowska-Schelling A., Schelling C. & Slota E. (2008) A case of an intersex horse with 63,X/64,XX/65,XX,del(Y)(q?) karyotype. *Cytogenetic and genome research* **120**, 123-6.

Cantsilieris S. & White S.J. (2013) Correlating multiallelic copy number polymorphisms with disease susceptibility. *Hum Mutat* **34**, 1-13.

Carolino N. & Gama L.T. (2008) Inbreeding depression on beef cattle traits: estimates, linearity of effects and heterogeneity among sire-families. *Genet Sel Evol* **40**, 511-27.

Cervantes I., Molina A., Goyache F., Gutiérrez J.P. & Valera M. (2008) Population history and genetic variability in the Spanish Arab Horse assessed via pedigree analysis. *Livest Sci* **113**, 24-33.

Charlesworth D. & Willis J.H. (2009) The genetics of inbreeding depression. *Nature reviews. Genetics* **10**, 783-96.

Charpentier M.J.E., Widdig A. & Alberts S.C. (2007) Inbreeding depression in non-human primates: a historical review of methods used and empirical data. *Am J Primatology* **69**, 1370-86.

Chowdhary B.P. (2013) Defining the Equine Genome: The Nuclear Genome and the Mitochondrial Genome. In: *Equine Genomics* (pp. 1-9. Blackwell Publishing Ltd.

Chowdhary B.P. & Raudsepp T. (2000) Cytogenetics and physical gene maps. In: *The genetics of the horse* (eds. by Bowling AT & Ruvinsky A), pp. 171-241. CABI Publishing, Oxon, UK.

Chowdhary B.P. & Raudsepp T. (2008) The horse genome derby: racing from map to whole genome sequence. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **16**, 109-27.

Cipriano F.G.F. (2007) Diversidade genética da raça garrana baseada em MTDNA, microssatélites e dados genealógicos. p. 137.

Collins C.W., Songsasen N.S., Vick M.M., Wolfe B.A., Weiss R.B., Keefer C.L. & Monfort S.L. (2012) Abnormal reproductive patterns in Przewalski's mares are associated with a loss in gene diversity. *Biol Reprod* **86 (2)**, 1-10.

Crackower M.A., Kolas N.K., Noguchi J., Sarao R., Kikuchi K., Kaneko H., Kobayashi E., Kawai Y., Kozieradzki I., Landers R., Mo R., Hui C.C., Nieves E., Cohen P.E., Osborne L.R., Wada T., Kunieda T., Moens P.B. & Penninger J.M. (2003) Essential role of Fkbp6 in male fertility and homologous chromosome pairing in meiosis. *Science* **300**, 1291-5.

Croquet C., Mayeres P., Gillon A., Vanderick S. & Gengler N. (2006) Inbreeding Depression for Global and Partial Economic Indexes, Production, Type, and Functional Traits. *J Dairy Sci* **89**, 2257-67.

Cunningham E.P., Dooley J.J., Splan R.K. & Bradley D.G. (2001) Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. *Anim Genet* **32**, 360-4.

Curik I. (2003) Inbreeding, Microsatellite Heterozygosity, and Morphological Traits in Lipizzan Horses. *J Hered* **94**, 125-32.

Der Sarkissian C., Ermini L., Schubert M., Yang Melinda A., Librado P., Fumagalli M., Jónsson H., Bar-Gal Gila K., Albrechtsen A., Vieira Filipe G., Petersen B., Ginolhac A., Seguin-Orlando A., Magnussen K., Fages A., Gamba C., Lorente-Galdos B., Polani S., Steiner C., Neuditschko M., Jagannathan V., Feh C., Greenblatt Charles L., Ludwig A., Abramson Natalia I., Zimmermann W., Schafberg R., Tikhonov A., Sicheritz-Ponten T., Willerslev E., Marques-Bonet T., Ryder Oliver A., McCue M., Rieder S., Leeb T., Slatkin M. & Orlando L. (2015) Evolutionary Genomics and Conservation of the Endangered Przewalski's Horse. *Curr Biol* **25**, 2577-83.

Doan R., Cohen N., Harrington J., Veazey K., Juras R., Cothran G., McCue M.E., Skow L. & Dindot S.V. (2012a) Identification of copy number variants in horses. *Genome Res* **22**, 899-907.

Doan R., Cohen N.D., Sawyer J., Ghaffari N., Johnson C.D. & Dindot S.V. (2012b) Whole-genome sequencing and genetic variant analysis of a Quarter Horse mare. *BMC Genomics* **13**, 78.

Dupuis M.C., Zhang Z., Durkin K., Charlier C., Lekeux P. & Georges M. (2013) Detection of copy number variants in the horse genome and examination of their association with recurrent laryngeal neuropathy. *Anim Genet* **44**, 206-8.

Durkin K.W., Raudsepp T. & Chowdhary B.P. (2011) Cytogenetic evaluation. In: *Equine Reproduction* (eds. by McKinnon AO, Squires EL, Vaala WE & Varner DD), pp. 1462-8. Wiley-Blackwell.

Eggert L.S., Powell D.M., Ballou J.D., Malo A.F., Turner A., Kumer J., Zimmerman C., Fleischer R.C. & Maldonado J.E. (2010) Pedigrees and the Study of the Wild Horse Population of Assateague Island National Seashore. *The Journal of Wildlife Management* **74**, 963-73.

Ellegren H., Johansson M., Sandberg K. & Andersson L. (1992) Cloning of highly polymorphic microsatellites in the horse. *Anim Genet* **23**, 133-42.

FAO (2012) Phenotypic characterization of animal genetic resources. FAO Animal Production and Health Guidelines, Rome.

FAO (2015) The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture. (eds. by Scherf BD & Pilling D). FAO Commission on Genetic Resources for Food and Agriculture Assessments, Rome.

FEI (2014) Dressage Rules. (ed. by Internationale FE), p. 124. Fédération Equestre Internationale, Lausanne, Switzerland.

Fontanesi L., Martelli P.L., Beretti F., Riggio V., Dall'Olio S., Colombo M., Casadio R., Russo V. & Portolano B. (2010) An initial comparative map of copy number variations in the goat (Capra hircus) genome. *BMC Genomics* **11**, 639.

Frankham R., Ballou J.D. & Briscoe D.A. (2004) *A Primer of Conservation Genetics*. Cambridge University Press.

Frankham R., Briscoe D.A. & Ballou J.D. (2002) *Introduction to Conservation Genetics*. Cambridge University Press.

Gandini G.C., Bagnato A., Miglior F. & Pagnacco G. (1992) Inbreeding in the Italian Haflinger horse. *J Anim Breed Genet* **109**, 433-43.

Ghosh S., Das P.J., Arnold C., Jaxheimer J., Varner D.D., Chowdhary B.P. & Raudsepp T. (2014a) Contribution of DNA copy number variants to equine cryptorchidism. *Journal of Equine Veterinary Science* **34**, 29.

Ghosh S., Das P.J., McQueen C.M., Gerber V., Swiderski C.E., Lavoie J.P., B P.C. & Raudsepp T. (2016) Analysis of genomic copy number variation in equine recurrent airway obstruction (heaves). *Anim Genet*.

Ghosh S., Qu Z., Das P.J., Fang E., Juras R., Cothran E.G., McDonell S., Kenney D.G., Lear T.L., Adelson D.L., Chowdhary B.P. & Raudsepp T. (2014b) Copy Number Variation in the Horse Genome. *PLoS Genet* **10**, e1004712.

Gómez M.D., Valera M., Molina A., Gutiérrez J.P. & Goyache F. (2009) Assessment of inbreeding depression for body measurements in Spanish Purebred (Andalusian) horses. *Livest Sci* **122**, 149-55.

Gonzaga P.G. (2004) *A History of the Horse, Volume 1: The Iberian Horse from Ice Age to Antiquity*. JA Allen & Co. Ltd., London, UK.

Guerier A.S., Bishop J.M., Crawford S.J., Schmidt-Küntzel A. & Stratford K.J. (2012) Parentage analysis in a managed free ranging population of southern white rhinoceros: genetic diversity, pedigrees and management. *Conservation Genetics* **13**, 811-22.

Gurgul A., Semik E., Pawlina K., Szmatola T., Jasielczuk I. & Bugno-Poniewierska M. (2014) The application of genome-wide SNP genotyping methods in studies on livestock genomes. *J Appl Genet* **55**, 197-208.

Höglund J. (2009) *Evolutionary Conservation Genetics*. OUP Oxford.

Holl H.M., Lear T.L., Nolen-Walston R.D., Slack J. & Brooks S.A. (2013) Detection of two equine trisomies using SNP-CGH. *Mamm Genome* **24**, 252-6.

ISCNH (1997) International system for cytogenetic nomenclature of the domestic horse. *Chromosome research* **5**, 433-43.

ISIS (2011) SPARKS (Single Population Analysis and Records Keeping System). International Species Information System, Eagan,MN.

Kalinowski S.T. & Hedrick P.W. (2001) Inbreeding depression in captive bighorn sheep. *Anim Conserv* **4**, 319-24.

Kalinowski S.T., Hedrick P.W. & Miller P.S. (2000) Inbreeding Depression in the Speke's Gazelle Captive Breeding Program. *Conserv Biol* **14**, 1375-84.

Kjöllerström H.J. (2005) O Studbook como instrumento na análise genética e demográfica de populações ameaçadas: o caso do Cavalo do Sorraia (in Portuguese). In: *Animal Biology Department*, p. 50. Faculty of Sciences, University of Lisbon, Portugal, University of Lisbon, Portugal.

Lacy R.C. (2000) Considering Threats to the Viability of Small Populations Using Individual-Based Models. *Ecological Bulletins*, 39-51.

Lear T.L. & Bailey E. (2008) Equine clinical cytogenetics: the past and future. *Cytogenetic and genome research* **120**, 42-9.

Leroy G. (2014) Inbreeding depression in livestock species: review and meta-analysis. *Anim Genet* **45**, 618-28.

Leroy G., Callede L., Verrier E., Meriaux J.C., Ricard A., Danchin-Burge C. & Rognon X. (2009) Genetic diversity of a large set of horse breeds raised in France assessed by microsatellite polymorphism. *Genet Sel Evol* **41**, 5.

Lopes M.S., Mendonça D., Rojer H., Cabral V., Bettencourt S.X. & da Câmara Machado A. (2015) Morphological and genetic characterization of an emerging Azorean horse breed: the Terceira Pony. *Front Genet* **6**, 62.

Luis C., Bastos-Silveira C., Cothran E.G. & Oom M.M. (2006) Iberian origins of New World horse breeds. *J Hered* **97**, 107-13.

Luís C., Bastos-Silveira C., Cothran E.G. & Oom M.M. (2002a) Variation in the mitochondrial control region sequence between the two maternal lines of the Sorraia horse breed. *Genetics and Molecular Biology* **25**, 309-11.

Luis C., Cothran E.G. & Oom M.M. (2007a) Inbreeding and genetic structure in the endangered Sorraia horse breed: implications for its conservation and management. *J Hered* **98**, 232-7.

Luís C., Cothran E.G. & Oom M.M. (2002b) Microsatellites in Portuguese autochthonous horse breeds: usefulness for parentage testing. *Genetics and Molecular Biology* **25**, 131-4.

Luis C., Cothran E.G., Oom M.M. & Bailey E. (2005) Major histocompatibility complex locus DRA polymorphism in the endangered Sorraia horse and related breeds. *J Anim Breed Genet* **122**, 69-72.

Luis C., Juras R., Oom M.M. & Cothran E.G. (2007b) Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. *Anim Genet* **38**, 20-7.

Mc Parland S., Kearney J.F., Rath M. & Berry D.P. (2007) Inbreeding Effects on Milk Production, Calving Performance, Fertility, and Conformation in Irish Holstein-Friesians. *J Dairy Sci* **90**, 4411-9.

McCue M. & Mickelson J. (2013) Genomic Tools and Resources: Development and Applications of an Equine SNP Genotyping Array. In: *Equine Genomics* (pp. 113-24. Blackwell Publishing Ltd.

McCue M.E., Bannasch D.L., Petersen J.L., Gurr J., Bailey E., Binns M.M., Distl O., Guerin G., Hasegawa T., Hill E.W., Leeb T., Lindgren G., Penedo M.C., Roed K.H., Ryder O.A., Swinburne J.E., Tozaki T., Valberg S.J., Vaudin M., Lindblad-Toh K., Wade C.M. & Mickelson J.R. (2012) A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. *PLoS Genet* **8**, e1002451.

McQueen C.M., Doan R., Dindot S.V., Bourquin J.R., Zlatev Z.Z., Chaffin M.K., Blodgett G.P., Ivanov I. & Cohen N.D. (2014) Identification of genomic loci associated with Rhodococcus equi susceptibility in foals. *PloS one* **9**, e98710.

Melo J.B., Pires D.A.F., Ribeiro M.N., Santos D.O. & Silva H.G.O. (2011) Estudo zoométrico de remanescentes da raça equina Nordestina no município de Floresta, Pernambuco – Brasil. *AICA* **1**, 71-4.

Metzger J., Philipp U., Lopes M.S., da Camara Machado A., Felicetti M., Silvestrelli M. & Distl O. (2013) Analysis of copy number variants by three detection algorithms and their association with body size in horses. *BMC Genomics* **14**, 487.

Miyamato T., Sato H., Yogev L., Kleiman S., Namiki M., Koh E., Sakugawa N., Hayashi H., Ishikawa M., Lamb D.J. & Sengoku K. (2006) Is a genetic defect in Fkbp6 a common cause of azoospermia in humans? *Cell Mol Biol Lett* **11**, 557-69.

Oliehoek P.A. & Bijma P. (2009) Effects of pedigree errors on the efficiency of conservation decisions. *Genet Sel Evol* **41**, 9.

Oom M. & Ferreira J. (1987) Estudo biométrico do cavalo Alter. *Rev Port Cienc Vet* **83**, 101-48.

Oom M.M. (1992) O cavalo Lusitano. Uma raça em Recuperação (in Portuguese). p. 299, University of Lisbon, Portugal.

Oom M.M. (2006) *O Cavalo do Sorraia. Um património histórico e genético a conservar (in Portuguese)*. Edições Cosmos, Portugal.

Oom M.M., d'Andrade J.L. & Costa-Ferreira J. (2004) *Studbook da raça Sorraia (in Portuguese)*. Garrido Artes Gráficas, Alpiarça, Portugal.

Orlando L., Ginolhac A., Zhang G., Froese D., Albrechtsen A., Stiller M., Schubert M., Cappellini E., Petersen B., Moltke I., Johnson P.L.F., Fumagalli M., Vilstrup J.T., Raghavan M., Korneliussen T., Malaspinas A.-S., Vogt J., Szklarczyk D., Kelstrup C.D., Vinther J., Dolocan A., Stenderup J., Velazquez A.M.V., Cahill J., Rasmussen M., Wang X., Min J., Zazula G.D., Seguin-Orlando A., Mortensen C., Magnussen K., Thompson J.F., Weinstock J., Gregersen K., Røed K.H., Eisenmann V., Rubin C.J., Miller D.C., Antczak D.F., Bertelsen M.F., Brunak S., Al-Rasheid K.A.S., Ryder O., Andersson L., Mundy J., Krogh A., Gilbert M.T.P., Kjær K., Sicheritz-Ponten T., Jensen L.J., Olsen J.V., Hofreiter M., Nielsen R., Shapiro B., Wang J. & Willerslev E. (2013) Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*.

Orr N., Back W., Gu J., Leegwater P., Govindarajan P., Conroy J., Ducro B., Van Arendonk J.A., MacHugh D.E., Ennis S., Hill E.W. & Brama P.A. (2010) Genome-wide SNP association-based localization of a dwarfism gene in Friesian dwarf horses. *Anim Genet* **41 Suppl 2**, 2-7.

Pariacote F., Van Vleck L.D. & MacNeil M.D. (1998) Effects of inbreeding and heterozygosity on preweaning traits in a closed population of Herefords under selection. *J Anim Sci* **76**, 1303-10.

Payne H.W., Ellsworth K. & DeGroot A. (1968) Aneuploidy in an infertile mare. *J Am Vet Med Assoc* **153**, 1293-9.

PDR2020 (2014) Ação 7.8 - Recursos Genéticos. In: *Programa de Desenvolvimento Rural do Continente para 2014-2020* (ed. by Governo de Portugal MdAedM). Governo de Portugal, Ministério da Agricultura e do Mar, Portugal.

Penedo M.C.T. & Raudsepp T. (2013) Molecular Genetic Testing and Karyotyping in the Horse. In: *Equine Genomics* (pp. 241-54. Blackwell Publishing Ltd.

Petersen J.L., Mickelson J.R., Cothran E.G., Andersson L.S., Axelsson J., Bailey E., Bannasch D., Binns M.M., Borges A.S., Brama P., da Camara Machado A., Distl O., Felicetti M., Fox-Clipsham L., Graves K.T., Guerin G., Haase B., Hasegawa T., Hemmann K., Hill E.W., Leeb T., Lindgren G., Lohi H., Lopes M.S., McGivney B.A., Mikko S., Orr N., Penedo M.C., Piercy R.J., Raekallio M., Rieder S., Roed K.H., Silvestrelli M., Swinburne J., Tozaki T., Vaudin M., C M.W. & McCue M.E. (2013) Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PloS one* **8**, e54997.

Putman A.I. & Carbone I. (2014) Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecol Evol* **4**, 4399-428.

Qu Z., Chowdhary B.P., Raudsepp T. & Adelson D.L. (2011) Design of an equine whole genome tiling array. *Plant & Animal Genome XIX, January 15–19, San Diego, CA*.

Ralls K., Brugger K. & Ballou J. (1979) Inbreeding and juvenile mortality in small populations of ungulates. *Science* **206**, 1101-3.

Raudsepp T. & Chowdhary B. (2016) Chromosome Aberrations and Fertility Disorders in Domestic Animals. In: *Annual Review of Animal Biosciences* (ed. by L A). Annual Reviews, Palo Alto, CA.

Raudsepp T. & Chowdhary B.P. (2013) Physical and Comparative Maps. In: *Equine Genomics* (pp. 49-71. Blackwell Publishing Ltd.

Raudsepp T., Das P.J. & Chowdhary B.P. (2013) Genomics of Reproduction and Fertility. In: *Equine Genomics*, pp. 199-215. Blackwell Publishing Ltd.

Raudsepp T., McCue M.E., Das P.J., Dobson L., Vishnoi M., Fritz K.L., Schaefer R., Rendahl A.K., Derr J.N., Love C.C., Varner D.D. & Chowdhary B.P. (2012) Genome-wide association study implicates testis-sperm specific FKBP6 as a susceptibility locus for impaired acrosome reaction in stallions. *PLoS Genet* **8**, e1003139.

Reece R.J. (2004) *Analysis of Genes and Genomes*. John Wiley & Sons, LTd., West Sussex, England.

Roos L., Hinrichs D., Nissen T. & Krieter J. (2015) Investigations into genetic variability in Holstein horse breed using pedigree data. *Livest Sci* **177**, 25-32.

Ryan K.K., Robert C.L. & Margulis S.W. (2002) Impacts of inbreeding on components of reproductive success. In: *Reproduction Science and Integrated Conservation* (eds. by Holt WV, Pickard AR, Rodger JC & Wildt JC), pp. 82-96. Cambridge University Press, Cambridge, United Kingdom.

Sairanen J., Nivola K., Katila T., Virtala A.M. & Ojala M. (2009) Effects of inbreeding and other genetic components on equine fertility. *Animal* **3**, 1662-72.

Santana M.L., Jr., Oliveira P.S., Eler J.P., Gutierrez J.P. & Ferraz J.B. (2012) Pedigree analysis and inbreeding depression on growth traits in Brazilian Marchigiana and Bonsmara breeds. *J Anim Sci* **90**, 99-108.

Santana M.L.J., Oliveira P.S., Pedrosa V.B., Eler J.P., Groeneveld E. & Ferraz J.B.S. (2010) Effect of inbreeding on growth and reproductive traits of Nellore cattle in Brazil. *Livest Sci* **131**, 212-7.

Santos A.S. & Ferreira L.M.M. (2012) The Portuguese Garrano breed: an efficient and sustainable production system. In: *Forages and grazing in horse nutrition* (eds. by Saastamoinen M, Fradinho M, Santos A & Miraglia N), pp. 481-4. Wageningen Academic Publishers.

Schaid D.J., Guenther J.C., Christensen G.B., Hebbring S., Rosenow C., Hilker C.A., McDonnell S.K., Cunningham J.M., Slager S.L., Blute M.L. & Thibodeau S.N. (2004) Comparison of Microsatellites Versus Single-Nucleotide Polymorphisms in a Genome Linkage Screen for Prostate Cancer–Susceptibility Loci. *American journal of human genetics* **75**, 948-65.

Schrimpf R., Dierks C., Martinsson G., Sieme H. & Distl O. (2014) Genome-wide association study identifies phospholipase C zeta 1 (PLCz1) as a stallion fertility locus in Hanoverian warmblood horses. *PloS one* **9**, e109675.

Schrimpf R., Metzger J., Martinsson G., Sieme H. & Distl O. (2015) Implication of FKBP6 for male fertility in horses. *Reprod Domest Anim* **50**, 195-9.

Siderits M., Baumung R. & Fuerst-Waltl B. (2013) Pedigree analysis in the German Paint Horse: Genetic variability and the influence of pedigree quality. *Livest Sci* **151**, 152-7.

Smith L.A., Cassell B.G. & Pearson R.E. (1998) The effects of inbreeding on the lifetime performance of dairy cattle. *J Dairy Sci* **81**, 2729-37.

Solé M., Santos R., Gómez M.D., Galisteo A.M. & Valera M. (2013) Evaluation of conformation against traits associated with dressage ability in unridden Iberian horses at the trot. *Res Vet Sci* **95**, 660-6.

Sundstrom E., Imsland F., Mikko S., Wade C., Sigurdsson S., Pielberg G.R., Golovko A., Curik I., Seltenhammer M.H., Solkner J., Lindblad-Toh K. & Andersson L. (2012) Copy number expansion of the STX17 duplication in melanoma tissue from Grey horses. *BMC Genomics* **13**, 365.

Swiger L.A., Gregory K.E., Koch R.M. & Arthaud V.A. (1961) Effect of Inbreeding on Performance Traits of Beef Cattle. *J Anim Sci* **20**, 626-30.

Swinburne J. & Lindgren G. (2013) Genetic Linkage Maps. In: *Equine Genomics* (pp. 11-47. Blackwell Publishing Ltd.

Takezaki N. & Nei M. (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**, 389-99.

Toro M.A., Fernández J. & Caballero A. (2009) Molecular characterization of breeds and its use in conservation. *Livest Sci* **120**, 174-95.

Valera M., Esteves L., Oom M.M. & Molina A. (2000) La raça equina autóctone Puro Sangue Lusitano: estudo genético dos parâmetros reproductivos de importância nos esquemas de conservação e melhoramento. *Archivos de zootecnia* **49**, 147-56.

Valera M., Molina A., Gutiérrez J.P., Gómez J. & Goyache F. (2005) Pedigree analysis in the Andalusian horse: population structure, genetic variability and influence of the Carthusian strain. *Livest Prod Sci* **95**, 57-66.

Vanderwall D.K. (2008) Early Embryonic Loss in the Mare. *Journal of Equine Veterinary Science* **28**, 691-702.

Vicente A.A., Carolino N. & Gama L.T. (2012) Genetic diversity in the Lusitano horse breed assessed by pedigree analysis. *Livest Sci* **148**, 16-25.

Vicente A.A., Carolino N., Ralão-Duarte J. & Gama L.T. (2014a) Selection for morphology, gaits and functional traits in Lusitano horses: I. Genetic parameter estimates. *Livest Sci* **164**, 1-12.

Vicente A.A., Carolino N., Ralão-Duarte J. & Gama L.T. (2014b) Selection for morphology, gaits and functional traits in Lusitano horses: II. Fixed effects, genetic trends and selection in retrospect. *Livest Sci* **164**, 13-25.

Vostrá-Vydrová H., Vostrý L., Hofmanová B., Krupa E. & Zavadilová L. (2016) Pedigree analysis of the endangered Old Kladruber horse population. *Livest Sci* **185**, 17-23.

Wang W., Wang S., Hou C., Xing Y., Cao J., Wu K., Liu C., Zhang D., Zhang L., Zhang Y. & Zhou H. (2014) Genome-Wide Detection of Copy Number Variations among Diverse Horse Breeds by Array CGH. *PloS one* **9**, e86860.

Warmuth V., Eriksson A., Bower M.A., Canon J., Cothran G., Distl O., Glowatzki-Mullis M.L., Hunt H., Luis C., do Mar Oom M., Yupanqui I.T., Zabek T. & Manica A. (2011) European domestic horses originated in two holocene refugia. *PloS one* **6**, e18194.

Warmuth V., Manica A., Eriksson A., Barker G. & Bower M. (2013) Autosomal genetic diversity in non-breed horses from eastern Eurasia provides insights into historical population movements. *Anim Genet* **44**, 53-61.

Wyrobek A.J., Alhborn T., Balhorn R., Stanker L. & Pinkel D. (1990) Fluorescence in situ hybridization to Y chromosomes in decondensed human sperm nuclei. *Mol Reprod Dev* **27**, 200-8.

Zechner P., Sölkner J., Bodo I., Druml T., Baumung R., Achmann R., Marti E., Habe F. & Brem G. (2002) Analysis of diversity and population structure in the Lipizzan horse breed based on pedigree information. *Livest Prod Sci* **77**, 137-46.

CHAPTER 2 - Breed Characterization

Genetic diversity, demographic structure, morphometry and inbreeding

Paper 1 - Pinheiro M., Kjöllerström H.J. & Oom M.M. (2013) Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis. Livestock Science 152, 1-10.

Paper 2 - Kjöllerström H.J. & Oom M.M. Morphological characterization and inbreeding effects in the endangered Sorraia Horse breed. (*submitted to the Journal of Animal Science*)

Paper 3 - Kjöllerström, H.J., Raudsepp, T., Khanshour, A., Gosh, S., Oom, M.M. & Chowdhary, B.P. Genome-wide variability in the Sorraia horse. (*in preparation for PLoS ONE*)

CHAPTER 2 - PAPER 1

Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis.

Pinheiro M., Kjöllerström H.J. & Oom M.M.

Livestock Science (2013) 152, 1-10

Livestock Science 152 (2013) 1–10

Contents lists available at SciVerse ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis

M. Pinheiro, H.J. Kjöllerström, M.M. Oom*

Universidade de Lisboa, Faculdade de Ciências, Centro de Biologia Ambiental/Departamento de Biologia Animal, Edifício C2 Piso 3, 1749-016 Lisboa, Portugal

ARTICLE INFO

Article history:
Received 21 March 2012
Received in revised form
21 November 2012
Accepted 21 November 2012

Keywords:
Genetic diversity
Inbreeding
Pedigree information
Population structure
Sorraia horse

ABSTRACT

The complete pedigree information included in the Sorraia horse studbook was analyzed to investigate on the breed's genetic variability and demographic structure and ascertain the status of the current population. The dataset included 653 animals registered since the breed's foundation, in 1937, until 2006. In some cases, only a reference population consisting of 206 living animals was considered for the calculations. The inbreeding coefficients (F) and the genetic contribution of each founder to the population's genetic pool were assessed and parameters such as average relatedness (AR) and genetic importance of the studs were, for the first time, computed for this breed. The average generation interval found for the whole pedigree was 7.94 yr. Consistent with previous results, low levels of genetic diversity were found for this breed. The unbalanced use of animals for reproduction during most part of the breed's history are reflected by the low values of effective number of founders (7.46) and effective number of ancestors (4) characterizing the whole population, with only 15 animals explaining its overall genetic variability. Extremely high values were reported for the F and AR coefficients and mean levels of 26.99% and 46.26% were observed. The rate of inbreeding per generation (5.2%) exceeds the 1% (0.01) maximum limit defined by FAO. Taking into account only the living population, roughly 97% of the animals showed $F \geq 25\%$ (some of them exhibiting values above 60%), and there were no cases with AR values lower than 50%. The average F, AR and mean kinship coefficients computed for this group were 36.90%, 55.11% and 0.34, respectively. Indicating high losses of genetic diversity through generations, great differences were found for the genetic contribution of each founder to the current genetic pool and some of these animals are no longer represented. In the more recent years, an improvement of the population's parameters can, although, be observed, especially as a consequence of the conservation strategies that are being carried out by the Breeders Association. The number of births registered in the studbook strongly increased in the last two decades and the average values of F and AR tended to stabilize and even decrease in the recent past. The results make clear that an efficient conservation plan is essential to insure the maintenance of this singular breed. Appropriate management measures must continue to be implemented and improved as a way to preserve the genetic diversity still present in the population.

1. Introduction

The Sorraia horse is, along with the Lusitano and the Garrano, one of the native horse breeds in Portugal. Besides its unusual ability to survive in harsh conditions, the breed presents primitive traits and markings such as

* Corresponding author. Tel.: +351 217500000; fax: +351 217500028.
E-mail address: mmoom@fc.ul.pt (M.M. Oom).

1871-1413/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.livsci.2012.11.017

dun color (yellow or mouse), black dorsal stripe (frequently axial as well), zebra marks (leg barring), face masks, ear tips and bars and lighter guard hairs along the edges of a dark mane and tail (frosting) (Andrade, 1945; Oom et al., 2004). According to historical and molecular genetic evidence, it is believed to be the most probable ancestor of Southern Iberian horse breeds and, thus, involved in the foundation of light saddle horse breeds around the world (Andrade, 1945; Luís et al., 2006). It is considered by FAO as a critical-maintained breed, with less than 100 breeding females (FAO, 2000). The current population consists of roughly 280 individuals worldwide distributed in several subpopulations, mainly in Portugal and in Germany. As a consequence of its reduced effective population size, small number of founders, complete genetic isolation and unbalanced use of reproductive animals, the breed exhibits low levels of genetic diversity and extremely high inbreeding coefficients (e.g. Luís et al., 2007). The success of any conservation strategy strongly depends on the regular evaluation of parameters characterizing the target population (FAO, 1998, 2007). So the aim of this work was to characterize the Sorraia horse breed's current status by analyzing the pedigree data included in the studbook (Oom et al., 2004) and contribute for a better knowledge about its demographic structure and genetic variability. The gathered results will provide an important complement to previously published studies and allow to better access the critical level of endangerment of the breed, and will be very useful to evaluate success of conservation strategies carried out by the Breeders Association and to suggest new ones to be implemented in the future.

2. Material and methods

Animal Care and Use Committee approval was not obtained for this study because the data were collected from the pre-existing database maintained by the Sorraia Horse Breeders Association (Associação Internacional de Criadores do Cavalo Ibérico de Tipo Primitivo—Sorraia).

2.1. Population data

The complete pedigree information registered in the Sorraia horse studbook since the breed's foundation in 1937, to December 31st 2006, was collected from the Breeders Association and analyzed in the current study. More recent years were not included in order to have complete and consistent data from all breeding years considered, as some breeders lack to register animals at early ages. The dataset comprised a total of 653 animals, including 316 males, 335 females and two individuals with unregistered sex.

2.2. Pedigree analysis

Several demographic and genetic parameters were estimated using ENDOG v4.8 (Gutiérrez and Goyache, 2005), a software for populations' genealogic analysis that has been widely used during the last years to study different domestic breeds, many of them at risk (e.g. Cecchi et al., 2006; Cervantes et al., 2008; Gutiérrez et al., 2005; Jordana et al., 2010; Rizzi et al., 2011; Royo et al., 2007; Valera et al., 2005). Most of the following parameters were computed for the whole pedigree, however, in some cases, only a reference population comprising the 206 living animals was considered for analysis.

2.3. Demographic parameters

Generation intervals: average age of parents at the birth of the progeny that is kept for reproduction (James, 1977). It was computed for the four parent–offspring pathways (father–son, father–daughter, mother–son and mother–daughter) using the birth dates of the registered animals and respective parents.

Average age of parents at the birth of their offspring: it is the average age of parents when all their progeny is considered (kept for reproduction or not). It was computed for each one of the four parent–offspring pathways using the birth dates registered in the studbook.

Pedigree completeness level: the quality of the pedigree information registered in the Sorraia horse studbook was assessed through the percentage of known ancestors in each parental generation for the first five generations.

Maximum number of generations: number of generations separating the animal from the furthest ancestor present in its genealogy. Ancestors with unknown parents were considered as founders (generation 0).

Equivalent complete generations: computed for each individual in the pedigree following the methodology proposed by Maignel et al. (1996) as the sum over all known ancestors of the terms computed as the sum of $(1/2)^n$ (n is the number of generations separating the individual to each known ancestor).

Genetic importance of studs: assessed as the contribution of the studs with reproductive males to the population (Vassalo et al., 1986). Studs were classified as (1) *nucleus studs*, if breeders use exclusively their own stallions for reproduction, never purchase stallions but allow transfers of their stallions to other studs; (2) *multiplier studs*, if breeders use stallions from other studs and provide stallions to other ones; (3) *commercial studs*, when breeders use stallions from other studs but never provide stallions to other groups; (4) *isolated studs*, if breeders use only their own stallions for reproduction, never purchasing or selling stallions to other studs.

2.4. Genetic parameters

Effective number of founders (f_e): number of founders that would be necessary to produce the same amount of genetic diversity as in the population under study if they had identical contributions to the descendant population. It is computed as $f_e = 1/(\sum_{k-1}^{f} q_k{}^2)$ where q_k is the probability of gene origin of the k ancestor (computed as the average relatedness coefficient—AR) and f is the total number of founders. It is equivalent to the parameter f_e obtained following the methodology proposed by James (1972) or Lacy (1989) (*founder equivalents*) if the whole pedigree is included in the calculations.

Effective number of ancestors (f_a): defined as the minimum number of ancestors (founders or not) that are necessary to explain the population's overall genetic diversity (Boichard et al., 1997). It is computed as $f_a = 1/\left(\sum_{j-1}^{a} q_j{}^2\right)$, q_j being the marginal contribution of an ancestor j (genetic contribution of an ancestor that is not explained by any other previously chosen one). This parameter takes into account genetic diversity losses derived from the unbalanced use of reproductive animals through generations producing bottlenecks, being considered a useful parameter and an important complement to the information supplied by the f_e. Only the animals with both parents known are considered for the calculations.

Inbreeding coefficient (F): it is the probability that an individual has two genes identical by descent (Wright, 1931) and is computed according to the methodology proposed by Meuwissen and Luo (1992).

Monitoring of genetic diversity has classically been carried out by assessing the evolution of inbreeding in the population (Wright, 1922), often converted to effective population size (N_e), which is regarded as a good indicator of the risk of genetic erosion (FAO, 1998, establishing a maximum acceptable inbreeding rate of 0.01 (1%) per generation). Alderson (2009) reports measuring actual level of inbreeding in a population is more relevant for immediate conservation action than considering the rate of inbreeding. We used average F values' evolution per year of birth to evaluate the increase in inbreeding per generation, according to the formula of Gutiérrez et al. (2003) ($F_n - F_{n-1} = l \times b$, where l is the average generation interval and b the regression coefficient of mean inbreeding coefficient per birth year).

Individual increase in inbreeding ΔF_i (Gutiérrez et al., 2008) was computed for each individual in the pedigree, following the modification proposed by Gutiérrez, et al. (2009) to account for the exclusion of self-fertilization, as $\Delta F_i = 1 - \sqrt[t-1]{1-F_i}$, where F_i is the inbreeding coefficient for each individual i and t the equivalent complete generations computed on the pedigree of this individual.

Average relatedness coefficient (*AR*) *of each individual*: defined as the probability that an allele randomly chosen from the whole population in the pedigree belongs to a given animal (Gutiérrez and Goyache, 2005). Numerically, AR corresponds to twice the probability of two random alleles (one from the animal and other from the complete population, including itself) that are identical by descent and it can be interpreted as the representation of a given animal in the whole pedigree regardless of the knowledge of its own pedigree (Dunner et al., 1998). It is computed as the average of the coefficients integrating the row from the individual in the numerator relationship matrix (Goyache et al., 2003; Gutiérrez et al., 2003) and it takes into account, simultaneously, the inbreeding and the coancestry coefficients (Gutiérrez and Goyache, 2005).

As a complement to the previous analyses, some additional parameters, broadly used in the management of captive populations, were computed using PM2000 software v1.213 (Pollak et al., 2005). These parameters were assessed only for the reference population containing the living animals. Besides the genetic representation of each founder to the existing genetic pool, PM2000, provided individual values for mean kinship coefficient (mK), which is defined as the average of the kinship values between a single individual and all the other individuals in the living population, including itself (Ballou and Lacy, 1995), and needs to be recalculated at each change in population size (births/deaths).

3. Results

3.1. Demographic analysis

Among the 653 animals included in the analyzed file, 640 (328 females, 310 males and two animals with unknown sex) are descendants from known parents and the remaining 13 (seven females and six males) were considered founders. Fig. 1 shows the evolution of the number of births registered in the Sorraia horse studbook (founders not included) from the breed's foundation till the end of 2006. Despite some visible fluctuations, the number of records has progressively increased through the years. This fact is particularly evident for the last two decades (1987–2006), where 50.6% of the registered animals are included, and is greatly justified by the appearance of new breeders and by the increasing interest developed in this breed.

Considering all the registered animals, only 36.1% produced offspring. This reproductive fitness unbalance is even more obvious when we consider each sex separately: while roughly half the breeding females (52.2%) became dams, only 61 of the 316 stallions (19.3%) became sires (data not shown).

Great differences can also be found when we analyze the progeny sizes (Fig. 2). Most of the dams (58.3%) produced one to three offspring, with a maximum of 11 descendants in three cases. Looking at the sires' distribution, it is easy to see that the difference is much higher. Although the majority (65.6%) has between one to eight foals, in two cases this value is higher than 50. As a consequence, the average number of foals per parent is quite different for the two sexes: 3.7 for the dams and 10.5 for the sires.

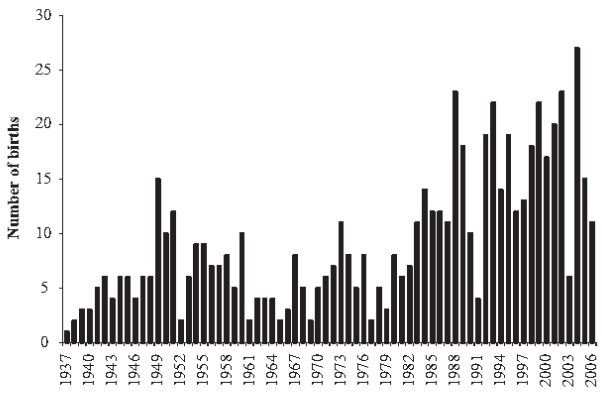


Fig. 1. Number of births registered each year in the Sorraia horse studbook (1937–2006).

For the Sorraia breed, we found an average generation length of 7.94 yr (8.39 yr if all the progeny is considered and not only the animals that would become sires and dams). Considering each parent–offspring pathway, dams' values are always slightly higher than those found for the sires (Table 1).

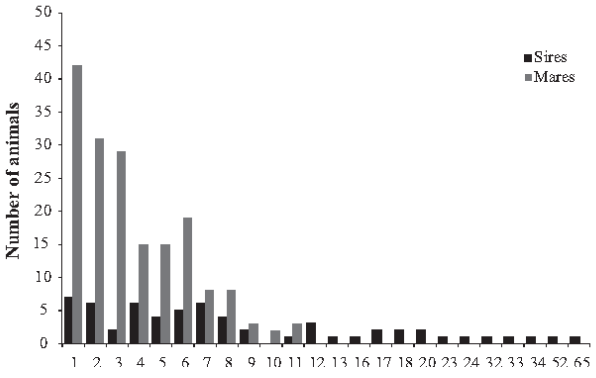


Fig. 2. Distribution of the sires and mares registered in the Sorraia horse studbook according to their progeny size.

Table 1
Generation intervals and average age of parents at the birth of their offspring (computed for each parent–offspring pathway) considering the complete Sorraia horse population.

Pathways	Generation interval		Average age	
	N	Years	*N*	Years
Sire–son	55	7.65	310	8.27
Sire–daughter	168	7.28	330	7.73
Dam–son	55	8.61	310	9.02
Dam–daughter	168	8.48	330	8.57
Average		7.94		8.39

The percentage of known ancestors is around 98% for the ancestors constituting the first generation, as shown in Fig. 3, remaining above 80% for all the parental positions in the second and third generations. With respect to the reference population, the pedigree knowledge is complete for all the positions in the genealogy till the fourth generation. The mean number of equivalent complete generations found for the breed was 6.14. Taking into account only the living population, we found an average value of 8.17 generations.

The classification of the Sorraia horse studs according to the origin and use of their reproductive males (Vassalo et al., 1986) is given in Table 2. The breed's first subpopulation was founded in 1975, when four foals (two males and two females) were transferred to the National Stud. Since then, and particularly in the most recent years, the number of subpopulations has increased, not only in Portugal but also in several other countries. From the ten studs registered in the studbook, only six were classified according to this methodology because no offspring were produced by the four remaining ones.

No *nucleus* studs, using exclusively their own males for reproduction and selling some of them to other groups, were found for this breed. Instead, they were all classified as *multiplier* or *commercial*, always making use of purchased males for reproduction, although to a variable extent. The differences among them are based on the use of their own stallions besides the purchased ones and in the possibility of transferring stallions to other groups. Table 3 depicts the values characterizing each one of these studs. As shown, the most influent ones (where the majority of the births occurred) were all classified as *multiplier* despite the great variation in values found between them.

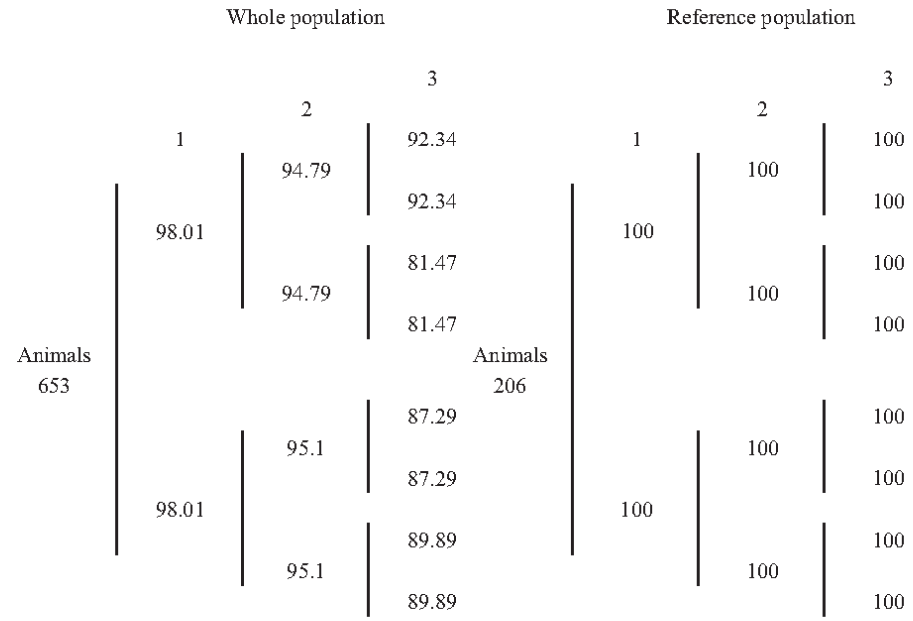


Fig. 3. Percentage of known ancestors per genealogical position in the first 3 parental generations. The values are given taking into account the whole population (left side) or the living animals only reference population (right side).

Table 2
Classification of the Sorraia horse studs according to the origin and use of their stallions (Vassalo et al. 1986). The column on the right shows the percentage of the reproductive males that were purchased from other studs.

Classification	Use purchased stallions	Use own stallions	Sell stallions	Number of studs	% Purchased males
Nucleus	No	Yes	Yes	0	0
Multiplier	Yes	Yes	Yes	3	16.39
Multiplier	Yes	No	Yes	0	100
Commercial	Yes	Yes	No	2	82.76
Commercial	Yes	No	No	1	100
Isolated	No	Yes	No	0	0

Table 3
Characterization of the studs registered in the Sorraia horse studbook: country of origin, classification according to the origin and use of the stallions (Vassalo et al., 1986), number of registered births, proportion of registered births whose father was born on the same stud, number of descendants produced by the stallions from the stud and percentage of these descendants that were born on the same stud.

Stud	Country	Classification	Births	% Father in	Total fathers	% Same stud
Andrade	Portugal	Multiplier	447	97.09	556	78.06
Schäfer	Germany	Multiplier	67	38.81	38	68.42
National Stud	Portugal	Multiplier	84	47.62	41	97.56
MA/FA	Portugal	Commercial	17	23.53	4	100
W. Springe	Germany	Commercial	17	0	0	0
Wilpferd	Germany	Commercial	12	8.33	1	100

3.2. Genetic analysis

Parameters characterizing the genetic variability of the Sorraia horse breed are given in Table 4, evidencing important diversity loss during the history of the breed. The effective number of founders (f_e) computed for the whole population was 7.46, which is nearly half the number of founders (13). The difference between the two values indicates that some of the initial genetic variability was lost due to the unbalanced contribution of these animals to the descendant population. Fig. 4 illustrates the genetic contribution of each Sorraia horse breed's founder to the whole population. Justifying the low f_e that was obtained for the whole population, great divergences are found among the contributions, with the three most represented founders (*Gaivota*, *Baio* and *Cunhal*) being responsible for more than half of the genetic variability (52.19%). The situation is even more dramatic when we consider just the living animals. Under ideal conditions, to avoid alleles' losses through generations, the contribution of each founder to the descendant population should be equivalent (Lacy, 1989). The historically reduced number of stallions used each breeding year led to the complete loss of diversity related to some of the initial founders and only ten are still represented in the current population (*Vigilante*, *Anselma* and *Freire* are no longer represented). Great discrepancies are also found among their individual contributions to the existing genetic pool, with the most important founder contributing with

Table 4
Main genealogical parameters illustrating the genetic variability of the Sorraia horse complete population.

Total number of animals	653
Number of founders	13
Effective number of founders (f_e)	7.46
Effective number of ancestors (f_a)	4
Number of ancestors explaining 100%	15
Number of ancestors explaining 50%	2
Mean inbreeding F (%)	26.99
Rate of inbreeding per generation $\overline{\Delta Fi}$ (%)	5.2
Mean average relatedness AR (%)	46.26

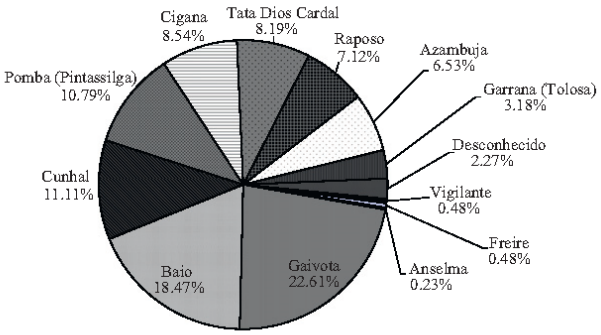


Fig. 4. Genetic contribution of each founder to the whole population registered in the Sorraia horse studbook.

25.00% to the overall diversity and the two less represented being responsible for only 1.96% and 2.75% (data not shown).

According to Boichard et al. (1997), the effective number of ancestors (f_a) computed for the whole population (animals with both parents known) was four. Following this methodology, only 15 animals (five males and ten females) are enough to explain the Sorraia horse's entire variability, some of them belonging to the founder population (Table 5). As it is shown, all these individuals were born before 1950 at the breed's main stud, Andrade, and great differences can be found among their contributions. A stallion born in 1948, *Manco*, is, unquestionably, the main ancestor of the breed, explaining 43.48% of the overall diversity. The second most important ancestor is a mare, *Borboleta* 17, but her contribution to the population (16.91%) is considerably lower than the previous one.

Table 5
Description of the ancestors explaining the overall genetic variability of the Sorraia horse breed. identified according to the methodology proposed by Boichard et al. (1997). Founder animals are shown in bold.

Name	Sex	Year	Stud	Contribution (%)
1. Manco	M	1948	Andrade	43.48
2. Borboleta 17	F	1946	Andrade	16.91
3. Engeitada II 29	F	1949	Andrade	13.51
4. **Baio**	M	1936	Andrade	7.82
5. Vassoura	F	1937	Andrade	4.31
6. **Pomba**	F	1933	Andrade	3.31
7. Esponja	F	1938	Andrade	2.29
8. Engeitada 2	F	1939	Andrade	1.95
9. Garrano	M	1938	Andrade	1.94
10. **Tata Dios Cardal**	M	1942	Andrade	1.45
11. **Gaivota**	F	1936	Andrade	1.17
12. Caraça	F	1939	Andrade	0.66
13. **Raposo**	M	1934	Andrade	0.65
14. **Cigana**	F	1933	Andrade	0.47
15. **Anselma**	F	1935	Andrade	0.08

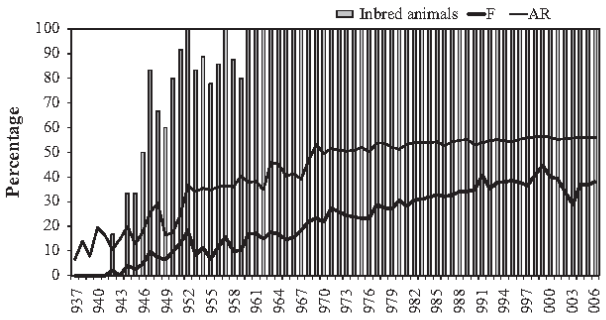


Fig. 5. Evolution of the average values of inbreeding (F) and relatedness (AR) coefficients per year of birth and percentage of inbred births registered each year in the Sorraia horse studbook.

Baio is the first founder arising in the list with the fourth major contribution (7.82%). According to the results, the two major contributors represent 60.39% of the breed's diversity and the first four are responsible for more than 80%.

Extremely high and unusual values of inbreeding were reported for the Sorraia horse breed and an average inbreeding coefficient (F) of 26.99% was found for the analyzed pedigree (27.54% if we exclude the founders). During the first five years after the breed's foundation, no inbred births were registered in the studbook (the majority resulted from the mating between founders) and the average F remained null (Fig. 5). However, due to inadequate breeding strategy, inbred animals started to appear as soon as 1942 and the percentage of inbred births shows a strong increase through the years (linear increase in inbreeding coefficient, not shown, is $y=0.0062\times-11.984$, $R^2=0.9434$, $p<0.001$). From 1960 onward, all the registered births are inbred. An accentuated increase in the average values of F per year of birth is also evident and a maximum value of 44.63% was reached for the animals born in 1999. Increase in inbreeding per generation was 0.049 (l x b=7.94 y r x 0.0062). Following Gutiérrez et al.'s (2008, 2009) approach, the rate of inbreeding per equivalent generation, $\overline{\Delta Fi}$, was 0.052 (5.2%).

Considering the individual F values computed for the complete population, we found that, from the 640 animals with known parents, only a small fraction (8.1%) is not inbred. Among the inbred ones, more than half (65.8%) revealed inbreeding levels of 25% or higher and, in some cases, this value exceeded 60%. Average values computed for each sex separately were found to be very similar: 28.53% for the males and 26.57% for the females. Particularly high values were found for the extant Sorraia horse population, with only six animals showing F values lower than 25% and an average F of 36.90% being recorded for this group (38.13% for the males and 35.90% for the females) (data not shown). The maximum F computed was 60.05% which is a very rare percentage given the adverse consequences usually arising from this kind of values.

As expected, high values were also reported for the average relatedness coefficients (AR) and an average of 46.26% was found for the whole pedigree. This parameter can be used as a tool to maintain the maximum genetic variability through generations if we choose as breeding animals those with lower levels of AR (Gutiérrez and Goyache, 2005). The average values of AR per year of birth quickly increased along the history of the breed and the maximum value of 56.54% was found for the animals born in 1999 (Fig. 5). However, from 1970 onward, these values tend to stabilize. Taking into account the individual values, most of the animals registered in the Sorraia horse studbook are characterized by AR levels between 50% and 60%, with an average of 55.11%, and a maximum of 59.72% was found for a stallion that sired 34 foals. The mean kinship coefficients (mK) computed for these animals were also very high, with an average of 0.34 being reported and the majority ranging from 0.315 to 0.360 (Fig. 6).

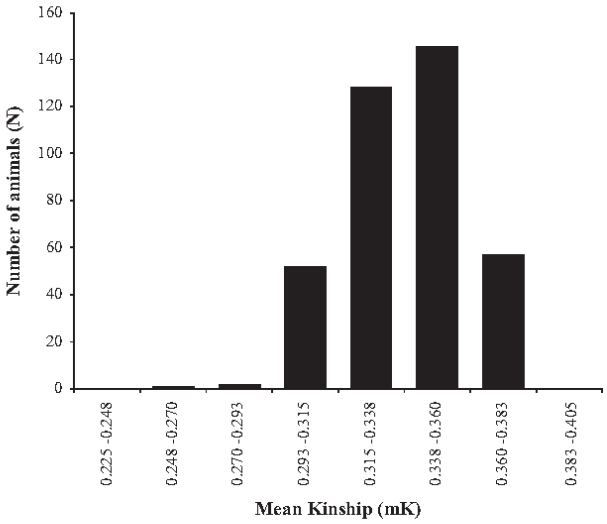


Fig. 6. Distribution of the animals constituting the living population (reference population) per class of mean kinship coefficient (mK).

4. Discussion

4.1. Demographic analysis

Despite some visible fluctuations of the evolution of births per year (Fig. 1), the number of records progressively increased through the years. This fact is particularly evident for the last two decades (1987–2006), where 50.6% of the registered animals are included, and is greatly justified by the appearance of new breeders and by the increasing interest developed in this breed.

The unbalanced use of animals for reproduction, which is one of the main causes for the loss of genetic diversity through generations, has been commonly practiced within many Sorraia horse's studs. Considering all registered animals, half (52.2%) the breeding females and 19.3% of the stallions produced offspring. Lower values were registered by Cervantes et al. (2008) for the Spanish Arab horse, with only 27.9% of the females and 12.6% of the males being used for reproduction. Similar values for dams and sires producing one offspring were reported by Schurink et al. (2012) in the Dutch harness horse (52.2% and 56.5%, respectively).

The divergence between the average number of foals produced per sire (10.5) and dam (3.7) reported in the present study is in accordance with the one described for several other horse breeds such as Mangalarga (23.8 for sires; 4.4 for dams) (Mota et al., 2006) or Campolina (22.2 for sires; 3.1 for dams) (Procópio et al., 2003). Furthermore, the average of 10.5 descendants per sire encountered for the Sorraia horse is smaller than the 12.3 found for the Lusitano horse (Valera et al., 2000).

Generation length represents one of the main factors influencing the rate of genetic progress and genetic conservation strategies of a specific population, being an important parameter to be measured in the context of any breeding program. According to previous literature, generation length is generally long for horses when compared with other domestic breeds, frequently ranging between 9 and 11 yr. The value obtained for the Sorraia breed (7.94 yr) is lower than that registered for the Andalusian horse (10.11 yr, Valera et al., 2005), the Lusitano (10.40 yr, Vicente et al., 2009), the Hanoverian warmblood (10 yr, Hamann and Distl, 2008) or the Dutch harness (8.60 yr, Schurink et al., 2012).

Considering each parent–offspring pathway (Table 1), the fact that dams' values are always slightly higher than those of sires is mainly justified by the fact that dams are kept for reproduction as long as possible (during most of their lifetime) whereas sires are usually replaced in breeding plans more frequently (only one sire is used per herd per breeding year). Effectively, one of the most important goals of the breed's conservation plan is to encourage the breeders to perform a regular substitution of their stallions in order to minimize genetic diversity loss through generations. The opposite scenario was reported by Moureaux et al. (1996) from the analysis of pedigree information concerning five French horse breeds—Thoroughbred, Trotteur Français, Arab, Anglo-Arab and Selle Français. The values described in the literature for the four parent–offspring pathways show that the sire–offspring ones are, for most of the cases, larger than those comprising dams, ranging between 8.2 yr (dam–daughter pathway in the Arab horse) and 13.1 yr (sire–son pathway in the Trotteur Français). The same scenario was found for the Lusitano horse, with values of 11.2 yr for sires and 9.6 yr for dams (Vicente et al., 2009). Schurink et al. (2012) found almost the same values for paternal (8.7 yr) and maternal lineages (8.5 yr) in the Dutch harness horse.

Given the particular relevance of the recovery of the Sorraia horse and its maintenance over the years, the quality of the genealogical information registered in the studbook is very high. Validated, complete pedigree information is available for all extant animals. Therefore, high pedigree completeness levels were expected in this population. In fact, 98% of the ancestors constituting the first generation are known, and over 80% for all the parental positions in the second and the third generations. This reduction is entirely justified by the fact that, mainly for the eldest animals, founders can be found in early genealogical positions and, so, there is no knowledge about former ancestors. The mean number of equivalent complete generations found for the breed was 6.14, which is very similar to the 5.70 generations reported by Cervantes et al. (2008) in the Spanish Arab horse but lower than the 8.26 generations found by Valera et al. (2005) in the Andalusian, the 12.28 generations in the Austrian Noriker draught (Druml et al., 2009) or the 15.22 in the Lipizzan (Zechner et al., 2002). However, in all cases, the maximum number of generations that can be traced is longer than the 13 generations found for the Sorraia horse breed, with records starting only in 1937. If considering only the living Sorraia horse population, the average value is 8.17 generations, which is considerably higher.

In recent years, the number of subpopulations has increased, not only in Portugal but also in several other countries, as Germany and Brazil. Currently, although most of the living animals are bred at Portuguese studs, a significant and growing number of animals can be found in Germany. One of its studs, Schäfer, is even one of the most important in the history of the breed and was the basis for the formation of several new subpopulations in the country.

From the ten studs registered in the studbook, only six were classified according to the origin and use of their reproductive males because no offspring were produced by the four remaining ones. No *nucleus* studs were found for this breed. Instead, they were all classified as *multiplier* or *commercial*. However, considering the results obtained for the Andrade's stud (the founder one and the most important in the history of the breed), we can say that although it was classified as *multiplier*, actually it behaves as *nucleus* because it has only used their own stallions for reproduction so far. The *multiplier* classification is due, exclusively, to the group of 13 founder animals, as they were considered to be descendants of males belonging to external groups.

4.2. Genetic analysis

The maintenance of a domestic breed is a process dependent not only on a minimum number of individuals

that can assure its survival but also on the underlying genetic diversity which represents the breed's ability to adapt in response to environmental changes being, thus, the basis of its evolutionary potential (Frankham et al., 2004). The assessment of this diversity within populations, as well as their structure, gene flow and demography, is essential to understand how endangered the population is and is important for the establishment of an effective selection and/or conservation program (Cervantes et al., 2008; Fernández et al., 2004; Goyache et al., 2003; Gutiérrez et al., 2003).

As evidenced in Table 4, there have been important losses of diversity during the history of the breed. The effective number of founders (f_e) is 7.46, which is nearly half the number of founders. The difference between the two values indicates that some of the initial genetic variability was lost due to the unbalanced contribution of founders to the descendant population. However, this ratio is higher than those obtained for other horse breeds, no matter the length and the completeness of the pedigree since the breed's foundation: with an initial number of founders between 1000 and 2000, Valera et al. (2005) reported a f_e of only 39.6 for the Andalusian horse, Cervantes et al. (2008) a value of 38.6 for the Spanish Arab horse and Druml et al. (2009) a value of 117.2 for the Austrian Noriker draught horse. The same situation was reported for the Lipizzan horse, with a f_e=42 computed from a pedigree with animals with more than 30 known generations and traced from 457 founders (Zechner et al., 2002).

Under ideal conditions, the contribution of each founder to the descendant population should be equivalent (Lacy, 1989), but in the Sorraia horse population just three founders are responsible for more than half (52.19%) of the population's genetic variability. The breeding historical management adopted led to the loss and uneven contribution of the ten founders still represented in the living population, as frequently the same stallion were used for consecutive years in the same breeding farm. Furthermore, results gathered from molecular analysis show that only two of the seven initial matrilineal lines from the founder females subsisted through the breeding years (Luís et al., 2002), corroborating genealogical data. Studying a reference population of the Lipizzan horse breed, Zechner et al. (2002) demonstrated that only 19 of the 457 animals identified as founders are necessary to explain more than 50% of the total genetic pool (the major contribution being around 7%). In the Andalusian horse breed, the ten most influent founders contributed with more than 40% (the main founder responsible for 8.19%) which explains the low f_e that was also reported for this population (Valera et al., 2005).

The effective number of ancestors (f_a) computed for the whole population (animals with both parents known) was four, which is quite lower than the obtained f_e value. This can be explained by the fact that usually the latter parameter is overestimated and does not account for possible bottlenecks occurring in a population, although f_a considers the contribution of any ancestor not yet accounted for by other ancestors (Boichard et al., 1997). This f_a value is much lower than those reported for other horse breeds as the Lipizzan (f_a=26.2) (Zechner et al., 2002), Spanish Arab horse (f_a=19.0) (Cervantes et al., 2008) or Andalusian (f_a=16.5) (Valera et al., 2005). In all these cases, a limited number of ancestors (less than ten) are enough to explain 50% of the breed's diversity, with the most influent animals being responsible for 10.74%, 13.80% and 15.77%.

The average inbreeding coefficient (F) of the analyzed pedigree was 26.99%. These results are justified, not only by the small number of founders, but also by the reduced effective population size, the complete genetic isolation over the years (resulting from the non-existence of further introductions) and, as previously presented, by the unbalanced use of animals for reproduction, especially among males. In the first five years no inbred births were registered in the studbook, and inbred animals started to appear as soon as 1942 and the percentage of inbred births shows a strong increase through the years. Nevertheless, as it is illustrated in Fig. 5, the average values have stabilized or even decreased in recent years (especially over the last decade) which is a consequence of the efforts made by the Breeders Association in defining breeding plans to minimize this parameter. Increase in inbreeding per generation was 4.90%, much higher than the 1.36% reported by Schurink et al. (2012) for the Dutch harness horse and far from the limit of 1% defined by FAO (1998). Using Gutiérrez et al.'s (2008, 2009) approach, a more reliable method that accounts for differences in pedigree knowledge and completeness and for the effects of frequent genetic events in small populations (such as drift, non-discrete generations, breeding plans, selection and unequal contributions from a different number of ancestors, all conditioning the pedigree of each individual), the obtained value for the rate of inbreeding per generation is quite similar (5.2%). This high value reinforces the importance of a breeding plan focusing on maximum avoiding inbreeding.

Only 8.1% of the animals registered in the Studbook are not inbred and among inbred ones, more than half have inbreeding levels of 25% or higher. The inbreeding levels reported for the Sorraia horse breed are much higher than those described by other authors for different domestic breeds. The average value of 26.99% in our study is very distant from the 1.33%, 3.42%, 5.30%, 7.00%, 8.48%, 9.03% or 10.81% reported for the analyzed populations of Hanoverian warmblood (Hamann and Distl, 2008), Losina (Valera et al., 2007), Dutch harness horse (Schurink et al., 2012), Spanish Arab (Cervantes et al., 2008), Andalusian (Valera et al., 2005), Lusitano (Costa-Ferreira and Oom, 1989) and Lipizzan (Zechner et al., 2002) horse breeds, respectively. However, notwithstanding the differences between the results, the F value is admittedly highly dependent on such features as the deepness and the quality of the available pedigree information and that comparisons between breeds must be made in a cautious way. As previously demonstrated, the pedigree knowledge in the Sorraia horse breed is complete till the founder animals, and, therefore, the computed values are not underestimated, assuming that founders were probably unrelated.

High values were also reported for the average relatedness coefficients (46.26%). Much lower values were found

in other studies regarding different domestic breeds. Valera et al. (2005) reported an average AR of 12.25% for the Andalusian horse and only 3.76% was found by Gutiérrez et al. (2005) in the Catalonian donkey. All living Sorraia horses exhibit AR values equal or higher than 50% with an average of 55.11%.

The average mean kinship coefficient (mK) was 0.34. Minimizing mean kinship coefficients is the most recommended method for the establishment of efficient genetic management strategies in captive populations (Frankham et al., 2004) and has been routinely used in the management of zoo populations. The implementation of reproduction plans based on individual mK values has been, as much as possible, the main strategy adopted by the Sorraia horse Breeders Association in the more recent years (as far as we know, for the first time in a horse breed population). However, due to the geographical constraints and the fact that the most important animals are post-reproductive, the possible management is to maximally avoid inbreeding in each generation by choosing, every breeding year, the optimal and available stallion to each stud farm. Awareness of breeders for these conservation strategies is increasing, as some information forums are promoted by the Breeders Association, and they routinely ask for genetic analysis and simulations in order to select the most genetically important stallions. Positive results have been achieved following this methodology, and we have recent registries of animals with inbreeding coefficient far below the average.

5. Conclusion

Conservation of genetic diversity is now universally accepted as being vital for sustainable management of animal genetic resources and it can be accomplished by selection programmes that will restore genetic diversity in industrial breeds (Ajmone-Marsan and The GLOBALDIV Consortium, 2010).

Used independently or as a complement to molecular approaches, pedigree analysis is considered a useful tool to describe genetic variability within populations and its evolution across generations (Boichard et al., 1997). The results gathered in the present study (some of them for the first time computed in the breed) reveal, once more, low levels of diversity in the Sorraia horse population and are consistent with the results obtained from molecular analyses. The inadequate breeding strategies practiced during most part of the breed's history, in particular the unbalanced use of stallions and the lack of veterinarian support on evaluating and improving reproductive performance, both of mares and stallions, justify the low values of f_e and f_a found for the population, with a small number of animals explaining the overall genetic diversity. The average levels of F and AR registered in this study are extremely high and far from the values reported for several other livestock breeds, the situation is even more alarming when considering only the living population. Despite the slight improvements noticed in the more recent years (especially as a consequence of the management measures carried out by the Breeders Association, as aforementioned in Discussion), increase in inbreeding per generation far exceeds the 1% FAO limit and the breed is still at critical risk of extinction, so an efficient conservation plan is essential to guarantee a long-term self-sustaining population. In recent years, some decisive management-breeding plans have been pursued in order to reverse this scenario, namely by yearly selecting the most suitable stallion for each herd and to promote their interexchange between breeding farms, focusing on decreasing the extremely high inbreeding coefficients of the current population. Also, assays for improving fertility and survival rates were implemented, as far as possible, by monitoring ovulation and pregnancy by ultrasonography standard practices and evaluating semen characteristics in stallions selected for reproduction. The number of births per year is slightly increasing (the population size almost reached 300 animals in 2011), as well as the number of breeders, which definitively allows a more efficient genetic management in order to minimize inbreeding and maximizing genetic diversity.

In this context, all efforts that are currently undergoing such as maximizing the use of all available sires and mares, improving fertility and progeny survival, defining breeding strategies to minimize inbreeding and maximize genetic diversity is extremely important, as well as a permanent analysis of the achieved results.

Conflict of interest statement

The authors Márcia Pinheiro, Helena Josefina Kjöllerström and Maria do Mar Oom hereby declare that there was no conflict of interest involved in the present study, which could possibly influence their work.

Acknowledgments

The authors would like to thank the Associação Internacional de Criadores do Cavalo Ibérico de Tipo Primitivo—Sorraia for supplying the data analyzed in the present study. This work was supported by Centro de Biologia Ambiental (PEst-OE/BIA/UI0329/2011) and by PRODER/Action 2.2.3 (2011–2013). "Conservation and improvement of genetic resources"/Sub-action 2.2.3.2 "Animal component" (Sorraia breed). H.J. Kjöllerström was supported by a Ph.D. Grant (SFRH/BD/81502/2011) from the Portuguese Foundation for Science and Technology (FCT/MEC).

References

Alderson, J., 2009. Breeds at risk: definition and measurement of the factors which determine endangerment. Livest. Sci. 123, 23–27.

Andrade, R.D., 1945. O Cavalo do Sorraia. Boletim Pecuário. Ano XIII, vol. 3, pp. 1–13.

Ajmone-Marsan, P., The GLOBALDIV Consortium, 2010. A global view of livestock biodiversity and conservation GLOBALDIV. Anim. Genet. 41 (1), 1–5.

Ballou, J.D., Lacy, R.C., 1995. Identifying genetically important individuals for management of genetic variation in pedigreed populations. In: Ballou, J.D., Gilpin, M., Foose, T.J. (Eds.), Population Management for Survival and Recovery: Analytical Methods and Strategies in Small Population Conservation, Columbia University Press, New York, pp. 76–111.

Boichard, D., Maignel, L., Verrier, É., 1997. The value of using probabilities of gene origin to measure genetic variability in a population. Genet. Sel. Evol. 29, 5–23.
Cecchi, F., Ciampolini, R., Ciani, E., Matteoli, B., Mazzanti, E., Tancredi, M., Presciuttini, S., 2006. Demographic genetics of the endangered Amiata donkey breed. Ital. J. Anim. Sci. 5, 387–391.
Cervantes, I., Molina, A., Goyache, F., Gutiérrez, J.P., Valera, M., 2008. Population history and genetic variability in the Spanish Arab Horse assessed via pedigree analysis. Livest. Sci. 113, 24–33.
Costa-Ferreira, J., Oom, M.M., 1989. Inbreeding in the Lusitanian Horse. Rev. Port. Ciên. Vet. 84, 77–89.
Druml, T., Baumung, R., Sölkner, J., 2009. Pedigree analysis in the Austrian Noriker draught horse: genetic diversity and the impact of breeding for coat colour on population structure. J. Anim. Breed. Genet. 126, 348–356.
Dunner, S., Checa, M.L., Gutiérrez, J.P., Martin, J.P., Cañon, J., 1998. Genetic analysis and management in small populations: the Asturcon pony as an example. Genet. Sel. Evol. 30, 397–405.
FAO, 1998. Secondary Guidelines for Development of National Farm Animal Genetic Resources Management Plans: Management of Small Populations at Risk. Rome, Italy.
FAO, 2000. World Watch List for Domestic Animal Diversity, third ed. Scherf, B.D. Rome, Italy.
FAO, 2007. The State of the World's Animal Genetic Resources for Food and Agriculture. Rischkowsky, B., Pilling, D. Rome, Italy.
Fernández, J., Toro, M.A., Caballero, A., 2004. Managing individuals' contributions to maximize the allelic diversity maintained in small, conserved populations. Conserv. Biol. 18, 1358–1367.
Frankham, R., Ballou, J.D., Briscoe, D.A., 2004. A Primer of Conservation Genetics. Cambridge University Press, Cambridge, UK.
Goyache, F., Gutiérrez, J.P., Fernández, I., Gomez, E., Alvarez, I., Díez, J., Royo, L.J., 2003. Using pedigree information to monitor genetic variability of endangered populations: the Xalda sheep breed of Asturias as an example. J. Anim. Breed. Genet. 120, 95–105.
Gutiérrez, J.P., Altarriba, J., Díaz, C., Quintanilla, R., Cañon, J., Piedrafita, J., 2003. Pedigree analysis of eight Spanish beef cattle breeds. Genet. Sel. Evol. 35, 43–63.
Gutiérrez, J.P., Goyache, F., 2005. A note on ENDOG: a computer program for analysing pedigree information. J. Anim. Breed. Genet. 122, 172–176.
Gutiérrez, J.P., Marmi, J., Goyache, F., Jordana, J., 2005. Pedigree information reveals moderate to high levels of inbreeding and a weak population structure in the endangered Catalonian donkey breed. J. Anim. Breed. Genet. 122, 378–386.
Gutiérrez, J.P., Cervantes, I., Molina, A., Valera, M., Goyache, F., 2008. Individual increase in inbreeding allows estimating effective sizes from pedigrees. Genet. Sel. Evol. 40, 359–378.
Gutiérrez, J.P., Cervantes, I., Goyache, F., 2009. Improving the estimation of realised effective population sizes in farm animals. J. Anim. Breed. Genet. 126, 327–332.
Hamann, H., Distl, O., 2008. Genetic variability in Hanoverian warmblood horses using pedigree analysis. J. Anim. Sci. 86, 1503–1513.
James, J.W., 1972. Computation of genetic contributions from pedigrees. Theor. Appl. Genet. 42, 272–273.
James, J.W., 1977. A note on selection differentials and generation length when generations overlap. Anim. Prod. 24, 109–112.
Jordana, J., Ferrando, A., Marmi, J., Avellanet, R., Aranguren-Méndez, J.A., Goyache, F., 2010. Molecular, genealogical and morphometric characterisation of the *Pallaresa*, a Pyrenean relic cattle breed: Insights for conservation. Livest. Sci. 132, 65–72.
Lacy, R.C., 1989. Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. Zoo Biol. 8, 111–123.
Luís, C., Bastos-Silveira, C., Cothran, E.G., Oom, M.M., 2002. Variation in the mitochondrial control region sequence between the two maternal lines of the Sorraia horse breed. Genet. Mol. Biol. 25, 309–311.
Luís, C., Bastos-Silveira, C., Cothran, E.G., Oom, M.M., 2006. Iberian origins of new world horse breeds. J. Hered. 97, 107–113.
Luís, C., Cothran, E.G., Oom, M.M., 2007. Inbreeding and genetic structure in the endangered Sorraia horse breed: implications for its conservation and management. J. Hered. 98, 232–237.
Maignel, L., Boichard, D., Verrier, E., 1996. Genetic variability of French dairy breeds estimated from pedigree information. Interbull Bull. 14, 49–54.
Meuwissen, T.I., Luo, Z., 1992. Computing inbreeding coefficients in large populations. Genet. Sel. Evol. 24, 305–313.
Mota, M.D.S., Prado, R.S.A., Sobreiro, J., 2006. Characterization of the Mangalarga horse population in Brazil. Arch. Zootec. 55, 31–37.
Moureaux, S., Verrier, É., Ricard, A., Mériaux, J.C., 1996. Genetic variability within French race and riding horse breeds from genealogical data and blood marker polymorphisms. Genet. Sel. Evol. 28, 83–102.
Oom, M.M., Andrade, J.L.d'.Costa-Ferreira, J., 2004. Stud Book da Raça Sorraia. Associação Internacional de Criadores do Cavalo Ibérico do Tipo Primitivo Sorraia. Garrido Artes Gráficas, Alpiarça, Portugal.
Pollak, J.P., Lacy, R.C., Ballou, J.D., 2005. Population Management 2000, Version 1.213. Chicago Zoological Society, Brookfield, IL.
Procópio, A.M., Bergmann, J.A.G., Costa, M.D., 2003. Formation and demographic structure of the Campolina horse breed. Arq. Bras. Med. Vet. Zootec. 55, 361–365.
Rizzi, R., Tullo, E., Cito, S.M., Caroli, A., Pieragostini, E., 2011. Monitoring of genetic diversity in the endangered Martina Franca donkey population. J. Anim. Sci. 89, 1304–1311.
Royo, L.J., Álvarez, I., Gutiérrez, J.P., Fernández, I., Goyache, F., 2007. Genetic variability in the endangered Asturcón pony assessed using genealogical and molecular information. Livest. Sci. 107, 162–169.
Schurink, A., Arts, D.J.G., Ducro, B.J., 2012. Genetic diversity in the Dutch harness horse population using pedigree analysis. Livest. Sci. 143, 270–277.
Valera, M., Esteves, L., Oom, M.M., Molina, A., 2000. The Lusitano Native Thoroughbred: a genetic study of the important reproductive parameters in plans for conservation and improvement. Arch. Zootec. 49, 147–156.
Valera, M., Molina, A., Gutiérrez, J.P., Gómez, J., Goyache, F., 2005. Pedigree analysis in the Andalusian horse: population structure, genetic variability and influence of the Carthusian strain. Livest. Prod. Sci. 95, 57–66.
Valera, M., Martínez-Sáiz, J., Azor, P.J., Goyache, F., Molina, A., 2007. El caballo Losino. Modelo de recuperación genética de una raza equina. In: García, M.H. (Ed.), La raza equina marismeña de Doñana. Estudios interdisciplinares, Imprenta Sand S.L., Camas, Sevilla, Spain, pp. 59–107.
Vassalo, J.M., Díaz, C., García-Medina, J.R., 1986. A note on the population structure of the Avileña breed of cattle in Spain. Livest. Prod. Sci. 15, 285–288.
Vicente, A., Carolino, N., Gama, L.T., 2009. Demographic parameters in Lusitano horse. Arch. Zootec. 58 (1), 501–504.
Wright, S., 1922. Coefficients of inbreeding and relationship. Am. Nat. 56, 330–338.
Wright, S., 1931. Evolution in Mendelian populations. Genetics 16, 97–159.
Zechner, P., Sölkner, J., Bodo, I., Druml, T., Baumung, R., Achmann, R., Marti, E., Habe, F., Brem, G., 2002. Analysis of diversity and population structure in the Lipizzan horse breed based on pedigree information. Livest. Prod. Sci. 77, 137–146.

CHAPTER 2 - PAPER 2

Morphological characterization and inbreeding effects in the endangered Sorraia Horse breed.

Kjöllerström H.J. & Oom, M.M.

(Submitted to the Journal of Animal Science)

Morphological characterization and inbreeding effects in the endangered Sorraia Horse breed

Kjöllerström, Helena Josefina[a], Oom, Maria do Mar[a*]

[a]*cE3c -Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal*

[*]*Corresponding author: mmoom@fc.ul.pt*

Abstract

The primitive Sorraia horse is one of the most endangered animal genetic resources in Portugal, with only ~150 breeding mares and a worldwide population of ~300 animals, and is known for its extremely high level of inbreeding. So far, inbreeding depression in the Sorraia was found to be statistically significant in mare fertility. Analysing 26 body measurements in two 20-years-apart groups of stallions (OLD, N=8 + NEW, N=32), we morphologically describe Sorraia stallions and search for traces of inbreeding depression on its conformation. We also evaluate the effect of coat colour and age on the body measurements analysed. On average, animals from the OLD group have higher morphometric characteristics. The sampled animals can be considered a very homogenous group regarding their conformation, with low variability in some of the body measurements analysed. We found that this breed became smaller, except for head size, although overall body conformation ratio has not changed. Animals from the NEW group have higher values for inbreeding coefficients and age. Average inbreeding coefficient of sampled animals is 37.29%, with a significant negative impact on thoracic width and chest circumference, and positive in hock height in the NEW group, and only a significant positive effect on elbow height, in the OLD group. Age revealed to be significantly correlated with a few measurements, mostly those without any bone support in their definition. Yellow dun horses are taller and longer than mouse dun ones, except for tail insertion height. Although dark mouse dun horses (least common) seem to be shorter than mouse duns ("grulla", most common) and yellow duns, most differences are not statistically significant. Although pernicious effects of inbreeding on the majority of traits analysed on Sorraia horse are, so far, not alarming, studying inbreeding depression in this breed is of paramount importance for its long-term conservation and sustainability. These results can be used as a complementary tool to the breeds' current management breeding program.

Keywords: Sorraia horse; Morphometry; Inbreeding depression; Conformation; Inbreeding; Morphology

Introduction

The Sorraia horse is one of the most endangered animal genetic resources in Portugal recognized to be rare/critically endangered by the Portuguese government and in critical-maintained risk status by FAO (FAO, 2015; PDR2020, 2014). This breed has been managed as a closed population since its foundation in 1937 (Andrade, 1945; Oom et al., 2004), with only 12 founders and currently about 150 breeding mares. Sorraias have several morphological primitive characteristics including dun coat colour, bi-coloured mane and tail, dorsal stripe and zebra-like leg stripes (Oom et al., 2004). Phenotypic characterization of animal genetic resources is extremely important in defining conservation strategies (FAO, 2012). Sorraia morphology was previously studied by Oom (1992). With an average height at withers of 148cm for stallions the Sorraia is shorter than the Lusitano, but taller than the Garrano and the Terceira Pony.

The breed is known for extremely high inbreeding coefficients (Kjöllerström, 2005; Luis et al., 2007a; Pinheiro et al., 2013), with an average of 38% in the current living population (Kjöllerström et al., 2015). Inbreeding can expose recessive deleterious alleles due to increased homozygosity and result in reduced fitness, also known as inbreeding depression (Charpentier et al., 2007; Leberg and Firmin, 2008). Inbreeding depression has been reported for morphological traits in horses as well, namely in the Italian Haflinger (Gandini et al., 1992), Spanish Purebred (Gómez et al., 2009b) and Lusitano (Oom, 1992; Vicente et al., 2014b) breeds. Sorraia mare fertility has been shown to be affected by inbreeding depression: a 1% increase in inbreeding reduced mare fertility by 0.8% (Kjöllerström et al., 2015). The objective of this study was to morphologically describe Sorraia horse stallions and to assess the effects of inbreeding depression on its conformation over the last 20 years. We also evaluated the effect of coat colour and age on the analysed body measurements.

Material and Methods

In this study, 26 body measurements (Figure 1) were taken in 40 stallions organized in two groups for further analyses: the NEW group, comprising 32 stallions measured between 2012 and 2013, and the OLD group, comprising 8 stallions measured in 1984 (Oom, 1992). All measures were taken from the left side of the animals while standing in a square position, over a flat, hard floor, following the procedure described in Oom (1992) (Figure 2).

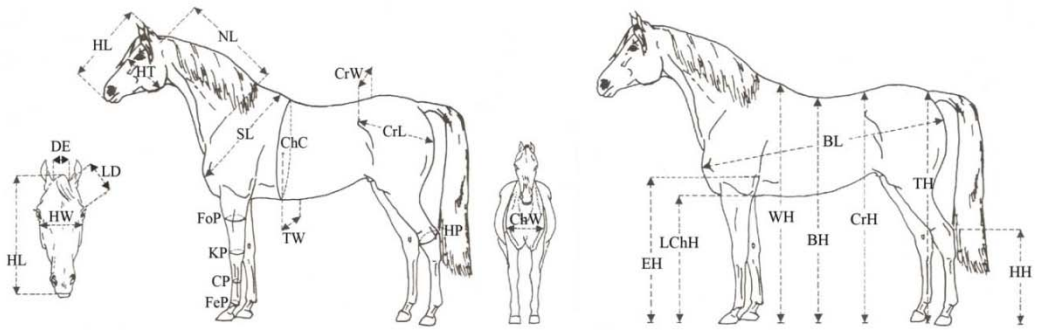


Figure 1 - Representation of the body measurements, following Oom (1992). WH = withers height; BH = back height; CrH = croup height; EH = elbow height; LChH = lowest part of chest height; HH = hock height (hind limb); TH = tail insertion height; BL = body length; CrL = croup length; SL = shoulder length; HL = head length; NL = neck length; CrW = croup width; HW = head width; DE = distance between ears; ChW = chest width; HT = head thickness; LD = longitudinal diameter; TW = thoracic width; ChC = chest circumference; FoP = forelimb perimeter; CP = cannon bone perimeter (forelimb); HP = hock perimeter (hind limb); KP = knee perimeter (forelimb); FeP = fetlock perimeter (forelimb).

Assessment of morphological measurements was always taken by the same operator in order to improve uniformity. Measurements were made using an extending measuring stick with a level on the horizontal arm (WH, BH, CrH, TH), a flexible measuring tape (EH, LChH, HH, NL, DE, LD, ChC, FoP, CP, HP, KP, FeP), a compass (CrL, SL, HL, CrW, HW, ChW, HT) and a calliper (TW, BL). All measures were taken three times to reduce error and averaged values were used in further analysis, except when taken with the compass (CrL, SL, HL, CrW, HW, ChW, HT were measured only once). Chest height (ChH, not shown in Figure 1) was calculated as the difference between WH and LChH.

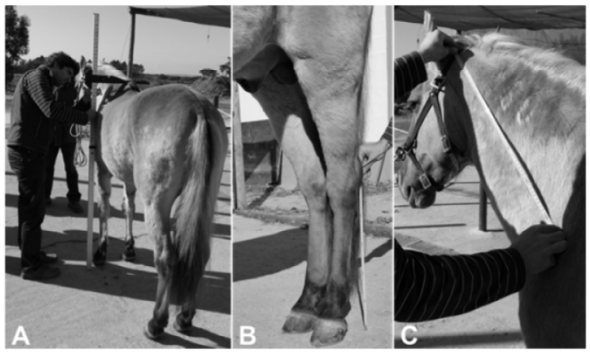


Figure 2 - Example of some measurements: (A) WH; (B) HH; (C) NL.

Inbreeding coefficients where calculated for all animals by the additive kinship matrix, considering complete pedigrees traced back to the founders (Oom et al., 2004). Age at measurement was calculated in years, as the difference between the measurement day and birth date, divided by 365. Only adult animals were measured (older than 3.5 years) so that comparisons could be done with other studies.

Statistical analysis

All statistical analyses were performed using Statistica 12© software (StatSoft, 2013) and all variables were tested for normality. Body conformation was assessed by the ratio between WH and BL, calculated for both groups and compared by Mann-Whitney U test (non-normal distribution). Morphological evolution of the breed over the past 20 years was assessed through comparison of OLD (N=8) and NEW (N=32) samples by t-test (only for variables with normal distributions) or Mann-Whitney U test (variables with non-normal distribution). The influence of inbreeding and age on the measured variables was done by regression analysis in the OLD (N=8) and NEW (N=32) samples. The quadratic effect of inbreeding and age was tested using the GLM procedure in Statistica© (StatSoft, 2013). Coat colour influence on height-related measurements and body length was tested by ANOVA in the NEW group only, due to the small sample size in the OLD subset. Principal component analysis (PCA) on body measurements was performed to determine which variables account for most of the phenotypic variation and how OLD and NEW populations are related. For each body measurement, data were standardized as in Oom (1992), in order to give each variable equal weight in the analysis.

Although the stallions analysed come from different breeders, they were kept by different owners and subjected to different care, management and training purposes, which could bias the results of the stud farm effect on the body measurements, whereby we did not analyse this variable effect.

Results & Discussion

The Sorraia horse breed has about 150 breeding mares, much less than other Portuguese autochthonous horse breeds: ~1400 in the Garrano (Santos and Ferreira, 2012) and ~5000 in the Lusitano (Vicente et al., 2012), but more than the ~100 animals in total of the Terceira pony (Lopes et al., 2015).Descriptive statistics of all variables analysed are given in Table 1. Due to the higher number of animals in the NEW (N=32) group in relation to the OLD (N=8), all variables have a wider distribution in the former (Figure 3, Table 1). On average, animals

from the OLD group have higher morphometric values, whereas animals from the NEW group have higher inbreeding coefficients (Figure 3A) and age values (Figure 3B).

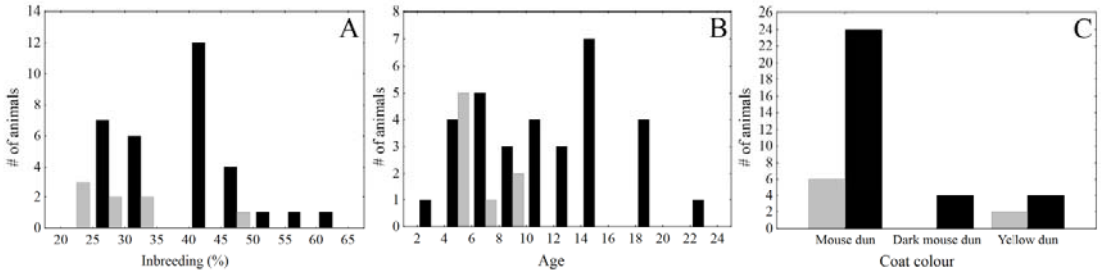


Figure 3 - Frequency distribution of (A) inbreeding coefficients, (B) age and (C) coat colour of OLD (grey bars) and NEW (black bars) measured animals.

Average WH was 148.0±2.7cm in the OLD group and 145.5±3.9cm in the NEW group (Table 1). The horses from the NEW group are shorter than the ones from the OLD group by approximately 2.5cm. Of the measured animals, 8% of the OLD group and 16% of the NEW group were taller than the described standard of 148cm (Oom et al., 2004). Average BL was 143.0±4.7cm in the NEW group and 147.7±2.4cm in the OLD (Table 1). With a larger sample, we corroborate the previous result of Oom (1992) mentioned above, when comparing with other Portuguese autochthonous horse breeds, that the Sorraia is shorter than the Lusitano, with an average of about 160cm (Oom and Ferreira, 1987; Oom, 1992; Solé et al., 2013; Vicente et al., 2014a), and taller than the Garrano, with 130cm (Oom, 1992; Santos and Ferreira, 2012), and the Terceira Pony, with 128cm (Lopes et al., 2015). The breed followed the same tendency as Finnhorses that became shorter over a 20 to 30-year period, although, in this breed, there was an assumed strategy in the breeding program where lighter sport horses were preferentially used for breeding (Saastamoinen, 1990). Our results are also in accordance with Gandini et al. (1992) that reported a decrease of 1.1cm on WH by for a 10% increase in inbreeding in Haflinger horses with 4 years and older over an eight-year period. In contrast, the Murguese and the Spanish Arab horses showed an increase in height over a similar 20-year period (4cm and 0.4cm, respectively) (Cervantes et al., 2009; Dario et al., 2006).

Sorraias are considered ponies by FEI regulations (FEI, 2014), but the Sorraia Breeders Association considers them as "small horses" (Oom et al., 2004). In fact, molecular genetics revealed that the Sorraia is not a pony, clustering with standard-sized horses (Andalusian and Lusitano) and not with Iberian pony breeds, like the Garrano (Luis et al., 2007b).

Table 1. Descriptive statistics of the variables analysed. Number of animals analysed (Valid N); Standard deviation (SD); coefficient of variation percentage (CV).

		N		Mean ± SD		Minimum		Maximum		CV (%)	
	Source	OLD	NEW	OLD	NEW	OLD	NEW	OLD	NEW	OLD	NEW
	Inbreeding (%)	8	32	29.3 ± 7.8	39.3 ± 8.9	22.5	26.5	45.6	60.1	26.8	22.7
	Age	8	32	6.8 ± 1.8	11.8 ± 5.2	4.7	3.8	9.4	23.6	26.6	43.9
HEAD	HL	8	32	54.8 ± 1.9	56.9 ± 1.8	52	54	57	60	3.5	3.2
	HW	8	32	18.5 ± 0.7	17.4 ± 1.7	18	11.5	19.5	19	3.5	9.7
	HT	8	32	26.8 ± 1.0	26.1 ± 1.8	26	23	28.5	29	3.6	6.9
	DE	8	32	10.8 ± 1.0	11.7 ± 0.8	10	9.8	12.5	13.8	8.9	7
	LD	8	32	16.1 ± 1.3	15.1 ± 1.0	14.5	12.8	18.5	17.5	8.1	6.7
BODY	NL	8	32	60.8 ± 3.3	61.6 ± 2.6	55	55.5	64	68	5.4	4.2
	WH	8	32	148.0 ± 2.7	145.5 ± 3.9	143	138.8	152.5	153.8	1.9	2.7
	BH	8	32	141.1 ± 2.2	138.6 ± 3.6	136.5	131.5	144	146	1.6	2.6
	CrH	8	32	146.4 ± 2.3	144.3 ± 3.5	143.5	136.5	150	152	1.5	2.4
	TH	7	32	134.1 ± 4.4	134.9 ± 3.9	125	125.3	138.5	144	3.3	2.9
	BL	8	32	147.7 ± 2.4	143.0 ± 4.7	144	133.5	151	153	1.6	3.3
	TW	7	32	42.6 ± 5.7	43.7 ± 6.0	37	36	52.5	55.5	13.3	13.8
	ChH	8	32	68.5 ± 1.3	68.5 ± 2.8	67	63.6	71	73.8	2	4.1
	ChW	8	32	36.3 ± 4.6	33.6 ± 3.3	32	24	44	43	12.6	9.9
	ChC	8	32	168.3 ± 8.3	168.6 ± 7.9	159	152.3	183	188	5	4.7
	LChH	8	32	79.5 ± 1.9	76.9 ± 2.1	75.5	72	81.5	80	2.4	2.7
	CrW	7	32	49.6 ± 2.4	47.4 ± 1.9	45	44	52.5	51	4.8	4.1
	CrL	8	32	47.9 ± 2.2	48.7 ± 3.0	45.5	44	51	54	4.5	6.1
LEGS	SL	8	32	61.2 ± 2.7	60.6 ± 2.6	57.5	56	65	70	4.3	4.2
	EH	8	32	88.7 ± 1.8	85.7 ± 2.1	86	81.5	91	89.8	2	2.4
	FoP	8	32	35.1 ± 1.7	35.1 ± 2.0	34	30.3	39	40.3	4.9	5.8
	KP	8	32	30.4 ± 0.8	29.5 ± 1.3	29	25.8	31	33.3	2.6	4.3
	CP	8	32	18.6 ± 1.1	18.2 ± 0.7	17	17	20	20	5.6	4.1
	FeP	8	32	25.4 ± 1.0	24.9 ± 1.0	23.5	23.3	26.5	27.3	3.9	4.1
	HH	8	32	57.4 ± 1.2	56.2 ± 2.0	56	51.5	59	61	2	3.5
	HP	8	32	39.7 ± 0.9	39.4 ± 1.6	38.5	35.8	41	43	2.2	4

The body conformation ratio of a breed (WH/BL) is an important tool to evaluate the purpose for which they are bred: sport, work, leisure or meat production. Ponies are described as being longer than taller, with a WH/BL ratio <1 (Aparicio Sánchez, 1944), contrary to horses (ratio >1), though there is no consistency for this value in different pony and horse breeds. The Sorraia has a WH/BL ratio ≈1 (1.002 in the OLD and 1.017 in the NEW, with no statistically difference between groups), being a well proportionate horse, with a square body shape (Oom and Ferreira, 1987; Oom, 1992). The Icelandic, Burguete and Jaca Navarra ponies, for example, have a WH/BL ratio of 0.983, 0.927 and 0.947, respectively (Gómez et

al., 2012; Kristjansson et al., 2013). On the other hand, the Cavall Pirinenc Català, the Hispano-Bretón from Burgos, the Hispano-Bretón from León, and the Hispano-Bretón from Palencia have a WH/BL ratio of 0.931, 0.941, 0.912, 0.939, respectively, and are all considered horse breeds (Gómez et al., 2012). The fact that these last mentioned horse breeds have a WH/BL ratio <1 could be due to the fact that they are mostly bred for meat production and so animals with bigger body mass in proportion to height will be preferred (Gómez et al., 2012). Therefore, considering that the Sorraia has a WH/BL ratio slightly above the previous examples, it should be considered a horse and not a pony.

Table 2 - T-test (normally distributed variables, t-value) and Mann-Whitney tests (non-normally distributed variables, Z) between OLD and NEW groups. Statistically significant differences in bold ($p<0.05$).

	Source	t-value	Z	p-value
	Inbreeding (%)	2.9		**0.0061**
	Age	2.66		**0.0115**
	WH/BL		1.59	0.112
HEAD	HL	2.88		**0.0065**
	HW		-1.74	0.0816
	HT		-1.18	0.2366
	DE		2.37	**0.0179**
	LD	-2.33		**0.0251**
BODY	NL	0.75		0.4552
	WH	-1.72		0.093
	BH	-1.85		0.0722
	CrH	-1.57		0.1253
	TH	0.45		0.6547
	BL	-2.72		**0.0097**
	TW		0.46	0.6473
	ChH	0.03		0.9758
	ChW	-1.88		0.0678
	ChC	0.13		0.9004
	LChH	-3.21		**0.0027**
	CrW	-2.6		**0.0133**
	CrL	0.75		0.458
LEGS	SL		-0.68	0.4989
	EH	-3.74		**0.0006**
	FoP		0.51	0.612
	KP		-2.21	**0.0268**
	CP	-1.08		0.2886
	FeP	-1.14		0.2596
	HH	-1.68		0.1008
	HP	-0.47		0.6445

Simultaneously, WH is slightly higher than CrH in both groups, which is in agreement with the results found in the Lusitano (Oom and Ferreira, 1987; Oom, 1992), the most relevant Portuguese saddle horse. Statistically significant differences between the OLD and NEW groups were found for inbreeding coefficient and age, as well as for HL, DE, LD, BL, LChH, CrW, EH, and KP (Table 2). However, as mentioned before, no statistically significant difference was found for body conformation (WH/BL) between OLD and NEW groups (Table 2).

There was a trend for chest circumference (ChC) to increase over time in Sorraia stallions as in the Murguese horse (Dario et al., 2006) (although not statistically significant), which might be explained by the fact that the Sorraias in the NEW group are almost daily used in riding centres where a bigger chest and higher aerobic capacity is preferred, whereas animals from the OLD were mainly work horses used for reproduction or for cattle driving. Cannon bone perimeter (CP) has decreased from OLD to NEW, contrary to the Murguese horse (Dario et al., 2006), probably due to the fact that more recently Sorraia are bred mostly for leisure and sports rather than for cattle-related work, and thus would no longer be selected for a sturdier bone support. Feeding particularities and genetics could also explain these results.

Only horses 3.5 years and older were measured, the oldest having 23.6 years. Although it may be considered that horses only reach adult size at 5-6 years of age, we measured horses 3.5 years and older due to reduced effective population and the difficulty to find animals easy to handle, as explained in Oom (1992). Even so, age at measurement is in accordance with other morphometric studies performed in horses, where only adult animals were measured: 6 years and older in the Lusitano (Oom and Ferreira, 1987); 6 years and older in the Lusitano, Pura Raza Española, Berber and Arab, and 4 years and older in the Sorraia and Garrano (Oom, 1992); 6.37 years in average in the Lusitano (Vicente et al., 2014b); 3 to 14 years in Spanish Arab horses (Cervantes et al., 2009); 4 to 10 years in Icelandic horses (Kristjansson et al., 2013); minimum of 4 years in Finnhorses (Saastamoinen, 1990) and Lipizzaner (Curik, 2003; Zechner et al., 2001) and Spanish heavy horse breeds (Gómez et al., 2012); 5 years in Nordestino (Melo et al., 2011); 3 and 3.94 years in Spanish Purebred horses (Gómez et al., 2009a; Gómez et al., 2009b); 4 to 34.5 years in a study with 65 different horse breeds (Brooks et al., 2010); 4.7 to 18.3 in Mangalarga Marchador (Cabral et al., 2004); 4 to 14 in Lusitano, Spanish and Menorca horses (Solé et al., 2013). Average age of measured animals was higher in the NEW group (11.8±5.2 years) than in the OLD (6.8±1.8 years), as was the maximum age (23.6 in the NEW VS 9.4 years in the OLD). Minimum age was similar in both groups (3.8 in the NEW and 4.7 in the OLD).

Table 3 - Regression results for inbreeding and age effect on body measurements. Statistically significant results in bold ($p<0.05$). SE - Standard error.

		Inbreeding (%)						Age (years)					
		Effect		SE		p-value		Effect		SE		p-value	
	Source	OLD	NEW	OLD	NEW	OLD	NEW	OLD	NEW	OLD	NEW	OLD	NEW
HEAD	HL	0.12	-0.01	0.09	0.04	0.2235	0.8719	0.67	-0.04	0.34	0.06	0.0965	0.4966
	HW	-0.01	-0.02	0.03	0.03	0.7686	0.5126	-0.11	0.04	0.14	0.06	0.4697	0.5031
	HT	0	0.05	0.05	0.04	0.9544	0.2169	-0.29	-0.02	0.18	0.06	0.1644	0.7095
	DE	-0.01	-0.03	0.05	0.02	0.7868	0.0873	0.49	0.04	0.08	0.03	**0.0011**	0.2118
	LD	-0.03	0.01	0.07	0.02	0.6849	0.5585	0.62	0.002	0.15	0.04	**0.0063**	0.9654
BODY	NL	0.12	-0.02	0.16	0.05	0.4901	0.7258	-1.43	-0.03	0.45	0.09	**0.0197**	0.7659
	WH	0.22	-0.04	0.11	0.08	0.0851	0.6264	0.16	-0.2	0.69	0.13	0.7451	0.1485
	BH	0.12	-0.01	0.1	0.07	0.282	0.8614	0.18	-0.2	0.5	0.12	0.7286	0.1107
	CrH	0.09	-0.07	0.11	0.07	0.4445	0.2986	0.62	-0.12	0.44	0.12	0.2129	0.3387
	TH	0.27	-0.12	0.22	0.08	0.2611	0.1358	-0.88	0.05	0.96	0.14	0.3997	0.7366
	BL	0.15	-0.08	0.11	0.09	0.2094	0.4332	0.57	0.03	0.48	0.17	0.2801	0.8361
	TW	-0.13	-0.38	0.29	0.1	0.6651	**0.0009**	1.13	0.15	1.51	0.21	0.4881	0.4739
	ChH	0.08	-0.07	0.06	0.06	0.2386	0.2329	-0.36	-0.19	0.26	0.09	0.2146	**0.045**
	ChW	-0.01	-0.06	0.24	0.07	0.9742	0.3712	2.09	-0.06	0.58	0.12	**0.0117**	0.5866
	ChC	-0.02	-0.39	0.43	0.15	0.9641	**0.0113**	2.06	-0.6	1.68	0.26	0.2681	**0.0261**
	LChH	0.14	0.03	0.08	0.04	0.1163	0.5007	0.57	-0.01	0.36	0.07	0.1597	0.9356
	CrW	0.29	0.02	0.2	0.04	0.197	0.5398	0.45	-0.02	0.53	0.07	0.4322	0.7609
	CrL	-0.04	-0.02	0.11	0.06	0.6857	0.7357	0.17	-0.29	0.48	0.09	0.7386	**0.0028**
LEGS	SL	0.09	-0.01	0.13	0.05	0.5094	0.8188	0.77	-0.09	0.51	0.09	0.1824	0.3036
	EH	0.17	0.01	0.06	0.04	**0.0295**	0.8305	0.08	-0.06	0.4	0.07	0.8408	0.3983
	FoP	-0.02	-0.01	0.09	0.04	0.8493	0.8844	0.45	-0.08	0.34	0.07	0.2403	0.2406
	KP	0.05	-0.01	0.04	0.03	0.1937	0.7236	0.16	0.004	0.17	0.04	0.3635	0.9311
	CP	0.02	-0.01	0.05	0.02	0.7354	0.3413	0.23	0.02	0.22	0.03	0.3369	0.4633
	FeP	0.04	-0.02	0.05	0.02	0.4094	0.4425	0.17	0.05	0.21	0.04	0.4668	0.1687
	HH	0.07	-0.08	0.05	0.04	0.2497	**0.0404**	0.33	-0.12	0.22	0.07	0.1904	0.0675
	HP	0.08	-0.01	0.03	0.03	0.0504	0.6632	0.18	-0.11	0.19	0.05	0.3752	**0.038**

As age was statistically different between OLD and NEW samples, its effects were analysed separately within each group. Overall, age had no significant effect on the majority of the body measurements. In the OLD group, age had a significant positive impact on DE, LD, and ChW, as well as a significant negative impact on NL (Table 3). In the NEW group, age had a significant negative effect on ChH and a significant positive effect on ChC, CrL, and HP (Table 3). The quadratic effect of age was significant only for the OLD group with a negative impact on NL, WH, BH, ChH, LChH, and EH (Supplementary Table 1). The negative quadratic impact of age on body measurements can be explained by the overall loss of body condition with advancing age. Particularly in the OLD group and contrary to the

results of Carolino and Gama (2008) and Gómez et al. (2009b), the quadratic effect of age was significant and negative for NL, WH, BH, ChH, LChH and EH, all of which are also related to height that tends to reduce as horses get older. In humans, height reduction with increased age results from osteoporotic changes, while it can increase due to improved nutrition and living standards over the last century (McQuillan et al., 2012). However, the latter was not a confounding effect in the analysis of the influence of age in the reduction of height in humans. In Icelandic horses, Kristjansson et al. (2013) reported no effect of age on body measurements given that growth plates are usually closed by the age of 3. Our results are in agreement with the findings in the Haflinger horse where age had a negative impact on height at withers and girth circumference, but not on cannon bone circumference (Gandini et al., 1992).

Table 4 - Mean values ± Standard Error (SE) and ANOVA results for height measurements and body length by coat colour on animals from the NEW group (N=38). Statistically significant differences in bold ($p<0.05$).

	Mean ± SE					
Source	Mouse dun	Dark mouse dun	Yellow dun	df	F	p-value
WH	145.6 ± 0.8	143.1 ± 2.0	147.1 ± 2.0	2	1.0516	0.3623
BH	138.7 ± 0.7	137.2 ± 1.8	139.3 ± 1.8	2	0.3932	0.6784
CrH	144.7 ± 0.7	140.9 ± 1.7	145.3 ± 1.7	2	2.4693	0.1023
TH	136.0 ± 0.7	129.8 ± 1.7	133.1 ± 1.7	2	6.8325	**0.0037**
ChH	68.5 ± 0.6	67.4 ± 1.4	69.9 ± 1.4	2	0.7965	0.4605
LChH	77.1 ± 0.4	75.8 ± 1.0	77.2 ± 1.0	2	0.7306	0.4903
EH	86.0 ± 0.4	84.0 ± 1.0	86.0 ± 1.0	2	1.6745	0.2050
HH	56.3 ± 0.4	54.1 ± 0.9	57.4 ± 0.9	2	3.4763	**0.0443**
BL	143.5 ± 0.9	138.9 ± 2.3	144.1 ± 2.3	2	1.8592	0.1739

Only two coat colours are described in the Sorraia horse Studbook: mouse dun and yellow dun (Oom et al., 2004), the former being most common (Figure 3C), even though yellow mouse dun may be considered a more primitive coat colour (Ludwig et al., 2009), as found in the Przewalski horse. In recent years (in the nineties), new coat colour phenotypes (dark mouse dun and dark yellow dun) were recorded in this breed (Figure 3C). These phenotypes are due to the presence of a not-Dun allele (nd1) in heterozygosity for the coat colour locus D (presence of dun dilution and primitive markings), as found by DNA tests performed on UCDAVIS Veterinary Genetics Laboratory (www.vgl.ucdavis.edu/services/dunhorse.php,

January 2016). The influence of coat colour in height-related measurements and body length is considered in Table 4. Dark mouse dun horses are shorter than mouse and yellow duns. However, this difference is only statistically significant for TH and HH. Yellow duns are taller and longer than mouse duns except for TH, while EH is approximately the same on both colours. To the best of our knowledge, to this date, there are no other studies that report a relation between coat colour and height or body length, which makes this the first study where it has been proven that a specific coat colour, dark mouse dun, is shorter than the others.

The Sorraia breed has extremely high inbreeding coefficients (Kjöllerström et al., 2015; Luis et al., 2007a; Pinheiro et al., 2013) (Figure 3A). Inbreeding has an average of 37.29% in all measured animals, 29.3%±7.8 in the OLD group and 39.3%±8.9 in the NEW (Table 1). In the OLD group, 87.5% of the animals have inbreeding coefficients below the current population average (38%) (Kjöllerström et al., 2015), whereas in the NEW one only 40.6% of the inbreeding coefficients are below this value. As mentioned before, these values are much higher than found in any other horse breed.

Such a high level of average inbreeding hasn't been found in any other horse breed: 0.65% in the Garrano (Cipriano, 2007), 8.2% in the Andalusian (Gómez et al., 2009b), 9.92% in the Lusitano (Vicente et al., 2014b), 10% in the Lipizzan (Curik, 2003), 13% in the Thoroughbred (Cunningham et al., 2001) and 15.7% in the Friesian (Sevinga et al., 2004), just to name a few. Even the endangered Przewalski horse, recovered from 13 founders, has an average inbreeding coefficient of 14% (Der Sarkissian et al., 2015), much lower than the Sorraia.

Studying mammals, Simpson et al. (1960) considered a sample to represent an adequate variability when coefficient of variation (CV) values ranged from 4 to 10, with 5 and 6 being optimal values. In our study, the CV values are low to optimal, with few exceptions (Table 1), considering both OLD and NEW groups, ranging from 1.5 to 13.8 in body measurements.

CV was highest in TW in both groups, lowest in CrH in the OLD, and in CrH and EH in the NEW (Table1). CV values for body measurements were very similar to those reported in the Sorraia by Oom 1992, where CV ranged from 1.55 (CrH) to 13.32 (TW) (Supplementary Table 2). They were also similar to other CV values obtained by Cervantes et al. (2009) in Spanish Arab horses (2.07 in WH to 10.09 in ChW), by Portas (2001) in the Garrano (4 in BH to 10 in NL) and by Oom (1992) also in the Garrano (1.90 in CrH to 10.90 in ChW). As we can also see in Supplementary Table 2, other studied horse breeds have shown lower CV values for body measurements than those described for the Sorraia horse. As in Oom (1992), we also found that the body measurements with the highest CV values were harder to measure

precisely and were related to the head (DE and LD), chest (ChW, TW and ChC) and upper leg region (FoP), due to lack of precision in the points chosen as reference, to animal temperament and posture, nutritional status and muscle development. Values shown in Supplementary Table 2 indicate WH as the body measurement that most often has the lowest CV value in all studied breeds. In the Sorraia, CV value for WH was followed by other measurements that define the horse's body shape (BODY, Table 1) and those related with the legs (LEGS, Table 1). Therefore, the Sorraia stallions sampled should be considered a very homogenous group regarding their general conformation, with low variability in some of the body measurements analysed. Head shape is getting more variable in the NEW group, reflecting the recent emergence of some heads of more triangular shape, rather than the rectangular shape characteristic of the breed.

As for inbreeding and age variables, with 22.7 - 26.8 and 43.9 - 26.6 in NEW and OLD groups, respectively (Table 1), CV values are much higher, reflecting a great variability, due to the effect of the recent management breeding plan to minimize inbreeding, in some cases, and to the high range of ages considered.

It has been extensively shown that inbreeding has significant negative impact on performance, reproductive and fitness traits of farm animals, such as fertility, growth rates, birth weight, survival and disease resistance, representing significant losses to breeders and producers (Carolino and Gama, 2008; Lacy et al., 1996; Leroy, 2014; Mc Parland et al., 2007; Queiroz et al., 2000; Ralls et al., 1988; Ralls et al., 1979; Reed and Frankham, 2003; Santana et al., 2012; Santana et al., 2010; Smith et al., 1998; Swiger et al., 1961) and should be avoided. Inbreeding may also have negative impacts in adult life history traits, such as mate choice and acquisition, survival to adulthood and parental care (Ryan et al., 2002). In the Sorraia, inbreeding significantly decreased mare fertility but had no significant effect on stallion fertility, foaling intervals, age at first parturition (Kjöllerström et al., 2015) and rate of chromosome abnormalities in stallion sperm or sperm morphology (Kjöllerström et al., 2016).

In some studies, recessive deleterious alleles were not found with increasing homozygosity, suggesting that they might have been purged from the genetic pool of the population (Frankham et al., 2002; Leberg and Firmin, 2008), though this purging can still impact the viability of small inbred populations through inbreeding depression (Charpentier et al., 2007), as the fixation of alleles can lead to reduction of heterozygote advantage (e.g. Lacy et al. (1996)). However, even after several generations of inbreeding in wild populations, purging of lethal recessive alleles may not occur, neither their fixation (Kennedy et al., 2014). The extent and specific effects of inbreeding depression will vary greatly depending on the

population's genetic constitution and the interaction between their genotypes and the environment (Hedrick and Kalinowski, 2000). Inbreeding depression studies can be severely affected by pedigree depth and completeness (Cothran et al., 1986; Smith et al., 1998), which makes the Sorraia an excellent study model since all animals have complete pedigree information back to the founders, some with 12 generations (Oom et al., 2004).

Several studies in cattle have demonstrated that highly inbred animals are smaller (Young, 1984) and narrower than non-inbred ones (Croquet et al., 2006; Mc Parland et al., 2007; Santana et al., 2010; Smith et al., 1998). However, results regarding inbreeding depression on morphometric traits in horses are scarce. In the Lusitano horse, Oom (1992) found evidence of inbreeding depression for body measurements, reporting a significant ($p<0.01$) detrimental effect of inbreeding on withers height, as later did Vicente et al. (2014b). Gandini et al. (1992) found the same effect for withers height and girth circumference in the Italian Haflinger horse. In the Lipizzaner, only the length of pastern-hind limbs was negatively affected by inbreeding in the univariate analysis, losing its significance in multiple tests (Curik, 2003).

In this study, and because inbreeding coefficients were statistically different between groups, they were analysed separately for the evaluation of inbreeding depression on body measurements (Table 3). The negative impacts of inbreeding in the NEW group were visible on TW and ChC, with positive impact on HH. In the OLD group, inbreeding had a significant positive influence on EH. Even though Carolino and Gama (2008) and Gómez et al. (2009b) did not find the quadratic effect of inbreeding to be significant, we found it to be significant and negative for BL on the OLD group (Supplementary Table 1).

The reason inbreeding didn't have a more significant effect on morphological traits in the Sorraia horse might be due to the fact that inbreeding depression is known to be more significant on fitness and life-history traits than on morphological ones (Carolino and Gama, 2008; Coltman and Slate, 2003). It might also be explained by the fact that Sorraia stallions have now a better management and handling program (nutrition and training) then what they had 20 years ago, which can be masking the detrimental effects of the statistically significant increase in inbreeding coefficients from the OLD to the NEW groups. Another possible explanation, perhaps the most plausible one, is that deleterious recessive alleles have been purged (Lacy et al., 1996) from the Sorraia horse population through generations and natural selection. This is supported by the findings of Leberg and Firmin (2008) where populations with small sizes would suffer less from future inbreeding since they had been purged of deleterious recessive alleles. However, the same authors alert for the fact that purging success in eliminating inbreeding depression is uncertain and it also likely that purging will reduce

viability by fixation of deleterious alleles, meaning that any purging strategy should be carefully weighted to prevent negative irreversible results.

It is also possible that the decrease in size is due to stallion selection, since only one stallion is used per herd, per year, using a shorter stallion would lead to a decrease in height in the progeny that wasn't due to inbreeding but to this trait's heritability.

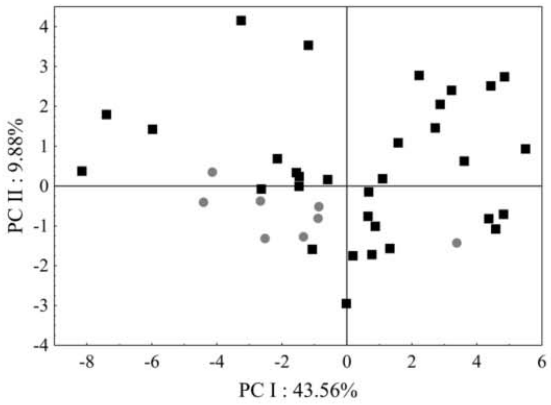


Figure 4 - Principal component analysis (PCA) scatterplot on the first two components of the 40 Sorraia stallions analysed using 26 body measurements (NEW-■, N=32; OLD-●, N=8). The fraction of the total variance explained by each of principal components is indicated on the corresponding axes.

Multivariate approach through PCA analysis has been previously used in horse morphometric studies based on a large set of variables to better understand relationship between different breeds (Gómez et al., 2012; Solé et al., 2013), breeding goals (Cervantes et al., 2009) or stud farms (Sobczuk and Komosa, 2012; Zechner et al., 2001). In our study, OLD and NEW samples did not cluster in separate groups when PCA analysis was performed using 26 body measurements (Figure 4). The first component (PC I) explains 43.56% of the total variation, whereas the second (PC II) and third (PC III) explain 9.88% and 7.48%, respectively (graphical data not shown for PC III), reaching a cumulative percentage of 60.91% (Table 5). The variables that most contributed to the discriminatory value of the first component (PC I) are WH, CrH and BH, all negatively correlated with it, whereas or the second axis (PC II), TW, DE, LD and NL, were the most correlated ones (Table 5). As in Oom (1992), height measurements and those relating to leg perimeter have an important discriminant role in the first component (PC I) (Figure 4), accounting for ordination of the animals along the PC I based on their body size and members bone structure. All variables

with higher impact on animal distribution along PC II revealed some difficulty in accurate measurement, since horses are sometimes reluctant to allow touching their ears, and also due to lack of bone structure in which to support those measurements (Oom, 1992). As hypothesised by Oom (1992), DE and LD should be taken directly from skulls, which will be very difficult due to the lack of such material in large animals, namely horses. The fact that OLD and NEW groups did not distribute in separate clusters by PCA analysis might be due, on one hand, to the great homogeneity of morphometric measurements and their high correlation, and secondly, to the small number of animals in the OLD group.

Table 5 - Factor loading, eigenvalues, percentage of total variance and cumulative variation for the first 3 principal components, obtained from PCA analysis based on the correlation matrix between the original data for each of the 26 body measurements. For body measurement legend see Figure 1.

		Principal Component					
		I		II		III	
	Body measurement	Factor loadings	Eigenvector	Factor loadings	Eigenvector	Factor loadings	Eigenvector
HEAD	HL	-0.6105	-0.1814	0.2901	0.181	0.3932	0.282
	HW	-0.3982	-0.1183	0.3813	0.2379	-0.534	-0.383
	HT	-0.6821	-0.2027	-0.238	-0.1485	-0.0907	-0.065
	DE	-0.0945	-0.0281	0.665	0.4149	0.4576	0.3282
	LD	-0.4793	-0.1424	-0.4144	-0.2586	0.3813	0.2735
BODY	NL	-0.2887	-0.0858	0.5516	0.3442	-0.4214	-0.3022
	WH	-0.9197	-0.2733	-0.0762	-0.0476	-0.0167	-0.012
	BH	-0.8512	-0.2529	-0.2268	-0.1415	0.0762	0.0547
	CrH	-0.9113	-0.2708	0.0022	0.0014	0.0781	0.056
	TH	-0.6365	-0.1891	0.1664	0.1038	0.0271	0.0195
	BL	-0.8105	-0.2408	0.1422	0.0887	-0.1437	-0.1031
	TW	0.2297	0.0683	0.8163	0.5093	0.0543	0.0389
	ChH	-0.7391	-0.2196	0.169	0.1054	0.2141	0.1535
	ChW	-0.5374	-0.1597	0.2693	0.168	-0.0019	-0.0014
	ChC	-0.6929	-0.2059	0.1908	0.119	0.4798	0.3442
	LChH	-0.7177	-0.2133	-0.3196	-0.1994	-0.2734	-0.1961
	CrW	-0.6419	-0.1907	-0.1308	-0.0816	-0.2056	-0.1475
	CrL	-0.5491	-0.1632	-0.1754	-0.1094	0.4341	0.3114
LEGS	SL	-0.7522	-0.2235	-0.0638	-0.0398	0.0962	0.069
	EH	-0.7175	-0.2132	-0.3234	-0.2018	-0.4163	-0.2986
	FoP	-0.6441	-0.1914	-0.2851	-0.1779	0.191	0.137
	KP	-0.7118	-0.2115	0.1392	0.0869	-0.1741	-0.1249
	CP	-0.6444	-0.1915	0.1479	0.0923	0.01	0.0071
	FeP	-0.7063	-0.2099	0.201	0.1254	-0.2684	-0.1925
	HH	-0.6296	-0.1871	0.1597	0.0996	-0.1507	-0.1081
	HP	-0.7965	-0.2367	0.0905	0.0565	0.0331	0.0237
Eigenvalue		11.3252		2.5687		1.9436	
% total variance		43.5586		9.8797		7.4754	
Cumulative %		43.5586		53.4382		60.9137	

Conclusion

Phenotypic characterization of a breed is one of the most important and basic information of AnGR Conservation Management Programs (Melo et al., 2011), that complement genetic and historical information, being therefore fundamental to the establishment of national inventories of AnGR and to effective monitoring of AnGR populations, by establishing the basis for which to start management breeding programs (FAO, 2012). Morphometric measurements also give precious data for the establishment or realistic breed standards, and to define breed's deviation from it in different generations.

Although effects of inbreeding on the majority of traits analysed on the Sorraia horse, so far, are not alarming, studying inbreeding depression in the Sorraia horse is of paramount importance for its conservation and sustainability. Pariacote et al. (1998) suggested that the negative effects of inbreeding depression could be mitigated by selection through management breeding programs, thus only by the continuous implementation of an updated breeding strategy will it be possible to preserve this extremely important and critically endangered genetic resource. Different horse breeds, such as the Italian Haflinger (Gandini et al., 1992), Spanish Purebred (Gómez et al., 2009b) and Lusitano (Oom, 1992; Vicente et al., 2014b) were shown to be affected by the negative effects of inbreeding on morphological traits. Researching these effects on Sorraia horses is very important as overall morphological conformation can determine athletic ability and vigour, very relevant in reproductive success, mainly if reproduction is performed extensively. They can also give us indications of the possible occurrence of inbreeding depression in other traits, even if they are not visible, and should be considered as part of the data analysed during breed management planning. As suggested by Lacy et al. (1996) the failure to detect significant inbreeding depression can be due to sample sizes too small to allow statistically significant results, unless it is abnormally strong and easily detected. This might be the case in our study, since only 40 animals were measured and increasing sample size would improve statistical power of the analysis.

Given the importance of inbreeding depression and its pernicious effects on life history and morphological traits, this study should be undertaken in Sorraia mares and the results compared to those presented here to see if inbreeding depression is differentially affecting both sexes, as it has been shown to be the case in primate species where females are more affected than males (Charpentier et al., 2007). Breeding strategies and conservation plans for the Sorraia horse should include data concerning evidences of inbreeding depression in order to establish a long-term self-sustaining population and long term-studies should be

implemented to track the breeds' evolution. Thus, further studies should be implemented, not only regarding different samples but also analysing different traits.

Acknowledgments

The authors wish to thank Rui P. de Matos for his collaboration with measurement data collection. These data were also intended to be used in the study developed for his Master's degree dissertation. We also want to thank Prof. Paulo Pardal and Prof. António Vicente, from Escola Superior Agrária de Santarém, Portugal, for their support in measurements sessions and by providing measurement equipment. The authors also wish to thank the Associação Internacional de Criadores do Cavalo Ibérico de Tipo Primitivo – Sorraia for providing privileged contacts with the breeders as well as pedigree data.

Funding sources

This work was supported by cE3c FCT Unit funding (Ref. UID/BIA/00329/2013) and by PRODER (Contract number 57882) (EU)/Action 2.2.3 (2011-2013) H. J. Kjöllerström was supported by a PhD grant (SFRH/BD/81502/2011) from the Portuguese Foundation for Science and Technology (FCT/MEC).

References

Andrade, R. 1945. O Cavalo do Sorraia. Boletim Pecuário XIII: 1-13.

Aparicio Sánchez, G. 1944. Fenotipología animal. Zootecnia V (7-8): 5-31.

Brooks, S. A. et al. 2010. Morphological variation in the horse: defining complex traits of body size and shape. Animal Genetics 41 Suppl 2: 159-165.

Cabral, G. C. et al. 2004. Avaliação morfométrica de eqüinos da raça Mangalarga Marchador: índices de conformação e proporções corporais. Revista Brasileira de Zootecnia 33: 1798-1805.

Carolino, N., and L. T. Gama. 2008. Inbreeding depression on beef cattle traits: estimates, linearity of effects and heterogeneity among sire-families. Genetics, selection, evolution 40: 511-527.

Cervantes, I. et al. 2009. Size and shape analysis of morphofunctional traits in the Spanish Arab horse. Livestock Science 125: 43-49.

Charpentier, M. J. E., A. Widdig, and S. C. Alberts. 2007. Inbreeding depression in non-human primates: a historical review of methods used and empirical data. American Journal of Primatology 69: 1370-1386.

Cipriano, F. G. F. 2007. Diversidade genética da raça garrana baseada em MTDNA, microssatélites e dados genealógicos. MsC Dissertation.

Coltman, D. W., and J. Slate. 2003. Microsatellite measures of inbreeding: a meta-analysis. Evolution; international journal of organic evolution 57: 971-983.

Cothran, E. G., J. W. MacCluer, L. R. Weitkamp, and S. A. Guttormsen. 1986. Genetic variability, inbreeding, and reproductive performance in standardbred horses. Zoo Biol 5: 191-201.

Croquet, C., P. Mayeres, A. Gillon, S. Vanderick, and N. Gengler. 2006. Inbreeding Depression for Global and Partial Economic Indexes, Production, Type, and Functional Traits. Journal of Dairy Science 89: 2257-2267.

Cunningham, E. P., J. J. Dooley, R. K. Splan, and D. G. Bradley. 2001. Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. Animal Genetics 32: 360-364.

Curik, I. 2003. Inbreeding, Microsatellite Heterozygosity, and Morphological Traits in Lipizzan Horses. Journal of Heredity 94: 125-132.

Dario, C., D. Carnicella, M. Dario, and G. Bufano. 2006. Morphological evolution and heritability estimates for some biometric traits in the Murgese horse breed. Genetics and molecular research : GMR 5: 309-314.

Der Sarkissian, C. et al. 2015. Evolutionary Genomics and Conservation of the Endangered Przewalski's Horse. Curr Biol 25: 2577-2583.

FAO. 2012. Phenotypic characterization of animal genetic resources. FAO Animal Production and Health Guidelines, Rome.

FAO. 2015. The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture. In: B. D. Scherf and D. Pilling (eds.). FAO Commission on Genetic Resources for Food and Agriculture Assessments, Rome.

FEI. 2014. Dressage Rules. In: F. E. Internationale (ed.). p 124. Fédération Equestre Internationale, Lausanne, Switzerland.

Frankham, R., D. A. Briscoe, and J. D. Ballou. 2002. Introduction to Conservation Genetics. Cambridge University Press.

Gandini, G. C., A. Bagnato, F. Miglior, and G. Pagnacco. 1992. Inbreeding in the Italian Haflinger horse. Journal of animal breeding and genetics = Zeitschrift fur Tierzuchtung und Zuchtungsbiologie 109: 433-443.

Gómez, M. D., P. J. Azor, M. E. Alonso, J. Jordana, and M. Valera. 2012. Morphological and genetic characterization of Spanish heavy horse breeds: Implications for their conservation. Livestock Science 144: 57-66.

Gómez, M. D., F. Goyache, A. Molina, and M. Valera. 2009a. Sire × stud interaction for body measurement traits in Spanish Purebred horses. Journal of animal science 87: 2502-2509.

Gómez, M. D., M. Valera, A. Molina, J. P. Gutiérrez, and F. Goyache. 2009b. Assessment of inbreeding depression for body measurements in Spanish Purebred (Andalusian) horses. Livestock Science 122: 149-155.

Hedrick, P. W., and S. T. Kalinowski. 2000. Inbreeding depression in conservation biology. Annu Rev Ecol Syst 31: 139-162.

Kennedy, E. S., C. E. Grueber, R. P. Duncan, and I. G. Jamieson. 2014. Severe inbreeding depression and no evidence of purging in an extremely inbred wild species--the Chatham Island black robin. Evolution; international journal of organic evolution 68: 987-995.

Kjöllerström, H. J. 2005. O Studbook como instrumento na análise genética e demográfica de populações ameaçadas: o caso do Cavalo do Sorraia (in Portuguese). Undergraduate Thesis, Faculty of Sciences, University of Lisbon, Portugal, University of Lisbon, Portugal.

Kjöllerström, H. J., L. T. Gama, and M. M. Oom. 2015. Impact of inbreeding on fitness-related traits in the highly threatened Sorraia horse breed. Livestock Science 180: 84-89.

Kjöllerström, H. J., M. M. Oom, B. P. Chowdhary, and T. Raudsepp. 2016. Fertility Assessment in Sorraia Stallions by Sperm-Fish and Fkbp6 Genotyping. Reproduction in domestic animals = Zuchthygiene 51: 351-359.

Kristjansson, T. et al. 2013. Objective quantification of conformation of the Icelandic horse based on 3-D video morphometric measurements. Livestock Science 158: 12-23.

Lacy, R. C., G. Alaks, and A. Walsh. 1996. Hierarchical Analysis of Inbreeding Depression in *Peromyscus polionotus*. Evolution; international journal of organic evolution 50: 2187-2200.

Leberg, P. L., and B. D. Firmin. 2008. Role of inbreeding depression and purging in captive breeding and restoration programmes. Molecular ecology 17: 334-343.

Leroy, G. 2014. Inbreeding depression in livestock species: review and meta-analysis. Animal Genetics 45: 618-628.

Lopes, M. S. et al. 2015. Morphological and genetic characterization of an emerging Azorean horse breed: the Terceira Pony. Frontiers in genetics 6: 62.

Ludwig, A. et al. 2009. Coat Color Variation at the Beginning of Horse Domestication. Science 324: 485.

Luis, C., E. G. Cothran, and M. M. Oom. 2007a. Inbreeding and genetic structure in the endangered Sorraia horse breed: implications for its conservation and management. Journal of Heredity 98: 232-237.

Luis, C., R. Juras, M. M. Oom, and E. G. Cothran. 2007b. Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. Animal Genetics 38: 20-27.

Mc Parland, S., J. F. Kearney, M. Rath, and D. P. Berry. 2007. Inbreeding Effects on Milk Production, Calving Performance, Fertility, and Conformation in Irish Holstein-Friesians. Journal of Dairy Science 90: 4411-4419.

McQuillan, R. et al. 2012. Evidence of Inbreeding Depression on Human Height. PLoS genetics 8: 1-14.

Melo, J. B., D. A. F. Pires, M. N. Ribeiro, D. O. Santos, and H. G. O. Silva. 2011. Estudo zoométrico de remanescentes da raça equina Nordestina no município de Floresta, Pernambuco – Brasil. AICA 1: 71-74.

Oom, M., and J. Ferreira. 1987. Estudo biométrico do cavalo Alter. Rev Port Cienc Vet 83: 101-148.

Oom, M. M. 1992. O cavalo Lusitano. Uma raça em Recuperação (in Portuguese). PhD Dissertation, University of Lisbon, Portugal.

Oom, M. M., J. L. d'Andrade, and J. Costa-Ferreira. 2004. Studbook da raça Sorraia (in Portuguese). Garrido Artes Gráficas, Alpiarça, Portugal.

Pariacote, F., L. D. Van Vleck, and M. D. MacNeil. 1998. Effects of inbreeding and heterozygosity on preweaning traits in a closed population of Herefords under selection. Journal of animal science 76: 1303-1310.

PDR2020. 2014. Ação 7.8 - Recursos Genéticos. In: M. d. A. e. d. M. Governo de Portugal (ed.) Programa de Desenvolvimento Rural do Continente para 2014-2020. Governo de Portugal, Ministério da Agricultura e do Mar, Portugal.

Pinheiro, M., H. J. Kjöllerström, and M. M. Oom. 2013. Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis. Livestock Science 152: 1-10.

Portas, M. C. P. 2001. La Conservación de la raza equina Garrana. Archivos de Zootecnia 50.

Queiroz, S. A., L. G. Albuquerque, and N. A. Lanzoni. 2000. Inbreeding effects on growth traits of Gyr cattle in Brazil (in Portuguese). Revista Brasileira de Zootecnia 29: 1014-1019.

Ralls, K., J. D. Ballou, and A. Templeton. 1988. Estimates of Lethal Equivalents and the Cost of Inbreeding in Mammals. Conservation Biology 2: 185-193.

Ralls, K., K. Brugger, and J. Ballou. 1979. Inbreeding and juvenile mortality in small populations of ungulates. Science 206: 1101-1103.

Reed, D. H., and R. Frankham. 2003. Correlation between Fitness and Genetic Diversity. Conservation Biology 17: 230-237.

Ryan, K. K., C. L. Robert, and S. W. Margulis. 2002. Impacts of inbreeding on components of reproductive success. In: W. V. Holt, A. R. Pickard, J. C. Rodger and J. C. Wildt (eds.) Reproduction Science and Integrated Conservation. p 82-96. Cambridge University Press, Cambridge, United Kingdom.

Saastamoinen, M. 1990. Factors Affecting Growth and Development of Foals and Young Horses. Acta Agriculturae Scandinavica 40: 387-396.

Santana, M. L., Jr., P. S. Oliveira, J. P. Eler, J. P. Gutierrez, and J. B. Ferraz. 2012. Pedigree analysis and inbreeding depression on growth traits in Brazilian Marchigiana and Bonsmara breeds. Journal of animal science 90: 99-108.

Santana, M. L. J. et al. 2010. Effect of inbreeding on growth and reproductive traits of Nellore cattle in Brazil. Livestock Science 131: 212-217.

Santos, A. S., and L. M. M. Ferreira. 2012. The Portuguese Garrano breed: an efficient and sustainable production system. In: M. Saastamoinen, M. Fradinho, A. Santos and N. Miraglia (eds.) Forages and grazing in horse nutrition. Forages and grazing in horse nutrition No. 132. p 481-484. Wageningen Academic Publishers.

Sevinga, M., T. Vrijenhoek, J. W. Hesselinks, H. W. Barkema, and A. F. Groen. 2004. Effect of inbreeding on the incidence of retained placenta in Friesian horses. Journal of animal science 82: 982-986.

Simpson, G. G., R. C. Lewontin, and A. Roe. 1960. Quantitative zoology. Harcourt, Brace & Co, United States.

Smith, L. A., B. G. Cassell, and R. E. Pearson. 1998. The effects of inbreeding on the lifetime performance of dairy cattle. Journal of Dairy Science 81: 2729-2737.

Sobczuk, D., and M. Komosa. 2012. Morphological Differentiation of Polish Arabian Horses - Multivariate Analysis. Bulletin of the Veterinary Institute in Pulawy 56: 623-629.

Solé, M., R. Santos, M. D. Gómez, A. M. Galisteo, and M. Valera. 2013. Evaluation of conformation against traits associated with dressage ability in unridden Iberian horses at the trot. Research in Veterinary Science 95: 660-666.

StatSoft. 2013. Statistica Version 12.

Swiger, L. A., K. E. Gregory, R. M. Koch, and V. A. Arthaud. 1961. Effect of Inbreeding on Performance Traits of Beef Cattle. Journal of animal science 20: 626-630.

Vicente, A. A., N. Carolino, and L. T. Gama. 2012. Genetic diversity in the Lusitano horse breed assessed by pedigree analysis. Livestock Science 148: 16-25.

Vicente, A. A., N. Carolino, J. Ralão-Duarte, and L. T. Gama. 2014a. Selection for morphology, gaits and functional traits in Lusitano horses: I. Genetic parameter estimates. Livestock Science 164: 1-12.

Vicente, A. A., N. Carolino, J. Ralão-Duarte, and L. T. Gama. 2014b. Selection for morphology, gaits and functional traits in Lusitano horses: II. Fixed effects, genetic trends and selection in retrospect. Livestock Science 164: 13-25.

Young, C. W. 1984. Inbreeding and the gene pool. Journal of Dairy Science 67: 472-477.

Zechner, P. et al. 2001. Morphological description of the Lipizzan horse population. Livestock Production Science 69: 163-177.

Supplementary Table 1 - Quadratic regression results of inbreeding and age on body measurements. Linear (b1) and quadratic (b2) terms ± standard error (SE). Statistically significant results in bold (*p<0.05, **p<0.01, ***p<0.001).

		Inbreeding (%)				Age			
		b1 (SE)		b2 (SE)		b1 (SE)		b2 (SE)	
	Source	OLD	NEW	OLD	NEW	OLD	NEW	OLD	NEW
HEAD	HL	1.2492 (0.7072)	-0.3673 (0.3011)	-0.0169 (0.0105)	0.0045 (0.0037)	5.6368 (4.4274)	0.1467 (0.2826)	-0.3404 (0.3028)	-0.0077 (0.0112)
	HW	0.3319 (0.2947)	-0.1343 (0.2856)	-0.0051 (0.0044)	0.0014 (0.0035)	0.6600 (2.0123)	0.1644 (0.2651)	-0.0527 (0.1376)	-0.0051 (0.0105)
	HT	0.6960 (0.3806)	-0.2458 (0.2933)	-0.0104 (0.0057)	0.0036 (0.0036)	2.4359 (2.3398)	0.1252 (0.2825)	-0.1868 (0.1600)	-0.0060 (0.0112)
	DE	0.0166 (0.4899)	-0.2223 (0.1280)	-0.0005 (0.0073)	0.0024 (0.0016)	-1.3362 (0.9105)	0.2070 (0.1223)	0.1253 (0.0623)	-0.0070 (0.0048)
	LD	-0.3595 (0.6399)	0.0179 (0.1732)	0.0049 (0.0095)	-0.0001 (0.0021)	-2.0342 (1.8256)	-0.0127 (0.1620)	0.1819 (0.1248)	0.0006 (0.0064)
BODY	NL	1.3827 (1.5070)	-0.3260 (0.4430)	-0.0189 (0.0224)	0.0038 (0.0054)	**11.5376 (3.0864)***	0.5869 (0.3989)	**-0.8895 (0.2110)****	-0.0250 (0.0158)
	WH	1.4790 (0.9128)	-0.5413 (0.6635)	-0.0187 (0.0136)	0.0062 (0.0082)	**16.9414 (4.7269)***	0.0397 (0.6023)	**-1.1475 (0.3232)***	-0.0097 (0.0238)
	BH	0.4131 (1.0187)	-0.5191 (0.6039)	-0.043 (0.0151)	0.0063 (0.00742)	**14.2260 (3.4932)****	0.1508 (0.5407)	**-0.9632 (0.2389)****	-0.0142 (0.0214)
	CrH	1.2742 (0.9619)	-0.2094 (0.5879)	-0.0177 (0.0143)	0.0017 (0.0072)	5.5711 (6.0384)	0.0690 (0.5476)	-0.3396 (0.4129)	-0.0076 (0.0217)
	TH	**3.8339 (1.2915)***	-0.4171 (0.6382)	-0.0524 (0.0189)	0.0037 (0.0078)	11.4156 (12.4173)	-0.0895 (0.6170)	-0.8455 (0.8510)	0.0055 (0.0244)
	BL	**1.9738 (0.6623)***	-0.9617 (0.7748)	**-0.0272 (0.0098)***	0.0110 (0.0095)	8.2278 (6.0853)	0.5036 (0.7387)	-0.5250 (0.4161)	-0.0191 (0.0292)
	TW	2.4434 (2.8562)	-0.6773 (0.8559)	-0.0386 (0.0425)	0.0037 (0.0105)	-2.2220 (18.9048)	0.7686 (0.9457)	0.2343 (1.3149)	-0.0250 (0.0374)
	ChH	0.6524 (0.5458)	-0.4982 (0.4597)	-0.0086 (0.0081)	0.0053 (0.0056)	**6.5359 (2.2086)***	0.0485 (0.4128)	**-0.4732 (0.1510)***	-0.0098 (0.0163)
	ChW	1.4067 (2.2539)	-0.4856 (0.5577)	-0.0211 (0.0335)	0.0053 (0.0068)	-2.0800 (8.2428)	0.2127 (0.5257)	0.2857 (0.5637)	-0.0113 (0.0208)
	ChC	4.6663 (3.7150)	-2.1258 (1.1720)	-0.0700 (0.0552)	0.0214 (0.0144)	8.2799 (24.2634)	0.1194 (1.1512)	-0.4268 (1.6592)	-0.0293 (0.0456)
	LChH	0.8265 (0.7111)	-0.0431 (0.3500)	-0.0102 (0.0106)	0.0009 (0.0043)	**10.4055 (2.6956)***	-0.0088 (0.3282)	**-0.6743 (0.1843)***	0.0001 (0.0130)
	CrW	4.1733 (4.2883)	-0.5950 (0.3060)	-0.0701 (0.0773)	0.0077 (0.0038)	15.8541 (5.7818)	-0.0837 (0.3056)	-1.0464 (0.3920)	0.0026 (0.0121)
	CrL	0.7169 (1.0353)	-0.6309 (0.4907)	-0.0113 (0.0154)	0.0075 (0.0060)	2.7034 (6.8678)	-0.5288 (0.4011)	-0.1739 (0.4696)	0.0096 (0.0159)
LEGS	SL	1.4521 (1.1568)	-0.3507 (0.4325)	-0.0203 (0.0172)	0.0042 (0.0053)	6.3000 (6.9748)	0.0578 (0.3986)	-0.3792 (0.4769)	-0.0061 (0.0158)
	EH	1.0281 (0.4493)	0.1143 (0.3510)	-0.0128 (0.4493)	-0.0013 (0.0043)	**10.5002 (3.3886)***	-0.0173 (0.3232)	**-0.7144 (0.2317)***	-0.0018 (0.0128)
	FoP	0.4168 (0.8559)	0.1053 (0.3456)	-0.0065 (0.0127)	-0.0014 (0.0042)	-1.3071 (4.9035)	0.0501 (0.3136)	0.1203 (0.3353)	-0.0054 (0.0124)
	KP	0.1993 (0.3411)	-0.3634 (0.2067)	-0.0022 (0.0051)	0.0044 (0.0025)	3.9174 (1.7086)	0.2596 (0.1963)	-0.2575 (0.1168)	-0.0104 (0.0078)
	CP	-0.0388 (0.5315)	-0.0857 (0.1245)	0.0009 (0.0079)	0.0009 (0.0015)	2.5778 (2.9777)	0.1597 (0.1140)	-0.1612 (0.2036)	-0.0057 (0.0045)
	FeP	0.1621 (0.4741)	0.0175 (0.1736)	-0.0018 (0.0070)	-0.0004 (0.0021)	4.3524 (2.4525)	-0.0108 (0.1576)	-0.2871 (0.1677)	0.0024 (0.0062)
	HH	0.5832 (0.4718)	0.0107 (0.3110)	-0.0077 (0.0070)	-0.0011 (0.0038)	3.2880 (2.9572)	-0.2944 (0.2921)	-0.2029 (0.2022)	0.0069 (0.0116)
	HP	0.1412 (0.3194)	-0.2443 (0.2682)	-0.0009 (0.0047)	0.0028 (0.0033)	3.9144 (2.1063)	-0.1630 (0.2354)	-0.2562 (0.1440)	0.0020 (0.0093)

Supplementary Table 2 - Coefficient of variation (CV) ranges found in different horse breeds. In brackets are the body measurement abbreviations.

Breed	CV (body measurement)	Reference
Sorraia	1.5 (CrH) - 13.8 (TW)	This study
Lusitano	2.02 (DE) - 7.06 (CrH)	Oom and Ferreira (1987)
Lusitano	1.51 (EH) - 8.06 (CrW)	Oom (1992)
Pura Raza Española	1.92 (TH) - 5.21 (TW)	
Berber	2.72 (WH) - 6.92 (TW)	
Sorraia	1.55 (CrH) - 13.32 (TW)	
Garrano	1.90 (CrH) - 10.90 (ChW)	
Arab	1.56 (CrH) - 4.30 (DE, TW)	
Garrano	4 (BH) - 10 (NL)	Portas (2001)
Spanish Purebred	2.60 (WH) - 7.80 (ChW)	Gómez et al. (2009a)
Andalusian	2.63 (WH) - 6.99 (ChW, KP)	Gómez et al. (2009b)
Spanish Arab	2.07 (WH) - 10.09 (ChW)	Cervantes et al. (2009)
Andalusian	2.31 (WH) - 5.12 (KP)	Molina et al. (1999)

Cervantes, I. et al. 2009. Size and shape analysis of morphofunctional traits in the Spanish Arab horse. Livestock Science 125: 43-49.

Gómez, M. D., F. Goyache, A. Molina, and M. Valera. 2009a. Sire × stud interaction for body measurement traits in Spanish Purebred horses. Journal of animal science 87: 2502-2509.

Gómez, M. D., M. Valera, A. Molina, J. P. Gutiérrez, and F. Goyache. 2009b. Assessment of inbreeding depression for body measurements in Spanish Purebred (Andalusian) horses. Livestock Science 122: 149-155.

Molina, A., M. Valera, R. Dos Santos, and A. Rodero. 1999. Genetic parameters of morphofunctional traits in Andalusian horse. Livestock Production Science 60: 295-303.

Oom, M., and J. Ferreira. 1987. Estudo biométrico do cavalo Alter. Rev Port Cienc Vet 83: 101-148.

Oom, M. M. 1992. O cavalo Lusitano. Uma raça em Recuperação (in Portuguese). PhD Dissertation, University of Lisbon, Portugal.

Portas, M. C. P. 2001. La Conservación de la raza equina Garrana. Archivos de Zootecnia 50.

CHAPTER 2 - PAPER 3

Genome-wide variability in the Sorraia horse.

Kjöllerström H.J., Raudsepp, T., Khanshour, A., Gosh, S., Oom, M.M. & Chowdhary, B.P.

(In preparation for PLoS One)

Genome-wide variability in the Sorraia horse

Kjöllerström, H.J.[1,2,*], Raudsepp, T.[2], Khanshour, A.[2], Gosh, S.[2], Oom, M.M.[1] & Chowdhary, B.P.[3]

[1]cE3c – Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal

[2]Department of Veterinary Integrative Biosciences, College of Veterinary Medicine & Biomedical Sciences, Texas A&M University, College Station, USA

[3]New Research Complex, Qatar University, Al Tarfa, Doha, Qatar

*Corresponding author: Tel.: +351 217500000; Fax: +351 217500028; E-mail: josefina.kjollerstrom@gmail.com

Abstract

The Sorraia horse is a Portuguese autochthonous breed and one of the most endangered horse breeds in the world. Due to reduced population size, small number of founders, genetic isolation and breeding practices, Sorraias have low genetic variability and extremely high inbreeding. Estimation of genetic variability is based on the analysis of markers that tend to differ between individuals. In this study, we used STRs, SNPs and CNVs to evaluate genetic variation in the Sorraia horse with the aim to improve the management of this highly endangered horse breed and to identify new efficient markers for parentage testing.

Thanks to systematic selection of informative STRs in Sorraias, the higher number of STRs used, and more animals genotyped, our results for autosomal STRs represent an overall improvement of analysed parameters over previously published data. With the exception of some Arabian horse populations, the Sorraia has the lowest described average number of alleles among the horse breeds studied so far. Overall, autosomal STRs results were better in Germany than in Portugal, showing that using one stallion per year per mare is a better management choice than using only one stallion per year per herd. Average inbreeding in Sorraias was higher than in other breeds. Heterozygosity of STRs was significantly higher than of SNPs. Runs of homozygosity were found in 11 chromosomes, with chr2, 11, 22 and 27 sharing some common regions between different horses. STRUCTURE analysis separated Sorraias into 3 clusters, while FCA (STRs) and PCA (SNPs) analyses separated our total population into two main groups based on the breed's historic evolution and connections between stud farms: Portugal and Germany. It would be important to mix these two populations again, potentially by bringing some selected stallions for extensive breeding in Portuguese stud

farms, in order to homogenise the genetic background and, in the short-term, increase genetic variability.

We found 213 CNVs, arranged into 71 CNVRs in Sorraias, 7 of which were Sorraia-specific.

To the best of our knowledge, this is the first time that a composite STR, SNP and CNV study has been done in a single horse breed. It provides crucial novel and genome-wide variability data useful for the improvement of the breed's genetic health and permanent loss prevention of this iconic and important animal genetic resource.

Keywords: Sorraia Horse; Genetic variability; STRs; SNPs; CNVs

Introduction

The Sorraia horse is a Portuguese autochthonous breed and one of the most endangered horse breeds in the world. It represents a primitive equine type that has inhabited the Iberian Peninsula since early Pleistocene and is related to other Iberian, as well as New World horse breeds [1]. Since foundation in 1937 the breed has been managed as a closed population [2]. In 1975 new breeders appeared in Portugal and abroad. German breeders, however, were selling animals to other German breeders with no further introduction of horses from Portugal [2]. Portuguese Sorraias were re-introduced to Germany only in year 2000 and currently there are 10 breeders in Portugal and eight in Germany [3]. As a consequence of reduced population size, small number of founders, genetic isolation and breeding practices, Sorraia horses have low genetic variability and extremely high rate of inbreeding, as revealed by molecular [4-10] and pedigree data [11,12] analysis. It has been shown that inbreeding depression significantly affects mare fertility, and negatively influences foaling intervals and stallion fertility [12].

Population genetic variability and viability analysis are important in conservation management programs [13] of endangered populations to evaluate health and survival capacity (e.g. [14]). Estimation of genetic variability is based on the analysis of markers that tend to differ between individuals. Overall, genomes within species vary in multifarious ways but the most commonly used markers for the study of intraspecific variation include polymorphic microsatellites or short tandem repeats (STRs), single nucleotide polymorphism (SNPs) and, more recently, copy number variations (CNVs). Of these, STRs have been extensively used in parentage testing, genetic diversity, population structure, conservation, evolutionary and domestication studies of different species [14-18] including horses [10,19-24]. Sampling can be non-invasive and PCR amplification can be done

from small amounts of DNA making STRs particularly useful in conservation genetic studies of endangered species [14].

Single nucleotide polymorphisms (SNPs) are highly informative (highly variable) markers [25] that have been used in livestock species to study genetic diversity, disease susceptibility or genes associated with phenotypic traits of interest [25]. The Affymetrix Axiom® Equine 670K Genotyping Array (Axiom MNEC670) and the Illumina 70K-Equine SNP chip have recently become commercially available, although most published studies so far have been done with the 50K-Equine SNP chip [26] and used for genome-wide association studies (GWAS) of equine disorders [27-29], to resolve breed relationships and study genetic diversity [30,31], or even screening for chromosomal abnormalities [32].

Copy number variants (CNVs) are considered important sources of genetic variability and have been used in livestock to study, for example, disease susceptibility, developmental disorders and morphological traits [25]. CNVs have also been described in the horse genome, in normal populations [33-36], as well as in relation to equine disorders/diseases or traits, such as melanomas [37], recurrent laryngeal neuropathy [38], susceptibility to *Rhodococcus equi* [39], disorders of sexual development [40], recurrent airway obstruction [41], and body size [42].

In this study, we will use STRs, SNPs and CNVs to evaluate genetic variation in the Sorraia horse with the aim to improve the genetic management of this highly endangered horse breed and to identify new efficient markers for parentage testing.

Materials and Methods

DNA isolation

Genomic DNA was extracted from blood and hair samples received at the Applied Genetics Laboratory (cE3c, Lisbon, Portugal) for parentage testing (n=190) following standard [43] and manufacturer's (Qiagen) protocols, with some modifications. Samples were received from 10 different breeders in Portugal (PT) (n=148) and 3 breeders in Germany (GER) (n=42).

Short tandem repeats (STRs)

We selected 150 highly polymorphic genome-wide equine STRs, as shown in Supplementary Table 1. The STRs were tested on pooled DNA of 4 most heterozygous Sorraia horses as previously revealed at the Applied Genetics Laboratory (cE3c, Lisbon, Portugal). Markers with three or more alleles (Supplementary Figure 1) and 21 STRs currently used for parentage testing in Sorraias (Luis et al. 2007a) (Supplementary Table 2) were selected for further analysis. These included 50 autosomal, 5 X-linked and 2 Y-linked STRs (Supplementary Table 2).

Table 1 – Summary statistics for 50 autosomal STRs studied in 190 Sorraia horses: chromosomal location (Chr), mean number of alleles (NA), number of animals (n), observed heterozygosity (H_O), unbiased expected heterozygosity (H_E), polymorphic information content (PIC), probability of paternity exclusion (PE), heterozygote deficiency coefficient (F_{IS}). [a] STRs used at the Applied Genetics Laboratory (cE3c, Lisbon). HWE - significance of deviation from Hardy-Weinberg equilibrium (with Bonferroni correction): * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

		TOTAL						PORTUGAL						GERMANY					
Locus	Chr	NA	H_O	H_E	PIC	PE	HWE	NA	n	H_O	H_E	F_{IS}	HWE	NA	n	H_O	H_E	F_{IS}	HWE
AHT58	1	3	0.593	0.626	0.553	0.490	NS	3	147	0.578	0.611	0.054	NS	3	42	0.643	0.661	0.027	NS
NVHEQ100	1	4	0.505	0.648	0.574	0.516	***	4	145	0.517	0.568	0.090	NS	3	41	0.463	0.513	0.098	ND
HMS007[a]	1	3	0.278	0.300	0.276	0.260	NS	3	134	0.194	0.190	-0.022	ND	3	42	0.548	0.561	0.024	NS
LEX020[a]	1	3	0.544	0.519	0.432	0.369	NS	3	120	0.550	0.533	-0.032	NS	3	40	0.525	0.477	-0.102	NS
COR065	2	4	0.615	0.612	0.531	0.481	NS	4	133	0.617	0.605	-0.020	NS	4	36	0.611	0.629	0.028	NS
TKY384	2	5	0.656	0.696	0.648	0.641	NS	5	147	0.626	0.685	0.087	NS	5	42	0.762	0.693	-0.101	NS
UMNE158	3	3	0.588	0.523	0.465	0.425	NS	3	146	0.575	0.514	-0.120	NS	3	41	0.634	0.542	-0.173	NS
COR089	4	4	0.692	0.721	0.666	0.634	NS	4	121	0.678	0.722	0.062	NS	4	35	0.743	0.725	-0.025	ND
HMS006[a]	4	4	0.618	0.667	0.594	0.532	NS	4	131	0.618	0.615	-0.005	NS	4	42	0.619	0.555	-0.117	NS
AHT107	5	6	0.408	0.668	0.611	0.600	***	5	101	0.396	0.675	0.415	***	4	19	0.474	0.626	0.248	ND
TKY412	6	4	0.654	0.708	0.654	0.629	NS	4	144	0.674	0.657	-0.025	NS	4	41	0.585	0.621	0.058	NS
TKY1001	6	3	0.654	0.652	0.575	0.505	NS	3	125	0.656	0.662	0.009	NS	3	31	0.645	0.565	-0.144	NS
TKY034	7	3	0.534	0.493	0.436	0.397	NS	3	147	0.578	0.523	-0.106	NS	3	42	0.381	0.326	-0.171	ND
COR003	8	3	0.563	0.555	0.494	0.449	NS	3	148	0.588	0.569	-0.033	NS	3	42	0.476	0.478	0.003	ND
COR012	8	4	0.659	0.707	0.648	0.609	NS	4	141	0.645	0.673	0.041	NS	4	41	0.707	0.737	0.041	NS
AHT005[a]	8	5	0.535	0.531	0.476	0.455	NS	5	118	0.449	0.452	0.006	NS	5	39	0.795	0.677	-0.177	NS
LEX023[a]	8	3	0.563	0.643	0.567	0.500	NS	3	122	0.549	0.622	0.118	NS	3	38	0.605	0.655	0.078	NS
HMS003[a]	9	3	0.190	0.214	0.193	0.170	ND	3	133	0.113	0.127	0.113	ND	3	41	0.439	0.426	-0.031	ND
HTG004[a]	9	4	0.727	0.699	0.642	0.612	NS	4	134	0.731	0.660	-0.108	NS	4	42	0.714	0.702	-0.017	NS
ASB009	10	3	0.112	0.108	0.105	0.103	ND	3	145	0.097	0.093	-0.034	ND	2	42	0.167	0.155	-0.079	ND
ABGe099	11	5	0.774	0.716	0.675	0.676	ND	4	148	0.743	0.662	-0.124	NS	5	42	0.881	0.787	-0.121	ND
SGCV24	11	4	0.599	0.552	0.497	0.466	NS	4	139	0.540	0.510	-0.058	NS	4	38	0.816	0.664	-0.232	NS
COR058	12	3	0.537	0.574	0.504	0.451	NS	3	147	0.524	0.514	-0.019	NS	3	41	0.585	0.661	0.116	NS
ASB002[a]	15	5	0.651	0.617	0.557	0.528	NS	4	133	0.647	0.582	-0.111	NS	4	42	0.667	0.682	0.023	NS
HTG006[a]	15	4	0.639	0.658	0.593	0.550	NS	4	120	0.642	0.655	0.021	NS	4	38	0.632	0.663	0.048	NS
COR007	17	4	0.604	0.674	0.625	0.608	NS	4	146	0.596	0.646	0.078	NS	4	41	0.634	0.643	0.015	NS
COR105	17	3	0.527	0.613	0.535	0.472	NS	3	147	0.497	0.566	0.124	NS	3	41	0.634	0.639	0.007	NS
ABGe151	18	3	0.625	0.579	0.496	0.436	NS	3	142	0.683	0.563	-0.215	*	3	42	0.429	0.605	0.294	NS
TKY741	18	4	0.633	0.680	0.614	0.566	NS	4	148	0.622	0.670	0.073	NS	4	40	0.675	0.675	0.000	NS
COR062	19	4	0.718	0.724	0.672	0.645	NS	4	129	0.698	0.727	0.040	NS	4	34	0.794	0.724	-0.099	ND
TKY448	19	3	0.624	0.625	0.551	0.489	NS	3	144	0.646	0.640	-0.009	NS	3	42	0.548	0.529	-0.036	NS
LEX036[a]	19	3	0.669	0.640	0.564	0.498	NS	3	120	0.683	0.638	-0.071	NS	3	40	0.625	0.623	-0.004	NS
TKY321	20	4	0.778	0.749	0.699	0.673	NS	4	147	0.769	0.749	-0.027	NS	4	42	0.810	0.744	-0.089	ND
TKY477	20	5	0.612	0.643	0.575	0.525	NS	5	147	0.612	0.625	0.021	NS	4	41	0.610	0.595	-0.026	NS
UM011	20	3	0.668	0.614	0.543	0.485	NS	3	148	0.689	0.608	-0.134	NS	3	42	0.595	0.618	0.037	NS
COR073	21	3	0.435	0.457	0.396	0.353	NS	3	143	0.455	0.456	0.003	NS	3	41	0.366	0.448	0.186	ND
TKY623	21	3	0.624	0.572	0.494	0.437	NS	3	146	0.610	0.580	-0.051	NS	2	40	0.675	0.503	-0.347	NS
TKY806	21	3	0.604	0.609	0.532	0.471	NS	3	147	0.605	0.627	0.034	NS	3	40	0.600	0.493	-0.221	ND
HTG010[a]	21	2	0.314	0.391	0.314	0.244	NS	2	133	0.368	0.406	0.094	NS	2	42	0.143	0.341	0.584	ND
ABGe121	22	3	0.667	0.628	0.554	0.492	NS	3	85	0.671	0.630	-0.065	NS	3	35	0.657	0.613	-0.073	NS
TKY568	23	6	0.785	0.740	0.695	0.693	NS	6	145	0.800	0.739	-0.083	NS	6	41	0.732	0.677	-0.082	NS
AHT004[a]	24	5	0.626	0.644	0.590	0.566	NS	5	129	0.636	0.654	0.029	NS	4	42	0.595	0.592	-0.005	NS
NVHEQ43	25	5	0.709	0.729	0.676	0.651	NS	5	147	0.707	0.730	0.031	NS	4	42	0.714	0.682	-0.047	NS
UCDEQ405[a]	25	3	0.513	0.519	0.458	0.415	NS	3	120	0.567	0.556	-0.019	NS	3	38	0.342	0.359	0.048	ND
TKY523	26	4	0.585	0.653	0.591	0.556	NS	4	143	0.587	0.629	0.066	NS	4	40	0.575	0.575	0.001	NS
TKY315	27	3	0.674	0.626	0.550	0.487	NS	3	141	0.688	0.641	-0.074	NS	3	40	0.625	0.561	-0.116	NS
UCDEQ005[a]	27	4	0.510	0.548	0.481	0.446	NS	4	118	0.585	0.583	-0.003	NS	3	27	0.185	0.338	0.456	ND
TKY478	29	3	0.132	0.139	0.133	0.129	NS	3	148	0.101	0.116	0.130	ND	2	42	0.238	0.212	-0.123	ND
VHL020[a]	30	5	0.670	0.650	0.612	0.623	ND	5	135	0.681	0.665	-0.025	NS	5	41	0.634	0.563	-0.129	NS
ABGe241	31	3	0.432	0.440	0.393	0.361	NS	3	143	0.434	0.440	0.015	NS	3	42	0.429	0.444	0.035	ND
Mean		**3.7**	**0.573**	**0.586**	**0.526**	**1**		**3.7**	–	**0.570**	**0.571**	**0.003**		**3.5**	–	**0.580**	**0.572**	**-0.009**	

Table 2 – Summary statistics for the 5 X-linked STRs studied in 98 Sorraia horses: mean number of alleles (NA), number of animals (n), observed heterozygosity (H_o), unbiased expected heterozygosity (H_e), polymorphic information content (PIC), probability of paternity exclusion (PE), heterozygote deficiency coefficient (F_{IS}). [a] STRs used at the Applied Genetics Laboratory (cE3c, Lisbon). HWE - significance of deviation from Hardy-Weinberg equilibrium (with Bonferroni correction): * p<0.05; ** p<0.01; *** p<0.001.

	TOTAL						PORTUGAL						GERMANY					
Locus	NA	H_O	H_E	PIC	PE	HWE	NA	n	H_O	H_E	F_{IS}	HWE	NA	n	H_O	H_E	F_{IS}	HWE
LEX003[a]	3	0.340	0.435	0.384	0.350	NS	3	77	0.403	0.480	0.162	NS	3	20	0.100	0.188	0.476	ND
LEX024[a]	2	0.206	0.233	0.205	0.174	ND	2	76	0.158	0.146	-0.079	ND	2	21	0.381	0.455	0.167	ND
LEX027	3	0.589	0.662	0.584	0.513	NS	3	74	0.581	0.654	0.112	NS	3	21	0.619	0.650	0.049	ND
TKY038	4	0.418	0.532	0.477	0.447	*	3	77	0.403	0.426	0.056	NS	4	21	0.476	0.733	0.356	ND
UCDEQ502[a]	2	0.105	0.253	0.220	0.184	ND	2	74	0.122	0.205	0.408	ND	2	21	0.048	0.396	0.882	ND
Mean	**2.8**	**0.332**	**0.423**	**0.374**	**0.882**		**2.6**	–	**0.333**	**0.382**	**0.132**		**2.8**	–	**0.325**	**0.484**	**0.386**	

All DNA samples (n=190) were used for the analysis of autosomal STRs, while X-linked markers were analysed only in females (n=98) and Y-linked markers only in males (n=92). Genotyping of the 21 parentage testing and X- and Y-linked STRs was performed following standard protocols [9] and resolved in a Li-Cor 4200S sequencer (Li-Cor, Lincoln, NE) with sizing determined in RFLP scan 3.1 software (Scanalytics CPS Inc., Rockville, MD). The remaining 36 STRs were genotyped using the three primer method [44] and resolved on an ABI 3730 DNA analyser (Applied Biosystems). Sizing was determined using GeneScan-500 LIZ Size Standard and GeneMapper® v4.1 (Applied Biosystems). Each 10µl PCR reaction contained 20ng genomic DNA, 1x Buffer, 1.5mM $MgCl_2$, 0.2mM dNTPs, 0.04U AmpliTaq Gold® DNA Polymerase (Applied Biosystems), 0.02µM forward primer, 0.3µM reverse primer and 0.3µM fluorescently labelled M13 primer. Amplification conditions for all STRs were as follows: 10min at 95ºC; 35 cycles of 30s at 95ºC, 30s at 58ºC, 30s at 72ºC; and a final extension of 10min at 72ºC.

Mean number of alleles (MNA), observed heterozygosity (H_O), expected heterozygosity (H_E), polymorphic information content (PIC), probability of paternity exclusion (PE) and deviations from Hardy-Weinberg equilibrium (HWE) with Bonferroni correction were calculated for all horses and separately for the Portuguese (PT) and German (GER) populations using CERVUS 3.0.3 [45] and GENEPOP 4.0.10 [46,47]. Wright's fixation coefficients [48] (F_{ST} and F_{IS}) were estimated according to [49] using GENEPOP 4.0.10 [46,47]. Inbreeding coefficients were calculated back to the founders in SPARKS v.1.6 [50] based on the additive relationship matrix [51]. Individual heterozygosity and mean d^2 were calculated as described in [9] for the total, Portuguese (PT) and German (GER) populations. All statistical analyses were performed using Statistica v12 [52]. Descriptive statistics and Pearson's correlations between inbreeding, d^2 and individual heterozygosity were calculated in

the total, Portuguese (PT) and German (GER) populations. Factorial correspondence analysis (FCA) to determine subpopulation relatedness was performed using GENETIX software [53]. STRUCTURE software package [54,55] was used to determine substructure in the breed using admixture and correlated allele frequencies, 10 independent runs for each K value (K=1 to 13), 50.000 burn-in period and 100.000 MCMC. Structure Harvester website and program were used to calculate and plot the best number of clusters in our data following the Evanno Delta K (ΔK) value model [56,57].

Single Nucleotide Polymorphisms (SNPs)

The most heterozygous animals for STRs (n=50) were genotyped on the Illumina® Equine SNP70 beadchip by GeneSeek®. Only autosomal SNPs were used for analysis. In order to ensure genotyping quality the following filters were applied: 0.1 maximum per-sample missing data rate (MIND), 0.01 minor allele frequencies (MAF), 0.1 maximum per-SNP missing rate (GENO) and 0.01 Hardy-Weinberg equilibrium p-value (HWE). Mean minor allele frequency (MAF) and missingness (MISSING) rate were calculated before and after data clean-up. Heterozygosity (Het), genome-wide linkage disequilibrium (LD), inbreeding coefficients (based on loss of heterozygosity, -het), runs of homozygosity (ROH) were calculated using PLINK [58,59] in cleaned data. Statistical analyses were done using Statistica v12 (StatSoft 2013). Principal component analysis (PCA) to determine population relatedness on cleaned data was performed using PLINK [58,59] with linkage disequilibrium based SNP pruning (INDEP PAIRWISE) of 50 5 0.2 (window size, number of SNPs to shift the window at each step and multiple correlation coefficient, respectively). Descriptive statistics of inbreeding calculated from pedigree or SNP data and SNP individual heterozygosity were calculated in the total, PT and GER populations. Using a Student's T-test for Independent Samples, we investigated the concordance between heterozygosities calculated with STRs and SNPs. Pearson's correlations were calculated between inbreeding by pedigree and by SNPs; inbreeding by pedigree and SNP individual heterozygosity; heterozygosity by STRs and by SNPs; and inbreeding by SNPs and year of birth.

Copy Number Variations (CNVs)

Seven males and one female from the original sample set (n=190) were randomly chosen for CNV analysis by array comparative genomic hybridization (aCGH) on the Texas-Adelaide horse 400K whole genome tiling oligoarray [35]. In addition, CNV data for one male (M#5) and one female (F) Sorraia were available from a prior study [35]. Probe labelling and aCGH followed Agilent

Technologies Protocol Version 7.3, March 2014 (http://www.agilent.com/cs/library/usermanuals/Public/G4410-90010_CGH_Enzymatic_7.4.pdf) described in detail in [35]. All hybridizations included two differently labelled probes: Cy3-labeled reference DNA (Thoroughbred) and Cy5-labeled Sorraia DNA. Slides were scanned on an Agilent SureScan DNA Microarray Scanner (AGILENT SCANNER CONTROL software v8.3) and data were extracted and normalized in AGILENT FEATURE EXTRACTION software v11.0.1.1. AGILENT GENOMIC WORKBENCH 5.0 software was used to discover CNV regions as described in [41]. Homozygous gains and losses were called when the $\log_2$ ratio was >2.0 and <-2.0, respectively. Our results were compared to previously published horse CNVs [33-36,38,40-42]. Sorraia-specific CNVs were further confirmed by qualitative and quantitative PCR, following [35], using the animals studied by aCGH (eight males and one female) and additional five females. Primers were designed using EquCab2 (http://genome.ucsc.edu/), Ensembl (http://www.ensembl.org/index.html) and PRIMER3 software [60]. The qPCR in a LightCycler® 480 (Roche Diagnostics) followed [35], with 20µl reactions of 50ng DNA, 10µM primers and HOT FIREPol® EvaGreen® qPCR Mix Plus (Solis BioDyne). Relative copy number variations were determined in comparison to the reference sample and normalized to an autosomal reference gene, GAPDH. Gene content was analysed using Ensembl Genebuild 77.2 (http://useast.ensembl.org/Equus_caballus/Info/Index) as well as USCS (http://genome.ucsc.edu/) and Ensembl (http://www.ensembl.org/index.html) to look for mammalian orthologues.

Results

STR analysis

The total number of alleles for autosomal STRs was 186 with an average of 3.7 alleles per locus in the entire population, varying from 2 (*HTG010*) to 6 (*AHT107* and *TKY568*) alleles for individual markers (Table 1). The highest H_O was found in *ABGe099* (0.881) and lowest in *ASB009* (0.097). H_E ranged from 0.093 (*ASB009*) to 0.787 (*ABGe099*). PIC varied from 0.105 (*ASB009*) to 0.699 (*TKY321*). The highest PE was found in *TKY568* (0.693) and lowest in *ASB009* (0.103), while the combined PE for all loci was 1. In the total population, only *AHT107* and *NVHEQ100* significantly deviated from HWE at the 0.1% level (Table 1). The average F_{IS} was 0.003 in PT and 0.009 in GER (Table 1), and F_{ST} between populations was 0.0672.

A total of 14 alleles were found for the five X-linked STRs, with an average of 2.8 alleles per locus, ranging from 2 (*LEX024* and *UCDEQ502*) to 4 (*TKY038*) alleles for individual markers (Table 2). The highest H_O was found in *LEX027* (0.619) and lowest in *UCDEQ502* (0.048). H_E ranged from 0.733 (*TKY038*) in GER to 0.146 (*LEX024*) in PT. Maximum PIC was 0.584 (*LEX027*) and

minimum 0.205 (*LEX024*). PE for X-linked STRs was 0.882 and was the highest for *LEX027* (0.513) and lowest for *LEX024* (0.174). In the total population, only *TKY038* deviated significantly from HWE at the 5% level (Table 2). Average F_{IS} was 0.132 and 0.386 in PT and GER (Table 2), respectively, and F_{ST} between populations was 0.1152.

There was no polymorphism in the two Y-linked loci.

Average inbreeding, d^2 and individual heterozygosity in the total population were 0.3816 ± 0.066, 60.6449 ± 17.556 (bp^2) and 0.5708 ± 0.097, respectively (Fig. 1; Supplementary table 3). In PT, these values were 0.3858 ± 0.070, 58.7353 ± 17.549 (bp^2) and 0.5678 ± 0.101 (Fig. 1; Supplementary table 3), and in GER 0.3671 ± 0.043, 67.3742 ± 16.032 (bp^2) and 0.5812 ± 0.081 (Fig. 1; Supplementary table 3). Correlations between inbreeding and mean d^2 (r=-0.404), and between inbreeding and individual heterozygosity (r=-0.504), were both negative and statistically significant (p<0.0001). The correlation between individual heterozygosity and mean d^2 was positive (r=0.562) and also statistically significant (p<0.0001).

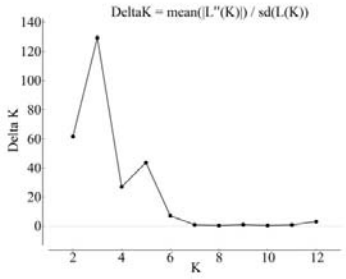


Figure 2 - Data plot of the ΔK distribution from the K values tested (K=1 to 13) considering 190 animals from 13 different breeders.

Table 3 - Breeder assignment to each of the three clusters at K=3. For each breeder (1 to 13), country (PT or GER) is indicated, as well as the cluster for which the assignment is highest (in bold).

		Cluster		
Breeder	Country	1	2	3
1	GER	0.019	0.068	**0.913**
2	PT	0.042	0.118	**0.840**
3	PT	0.035	0.032	**0.934**
4	PT	**0.557**	0.051	0.391
5	PT	0.038	0.250	**0.713**
6	PT	**0.419**	0.337	0.244
7	PT	0.319	0.135	**0.546**
8	PT	0.158	0.217	**0.625**
9	GER	0.026	**0.950**	0.023
10	GER	0.072	**0.883**	0.045
11	PT	0.226	0.146	**0.628**
12	PT	**0.491**	0.051	0.459
13	PT	0.046	**0.751**	0.203

The highest value for ΔK in STRUCTURE analysis was K=3 (Figs. 2, 3; Supplementary Figure 2). Stud farm assignment to each cluster is given in Table 3. Animals in Cluster 1 (stud farm 4, 6 and 12) are related to the second biggest stud farm in Portugal. Cluster number 2 contains animals from the first German (9 and 10) and a Portuguese stud farms (13). Cluster 3 contains stud farms related to the founder family (2, 3, 5 and 8), two Portuguese (7 and 11) and the most recent German stud farms (1), which all bought their horses from the founder stud farm.

SNP analysis

In raw SNP data, the total genotyping rate was 0.979 with MAF 0.137 and the proportion of missing SNPs in all animals was on average 0.021. After data pruning, these values were 0.991, 0.233 and 0.009, respectively, with a total of 35,037 SNPs. Average inbreeding by pedigree, inbreeding by SNPs and individual heterozygosity in the total population were 0.3393 ± 0.065, -0.0684 ± 0.084 and 0.3352 ± 0.026, respectively (Supplementary Table 4). In PT, these values were 0.3401 ± 0.070, -0.0569 ± 0.079 and 0.3316 ± 0.025 and in GER 0.3355 ± 0.334, -0.1210 ± 0.089 and 0.3516 ± 0.028 (Supplementary Table 4). Histograms of inbreeding by pedigree, inbreeding by SNPs and individual heterozygosity for the total, PT and GER populations are shown in Figure 4. The correlation between inbreeding by pedigree and by SNPs was positive (r=0.3352) and statistically significant ($p<0.05$). Inbreeding by pedigree and SNP heterozygosity were negatively correlated (r=-0.3345) and significant ($p<0.05$).

Comparison of STR and SNP analyses showed that Het by STRs was significantly higher ($p<0.05$) than Het by SNPs but the two were not significantly correlated. Correlation between inbreeding by SNPs and year of birth was positive and statistically significant in the total (r=0.3492, $p<0.05$) and PT (r=0.3335, $p<0.5$), but not in GER (r=0.3166, $p>0.05$). There were 47 ROH (Supplementary Table 5) with a 98.24% average proportion of homozygous sites within identified homozygous regions. Four ROHs were shared between more than one animal (Supplementary Figure 3). Average ROH size was 8,264,852 bp, with the longest on chromosome 15 (19,913,513 and 18,617,634 bp) and the shortest on chromosomes 2 and 20 (4,925,765 and 4,929,550 bp, respectively) (Supplementary Table 5). There were 172 SNPs on average per ROH, maximum 403 (chromosome 15) and minimum 105 (chromosome 2).

Factorial correspondence analysis (FCA) with autosomal STRs (N=50) separated our population into two main groups: PT and GER, with animals from the most recent German breeder clustering with the PT population (Figure 5A). The first axis explained 8.14% of the genetic variance, with the second and third explaining 5.06% and 4.05%, respectively. Likewise, PCA

analysis on SNP data revealed two main groups: PT and GER (Figure 5B) with animals from the recent German breeder clustering with Portuguese animals. The first axis explained 3.90% of the genetic variance, the second and third 2.89% and 2.58%, respectively.

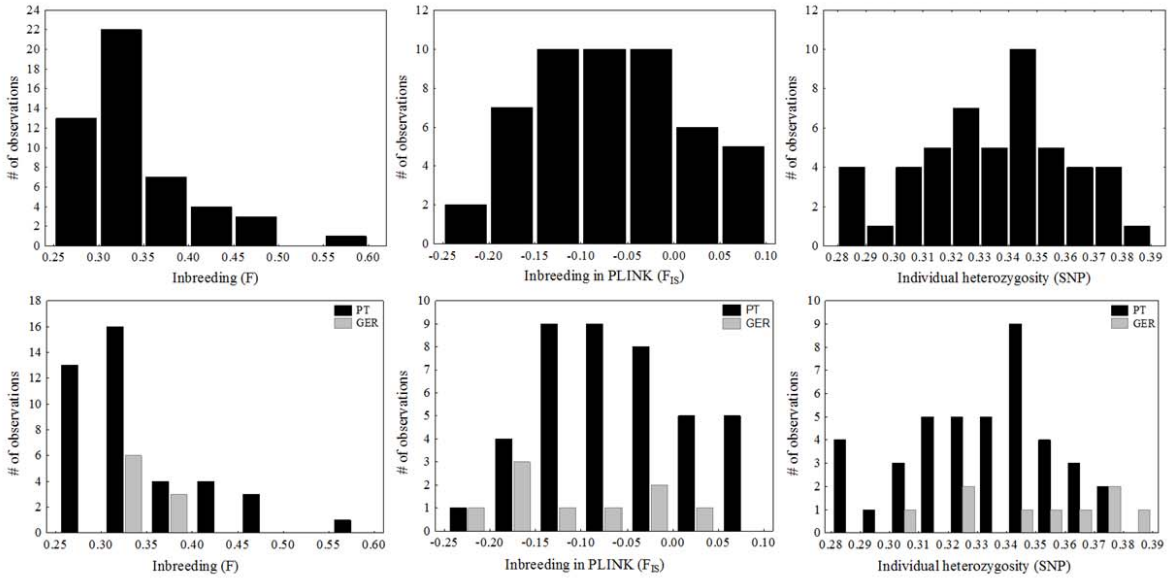


Figure 4 - Distribution of inbreeding by pedigree data (left), by SNP analysis (center), and individual heterozygosity in the whole population (top row) and in subpopulations (bottom row) based on the 50 animals sampled on Equine70KSNP chip.

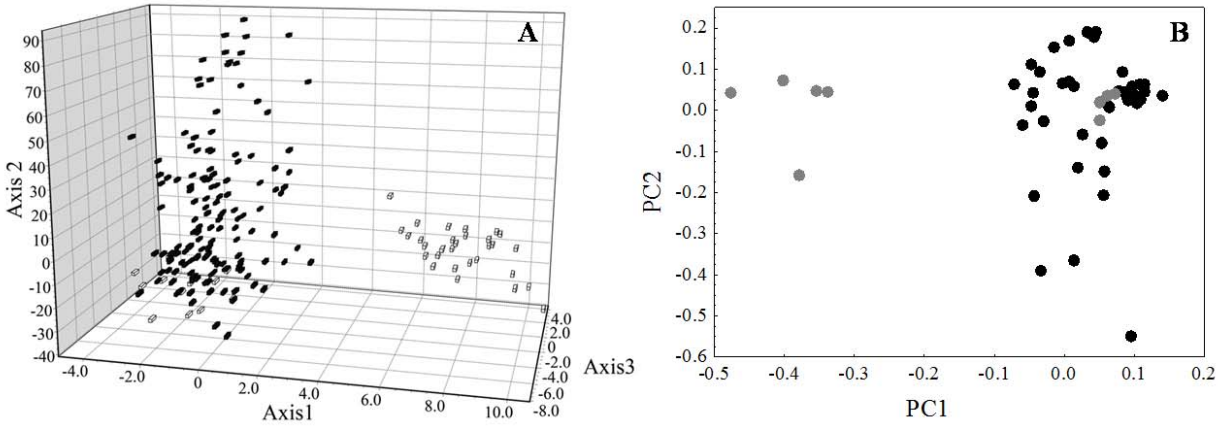


Figure 5 - (A) FCA on autosomal STRs (Portugal - black; Germany - grey)(n=190), and (B) PCA on SNPs (Portugal - black; Germany - grey)(n=50).

CNV analysis

Array CGH analysis revealed the presence of 213 CNVs: 53 gains and 160 losses, with an average 26.6 CNVs per horse (Table 4). Adjacent and overlapping CNV calls were arranged into 71 CNVRs. When compared with published results [33-36,38,40-42], 7 CNVRs (2 gains and 5 losses) were uniquely found in Sorraia horses. Of these, 4 CNVs were private (present in one individual), 2 were shared by two individuals, and one CNV was found in 4 animals (Table 4). Two novel CNVs involved genes: *ATP10D* in chr3:81,328,220-81,382,635 and *RFX3* in chr23:25,068,408-25,157,805.

Table 4 – Number of detected CNVs in the Sorraia horse: Sample identification (Sample), number of CNV calls (CNV calls), number of gains (Gains), number of losses (Losses), number of shared CNVs (Shared), number of CNVs unique to an animal (Unique)

Sample	CNV calls	Gains	Losses	Shared	Unique
M#1	31	12	19	29	2
M#2	25	6	19	24	1
M#3	29	4	25	29	0
M#4	30	10	20	30	0
M#5	36	8	28	35	1
M#6	26	10	16	26	0
M#7	18	2	16	18	0
F	18	1	17	18	0
Average	*26.6*	*6.6*	*20*	*26.1*	*0.5*

All novel CNVs were validated by qPCR. Validation by qPCR was also done for a gain in chr27. Though this CNV was not specific to the Sorraia, it was of interest due to the large size (230,312 bp; Table 5).

Table 5 –Sorraia-specific CNVs: chromosomal location (Chr), Start and Stop positions, CNV size in base pairs (bp), number of samples (# Samples) and CNV type (Type)

Chr	Start	Stop	Size (bp)	# Samples	Type
chr3	81,328,220	81,382,635	54,415	2	Loss
chr9	408,215	431,142	22,927	1	Loss
chr9	76,427,145	76,498,031	70,886	2	Loss
chr10	77,901,852	77,957,493	55,641	4	Loss
chr23	25,068,408	25,157,805	89,397	1	Loss
chr27	60,642	290,954	230,312	2	Gain
chrUn	1,521	1,589	68	1	Gain
chrUn	1,891	1,953	62	1	Gain

Discussion

To the best of our knowledge, this is the first time that a composite STR, SNP and CNV study has been done in a single horse breed, providing crucial novel and genome-wide variability data useful for the genetic management of this extremely endangered breed.

Thanks to systematic selection of informative STRs in Sorraias, the higher number of STRs used, and more animals genotyped, our results for autosomal STRs represent an overall improvement of analysed parameters over previously published data [9,61] (Supplementary Table 6), due to STR selection method based on higher variability in this breed and higher number of STRs and animals genotyped (Supplementary Table 6). Despite this, it is noteworthy that the average number of STR alleles in Sorraias is lower than in other Portuguese breeds such as the Lusitano, Garrano and Terceira ponies, other Iberian breeds like the Andalusian and the endangered Retuertas and Asturcón ponies (Supplementary Table 6), as well as the inbred Friesians, Thoroughbreds and Zanskari and Hucul ponies (Supplementary Table 6). With the exception of some Arabian horse populations [62], the Sorraia has the lowest described average number of alleles among the horse breeds studied so far. Likewise, even though the polymorphic information content (PIC) in this study was higher than previously described for the Sorraia [9], it remains lower than PIC in Hanoverian Warmblood and German cold-blood horses, Zanskari ponies, and Brazilian Criollo and Pataneiro horses (Supplementary Table 6).

Probability of exclusion (PE) increased slightly from that obtained by [9] due to high variability of the chosen STRs. However, it must be noted, that the number of loci used in this study was higher than in others breeds with comparable PE, showing the low variability of the Sorraia breed. Therefore, markers with higher variability (*ABGe099*, *NVHEQ43*, *TKY321* and *TKY568*; PE>0.65) will be of great interest in parentage testing and detection of false paternities in the Sorraia, potentially replacing the least variable or fixed STRs in the parentage testing panel currently in use.

Mean individual heterozygosity was lower than that found in Lipizzan horses but higher than found in this breed by [9] (Supplementary Table 6). Due to the way these markers were selected, H_O, H_E and mean d^2 were higher than previously described for this breed, showing that the markers used were, indeed, more variable than the ones routinely used in Sorraia parentage testing and standard genetic variability analysis (Supplementary Table 6). Comparison of H_O, H_E and F_{IS} of autosomal STRs between German (GER) and Portuguese (PT) Sorraia populations showed that these parameters were better in GER, while H_O and F_{IS} in X-linked STRs were better in PT. These results show that using one stallion per year per mare as in GER is a better management choice than using only one stallion per year per herd as in PT.

Regarding Y-linked STRs, our findings are consistent with prior studies showing that in contrast to autosomal STRs and mitochondrial DNA, the Y chromosome has no polymorphism in horses [63-65]. The only exceptions are the presence of 2 alleles of YA16 in indigenous Chinese and Tibetan horse breeds [66] and a few polymorphic haplotypes detected by next generation sequencing in modern horse breeds [67]. This very low Y polymorphism might be explained by the fact that throughout horse evolution and domestication there was strong selection, especially of males, with only a few stallions contributing to the following generations [22,65]. This is particularly true for the Sorraia where most of the breeding schemes have one stallion per herd per year [3]. Curiously, in contrast to horses, Y chromosome variants are present in other species in the genus *Equus,* including the Przewalski's horse, the donkey and zebras [63-65,67].

Average inbreeding in Sorraias was higher than in other breeds (Garrano, Andalusian, Lusitano, Lipizzan, Thoroughbred, Friesian, and even the endangered Przewalski's horse) (Supplementary Table 6). Since the prior study in Sorraias [9], inbreeding, mean d^2 and individual heterozygosity have increased 5.7%, 23.7% (in bp^2) and 10.7%, respectively. While the increase of inbreeding is detrimental and results in decreased genetic variation, increased homozygosity and inbreeding depression, increase in mean d^2 and individual heterozygosity denote improvement and show the positive results of the management-breeding plan of the Sorraia Breeders Association.

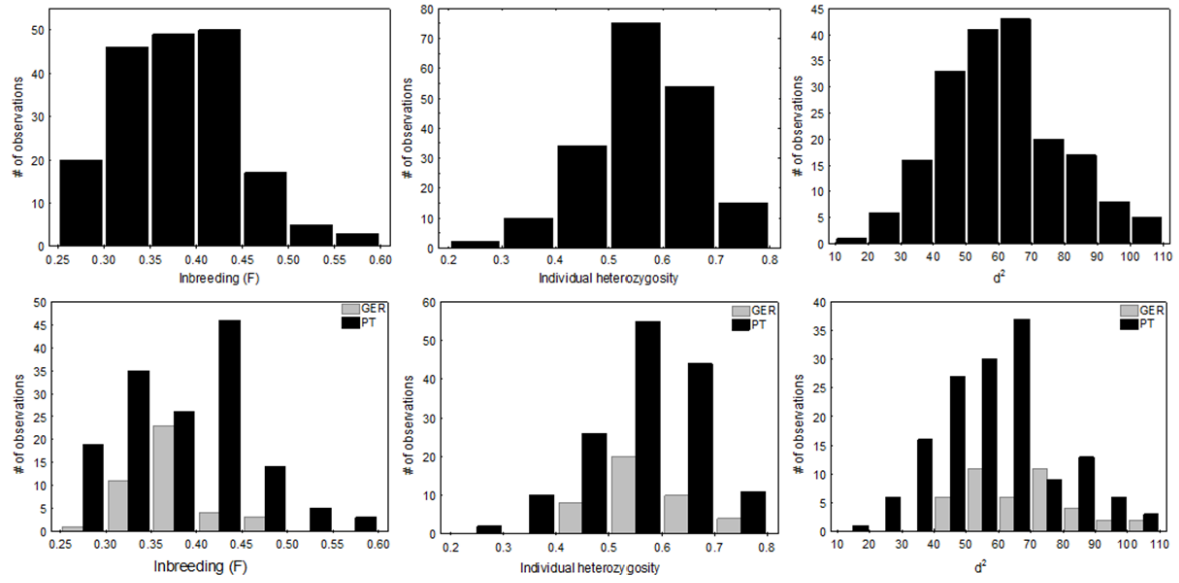


Figure 1 - Distribution of inbreeding, individual heterozygosities and d^2 in the whole population (top) and in subpopulations (bottom). Heterozygosities and d^2 calculated with up to 50 autosomal STRs in 190 animals.

Average inbreeding was higher in PT compared to GER, leading to a higher mean d^2 and individual heterozygosity in the latter (Figure 1), which is in agreement with STR results where GER had lower F_{IS}. The observed differences are direct consequences of different breeding systems: in PT, one stallion is chosen per herd of mares per year, while in GER, one stallion is chosen per mare per year almost. As a result, more stallions are used yearly in GER which, in turn increases genetic variation and decreases inbreeding in the population.

Compared to [9], negative correlation between inbreeding and both mean d^2 and individual heterozygosity increased and correlation between inbreeding and d^2 became statistically significant. Correlation between individual heterozygosity and mean d^2 remained significantly positive as in [9].

Average inbreeding calculated from SNP data was considerably lower than that calculated from pedigree data which is inconsistent with similar studies in cattle where inbreeding by pedigree was shown to be lower than by SNPs [68]. The differences are likely attributed to differences in pedigree data because a pedigree useful for inbreeding calculations must have sufficient depth and encompass the complete population over several generations [69]. This is the case in Sorraias where horses can be traced back to the founders, some with 12 complete generations [2]. Even though high-density SNP-chips are currently being used to calculate inbreeding in different species (e.g. [70-73]), pedigree-based inbreeding calculations should not be replaced by SNP-based ones, especially in highly inbred populations [69] like the Sorraia. Pedigrees can be particularly useful in calculations of more complicated inbreeding by descent coefficients, where SNPs might have restricted power [69]. While pedigree-based and SNP-based inbreeding rates were different, pedigree-based inbreeding in Sorraias (0.38; [12]) was in good agreement with inbreeding calculations based on whole genome sequencing and ROH [74]. The latter gave inbreeding coefficients for two Sorraia individuals as 0.38 and 0.33 [74] showing that this is an appropriate method for calculating inbreeding coefficients for this highly inbred population using sequencing data.

The observed positive correlation between SNP-based inbreeding and year of birth in the total, PT and GER populations indicates increased rate of inbreeding over the past 28-year period, being more pronounced in recent years. This is despite some recent negative trends in inbreeding, mostly thanks to the efforts of the Sorraia Breeders Association to promote the exchange of stallions between stud farms. This trend is similar to that found by SNP-based analysis in Thoroughbred horses [75] where inbreeding significantly increased over a 40-year period, but with lower correlation ($r=0.24$) than in this study.

The observed significantly ($p<0.05$) higher heterozygosity of STRs compared to SNPs in Sorraias is in good agreement with prior studies in horses [62] and other species [76-78]. One possible explanation is that SNPs are bi-allelic and have lower mutation rate than STRs [62].

Another reason might be that the STRs used in this study were chosen based on their higher variability in this particular breed, resulting in even higher heterozygosity rates. On the other hand, like shown for horses earlier [62], there was no correlation between the heterozygosity rates of STRs and SNPs corroborating with independent genomic distribution of the two types of polymorphisms.

Runs of homozygosity were found in 11 chromosomes (Supplementary Table 5), with chr2, 11, 22 and 27 sharing some common regions between different horses (Supplementary Figure 3). On chr27 a ROH was shared between three horses and there were two ROHs common in three horses in chr2 (Supplementary Figure 3, Supplementary Table 5). Interestingly, the largest ROHs were found in one of the largest (chr15) and one of the smallest (chr27) acrocentric chromosomes, while chr2, one of the largest in the complement, had the smallest ROH (Supplementary Table 5). It has been shown that short and long ROHs are signatures of ancient and recent inbreeding, respectively [68], revealing ancient inbreeding in the Sorraia due to small length. However, using sequencing data [74] found that Sorraias had the longest ROHs of all six studied breeds, and the second highest count (after a Thoroughbred horse). Notably, ROHs in chr4, 15, 11 and 24 in our study shared overlapping segments with those found by [74] (Supplementary Figure 4).

Principal component analysis (PCA) is a very useful tool to visualize genetic relationship between individuals from different breeds and within the same breed. It positions individuals in a 2D/3D graph based on genetic similarity as revealed by the analysis of molecular markers [79]. While STR-based STRUCTURE analysis separated Sorraias from 13 different stud farms into 3 clusters (Figure 3, Table 3), the STR-based FCA and SNP-based PCA analyses separated our total population into two main groups: PT and GER (Figure 5A). The FCA and PCA results are in agreement with the low F_{ST} values found between PT and GER and clearly separating both populations based on the breed's historic evolution and connections between stud farms. On the other hand, STRUCTURE analysis was able to demonstrate the extent of animal exchange between stud farms better than FCA or PCA analyses could. In both analyses, the first axis clearly separated horses from the Portuguese and "old" German breeders, whereas animals from the recent German breeder were indistinguishable from the Portuguese Sorraias due to the newly imported animals. The second axis separated the Portuguese breeders, though not completely, because of recent efforts by the Sorraia Breeders Association to promote the exchange of stallions between breeding farms, resulting in cluster 2 of STRUCTURE having a mix of different stud farms. This also explains why we could not see the 3 clusters in FCA or PCA analyses as proposed by STRUCTURE and Structure Harvester analyses.

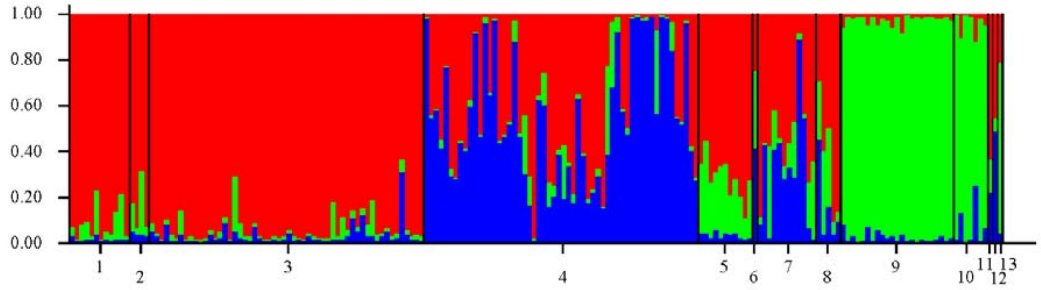


Figure 3 - STRUCTURE assignment of 190 animals from the 13 different breeders into K=3 clusters.

Despite some Sorraia specific CNVs none were ubiquitous to all animals, showing the variable nature of these markers. Although CNVs are an adequate tool to study diseases, they are not suitable to evaluate the degree of genetic diversity in this extremely inbred breed, since even with such low variability and high inbreeding levels there were no fixed CNVs and individual variation in CNVs was high. The two novel CNVs involved genes *ATP10D* and *RFX3*. *ATP10D* has functions in cation transport, metabolic processes, as well as phospholipid transport and translocation. *RFX3* (regulatory factor X, 3) is a transcription factor that influences HLA class II expression, DNA binding and protein binding and transcription. Of the 71 CNVRs found in Sorraia horses, 53 were shared with other horse breeds [35,40,41]. Of the remaining 18, 11 overlapped with results from [33], [34], [38], [42] and [36] (Supplementary Table 7) and 7 were Sorraia specific. The number of CNVs in Sorraia horses (n=213) was close to that found by [34] (n=282), higher than in [42] (n=50) but lower than in [33], [38], [35] and [36] (2368, 2797, 950 and 700, respetively). The number of CNVRs in Sorraia was also lower than found in [33] (775), [38] (478), [35] (258); [41] (245) and [36] (353).

The genome-wide data described herein provides improved evaluation of the genetic variation of the Sorraia breed and provides new possibilities to increase heterozygosity by choosing the right combination of sires and mares for mating. This data should be included in the current conservation management breeding program. The improvement of our study relative to previous ones in this breed is based on the multi-method approach to assess the genetic variability and differentiation of populations, the same used by [80] in wild horses from Doñana National Park with positive results.

Since STR genotyping can be done at a fraction of the cost of SNPs, and our results show higher variability and heterozygosity in the former, it is likely that parentage testing and genetic variability analysis for management purposes will continue to be done using STRs rather than SNPs. In the future, it would be of interest to use SNP genotyping to investigate signs of selection and also

to compare the Sorraia to other breeds and see how they are related using a more informative tool than used thus far.

Our results show that the Portuguese and "older" German populations now form two separate clusters with some genetic differentiation between them. It would be interesting and important to mix these two populations again, potentially by bringing some selected stallions for extensive breeding in Portuguese stud farms, in order to homogenise the genetic background and, in the short-term, increase genetic variability. Considering that our results show extremely high inbreeding levels, low genetic variability and that there is some evidence of inbreeding depression in the breed [12], the Sorraia Breeders Association might eventually consider the introduction of new animals from closely related breeds in order to improve the breeds' genetic health and prevent the permanent loss of this iconic and important animal genetic resource.

Acknowledgments

H. J. Kjöllerström was supported by a PhD grant (SFRH/BD/81502/2011) from the Portuguese Foundation for Science and Technology (FCT/MEC). This work was supported by cE3c FCT Unit funding (Ref. UID/BIA/00329/2013) and by PRODER (Contract number 57882) (EU)/Action 2.2.3 (2011-2013). The authors would like to thank the Sorraia Breeders Association (AICS) for privileged access to the breeders as well as for providing pedigree data. The authors would like to thank Dr. Jan Janečka for help in designing microsatellite with M13 tails, Dr. Pranab Das for sequencing non-parentage testing microsatellite data, and Raquel Silva for help in DNA extraction. The authors also wish to thank Dr. Gus Cothran and Dr. Rytis Juras for genotyping three samples used in SNP analysis.

References

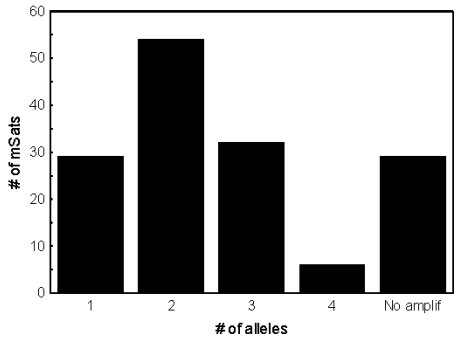
1. Gonzaga PG (2004) A History of the Horse, Volume 1: The Iberian Horse from Ice Age to Antiquity. London, UK: JA Allen & Co. Ltd.
2. Oom MM, d'Andrade JL, Costa-Ferreira J (2004) Studbook da raça Sorraia (in Portuguese); Sorraia AIdcdcidtp, editor. Alpiarça, Portugal: Garrido Artes Gráficas.
3. Oom MM (2006) O Cavalo do Sorraia. Um património histórico e genético a conservar (in Portuguese). Portugal: Edições Cosmos.
4. Oom MM, Cothran EG (1994) The genetic variation of an endangered breed, the sorraia horse. Animal Genetics 25: 15.
5. Luís C, Bastos-Silveira C, Cothran EG, Oom MM (2002) Variation in the mitochondrial control region sequence between the two maternal lines of the Sorraia horse breed. Genetics and Molecular Biology 25: 309-311.
6. Luís C, Cothran EG, Oom MM (2002) Microsatellites in Portuguese autochthonous horse breeds: usefulness for parentage testing. Genetics and Molecular Biology 25: 131-134.

7. Luis C, Cothran EG, Oom MM, Bailey E (2005) Major histocompatibility complex locus DRA polymorphism in the endangered Sorraia horse and related breeds. J Anim Breed Genet 122: 69-72.
8. Luis C, Bastos-Silveira C, Cothran EG, Oom MM (2006) Iberian origins of New World horse breeds. J Hered 97: 107-113.
9. Luis C, Cothran EG, Oom MM (2007) Inbreeding and genetic structure in the endangered Sorraia horse breed: implications for its conservation and management. J Hered 98: 232-237.
10. Luis C, Juras R, Oom MM, Cothran EG (2007) Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. Anim Genet 38: 20-27.
11. Pinheiro M, Kjöllerström HJ, Oom MM (2013) Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis. Livestock Science 152: 1-10.
12. Kjöllerström HJ, Gama LT, Oom MM (2015) Impact of inbreeding on fitness-related traits in the highly threatened Sorraia horse breed. Livestock Science 180: 84-89.
13. Bömcke E, Soyeurt H, Szydlowski M, Gengler N (2011) New method to combine molecular and pedigree relationships. J Anim Sci 89: 972-978.
14. Frankham R, Briscoe DA, Ballou JD (2002) Introduction to Conservation Genetics: Cambridge University Press.
15. Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics 144: 389-399.
16. Guerier AS, Bishop JM, Crawford SJ, Schmidt-Küntzel A, Stratford KJ (2012) Parentage analysis in a managed free ranging population of southern white rhinoceros: genetic diversity, pedigrees and management. Conservation Genetics 13: 811-822.
17. Putman AI, Carbone I (2014) Challenges in analysis and interpretation of microsatellite data for population genetic studies. Ecol Evol 4: 4399-4428.
18. Bruford MW, Ginja C, Hoffmann I, Joost S, Orozco-terWengel P, et al. (2015) Prospects and challenges for the conservation of farm animal genomic resources, 2015-2025. Front Genet 6: 314.
19. Bowling AT (2001) Historical development and application of molecular genetic tests for horse identification and parentage control. Livest Prod Sci 72: 111-116.
20. Bodó I, Alderson L, Langlois B (2005) Conservation Genetics of Endangered Horse Breeds: Wageningen Academic Publishers.
21. Leroy G, Callede L, Verrier E, Meriaux JC, Ricard A, et al. (2009) Genetic diversity of a large set of horse breeds raised in France assessed by microsatellite polymorphism. Genet Sel Evol 41: 5.
22. Warmuth V, Eriksson A, Bower MA, Canon J, Cothran G, et al. (2011) European domestic horses originated in two holocene refugia. PLoS One 6: e18194.
23. Penedo MCT, Raudsepp T (2013) Molecular Genetic Testing and Karyotyping in the Horse. Equine Genomics: Blackwell Publishing Ltd. pp. 241-254.
24. Warmuth V, Manica A, Eriksson A, Barker G, Bower M (2013) Autosomal genetic diversity in non-breed horses from eastern Eurasia provides insights into historical population movements. Anim Genet 44: 53-61.
25. Gurgul A, Semik E, Pawlina K, Szmatola T, Jasielczuk I, et al. (2014) The application of genome-wide SNP genotyping methods in studies on livestock genomes. J Appl Genet 55: 197-208.
26. McCue M, Mickelson J (2013) Genomic Tools and Resources: Development and Applications of an Equine SNP Genotyping Array. Equine Genomics: Blackwell Publishing Ltd. pp. 113-124.

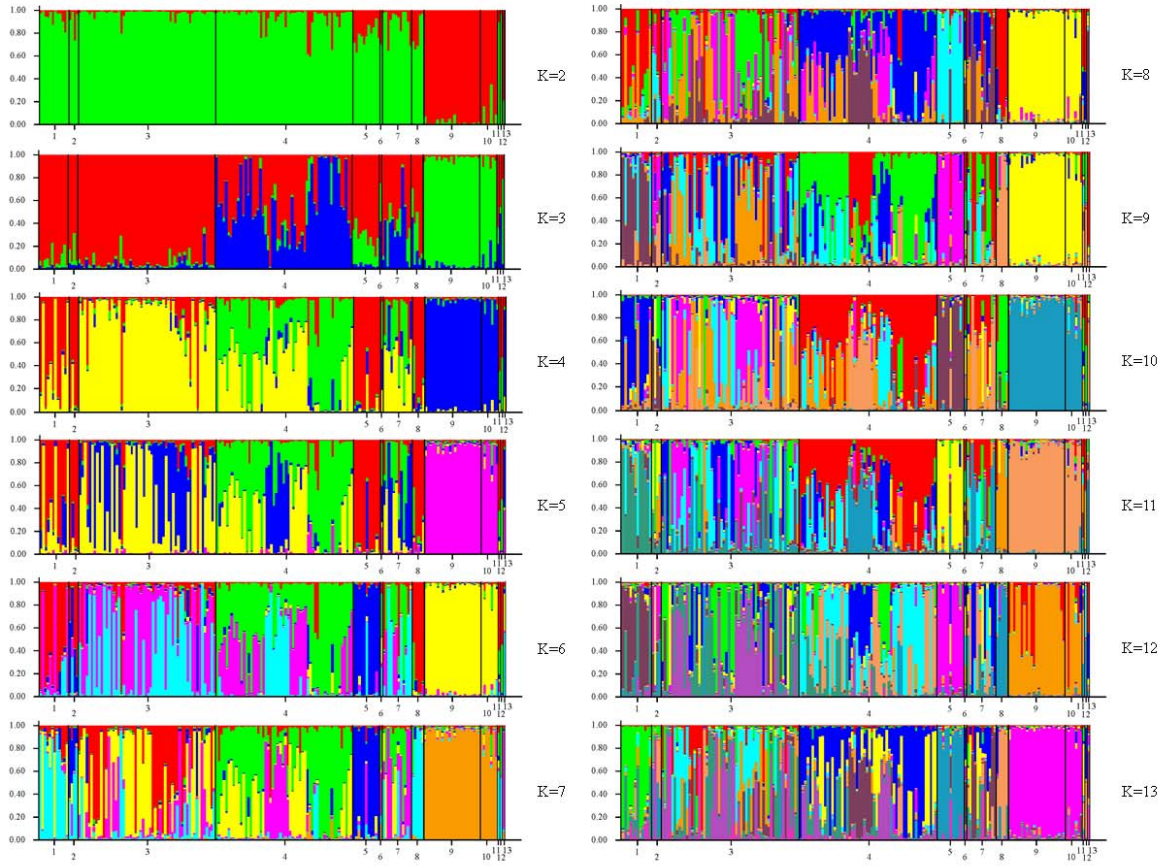
27. Orr N, Back W, Gu J, Leegwater P, Govindarajan P, et al. (2010) Genome-wide SNP association-based localization of a dwarfism gene in Friesian dwarf horses. Anim Genet 41 Suppl 2: 2-7.
28. Raudsepp T, McCue ME, Das PJ, Dobson L, Vishnoi M, et al. (2012) Genome-wide association study implicates testis-sperm specific FKBP6 as a susceptibility locus for impaired acrosome reaction in stallions. PLoS Genet 8: e1003139.
29. Signer-Hasler H, Flury C, Haase B, Burger D, Simianer H, et al. (2012) A Genome-Wide Association Study Reveals Loci Influencing Height and Other Conformation Traits in Horses. PLoS ONE 7: e37282.
30. McCue ME, Bannasch DL, Petersen JL, Gurr J, Bailey E, et al. (2012) A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. PLoS Genet 8: e1002451.
31. Petersen JL, Mickelson JR, Cothran EG, Andersson LS, Axelsson J, et al. (2013) Genetic diversity in the modern horse illustrated from genome-wide SNP data. PLoS One 8: e54997.
32. Holl HM, Lear TL, Nolen-Walston RD, Slack J, Brooks SA (2013) Detection of two equine trisomies using SNP-CGH. Mamm Genome 24: 252-256.
33. Doan R, Cohen N, Harrington J, Veazey K, Juras R, et al. (2012) Identification of copy number variants in horses. Genome Res 22: 899-907.
34. Doan R, Cohen ND, Sawyer J, Ghaffari N, Johnson CD, et al. (2012) Whole-genome sequencing and genetic variant analysis of a Quarter Horse mare. BMC Genomics 13: 78.
35. Ghosh S, Qu Z, Das PJ, Fang E, Juras R, et al. (2014) Copy Number Variation in the Horse Genome. PLoS Genet 10: e1004712.
36. Wang W, Wang S, Hou C, Xing Y, Cao J, et al. (2014) Genome-Wide Detection of Copy Number Variations among Diverse Horse Breeds by Array CGH. PLoS ONE 9: e86860.
37. Sundstrom E, Imsland F, Mikko S, Wade C, Sigurdsson S, et al. (2012) Copy number expansion of the STX17 duplication in melanoma tissue from Grey horses. BMC Genomics 13: 365.
38. Dupuis MC, Zhang Z, Durkin K, Charlier C, Lekeux P, et al. (2013) Detection of copy number variants in the horse genome and examination of their association with recurrent laryngeal neuropathy. Anim Genet 44: 206-208.
39. McQueen CM, Doan R, Dindot SV, Bourquin JR, Zlatev ZZ, et al. (2014) Identification of genomic loci associated with Rhodococcus equi susceptibility in foals. PLoS One 9: e98710.
40. Ghosh S, Das PJ, Arnold C, Jaxheimer J, Varner DD, et al. (2014) Contribution of DNA copy number variants to equine cryptorchidism. Journal of Equine Veterinary Science 34: 29.
41. Ghosh S, Das PJ, McQueen CM, Gerber V, Swiderski CE, et al. (2016) Analysis of genomic copy number variation in equine recurrent airway obstruction (heaves). Anim Genet.
42. Metzger J, Philipp U, Lopes MS, da Camara Machado A, Felicetti M, et al. (2013) Analysis of copy number variants by three detection algorithms and their association with body size in horses. BMC Genomics 14: 487.
43. Montgomery GW, Sise JA (1990) Extraction of DNA from sheep white blood cells. New Zealand Journal of Agricultural Research 33: 437-441.
44. Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotech 18: 233-234.
45. Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16: 1099-1106.
46. Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. Journal of Heredity 86: 248-249.
47. Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour 8: 103-106.
48. Wright S (1951) The genetical structure of populations. Ann Eugen 15: 323-354.

49. Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. Evolution 38: 1358-1370.
50. ISIS (2011) SPARKS (Single Population Analysis and Records Keeping System). v.1.6 ed. Eagan,MN.: International Species Information System.
51. Ballou J (1983) Calculating inbreeding coefficients from pedigrees. In: Schonewald-Cox CM, Chambers SM, MacBryde B, Thomas L, editors. Genetics and Conservation: A Reference for Managing Wild Animal and Plant Populations. Menlo Park, California: Benjamin/Cummings. pp. 509-520.
52. StatSoft (2013) Statistica Version 12.
53. Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).
54. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945-959.
55. Raj A, Stephens M, Pritchard JK (2014) fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. Genetics 197: 573-589.
56. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611-2620.
57. Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4: 359-361.
58. Purcell S PLINK v1.07.
59. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, et al. (2007) PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. The American Journal of Human Genetics 81: 559-575.
60. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, et al. (2012) Primer3--new capabilities and interfaces. Nucleic Acids Res 40: e115.
61. Aberle KS, Hamann H, Drogemuller C, Distl O (2004) Genetic diversity in German draught horse breeds compared with a group of primitive, riding and wild horses by means of microsatellite DNA markers. Animal genetics 35: 270-277.
62. Khanshour A (2013) Genetic diversity and population structure of the Arabion horse populations from Syria and other countries. College Station, Texas: VIBS, Texas A&M University. 135 p.
63. Wallner B (2004) Isolation of Y Chromosome-specific Microsatellites in the Horse and Cross-species Amplification in the Genus Equus. Journal of Heredity 95: 158-164.
64. Kakoi H, Hirota K, Gawahara H, Kurosawa M, Kuwajima M (2005) Genetic diagnosis of sex chromosome aberrations in horses based on parentage test by microsatellite DNA and analysis of X- and Y-linked markers. Equine Vet J 37: 143-147.
65. Lindgren G, Backstrom N, Swinburne J, Hellborg L, Einarsson A, et al. (2004) Limited number of patrilines in horse domestication. Nat Genet 36: 335-336.
66. Ling Y, Ma Y, Guan W, Cheng Y, Wang Y, et al. (2010) Identification of Y chromosome genetic variations in Chinese indigenous horse breeds. J Hered 101: 639-643.
67. Wallner B, Vogl C, Shukla P, Burgstaller JP, Druml T, et al. (2013) Identification of Genetic Variation on the Horse Y Chromosome and the Tracing of Male Founder Lineages in Modern Breeds. PLoS ONE 8: e60015.
68. Zhang Q, Calus MP, Guldbrandtsen B, Lund MS, Sahana G (2015) Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. BMC Genet 16: 88.
69. Wang J (2016) Pedigrees or markers: Which are better in estimating relatedness and inbreeding coefficient? Theoretical Population Biology 107: 4-13.

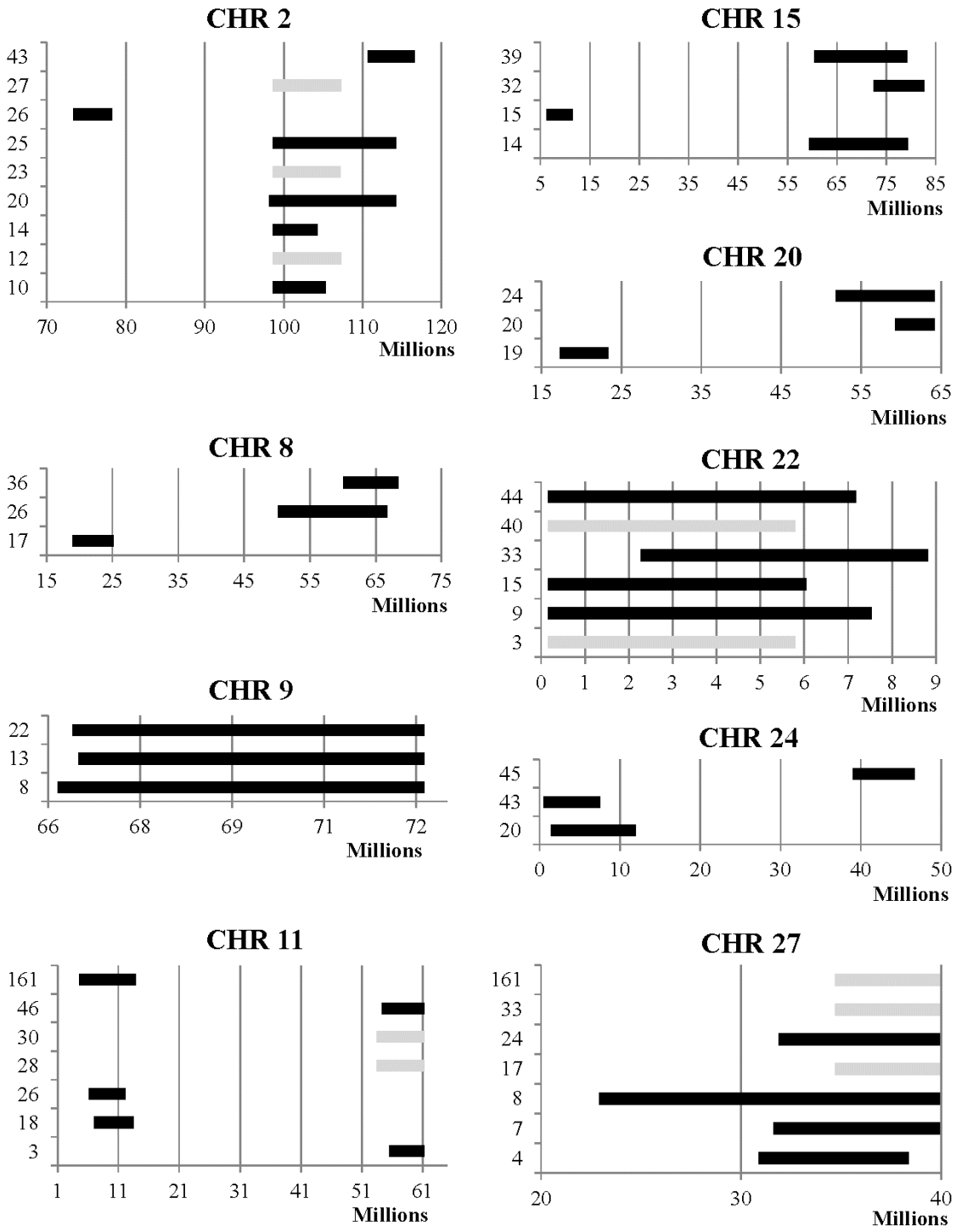
70. de Cara MAR, Fernández J, Toro MA, Villanueva B (2011) Using genome-wide information to minimize the loss of diversity in conservation programmes. Journal of Animal Breeding and Genetics 128: 456-464.
71. Lenstra JA, Groeneveld LF, Eding H, Kantanen J, Williams JL, et al. (2012) Molecular tools and analytical approaches for the characterization of farm animal genetic diversity. Animal Genetics 43: 483-502.
72. Curik I, Ferenčaković M, Sölkner J (2014) Inbreeding and runs of homozygosity: A possible solution to an old problem. Livestock Science 166: 26-34.
73. Knief U, Hemmrich-Stanisak G, Wittig M, Franke A, Griffith SC, et al. (2015) Quantifying realized inbreeding in wild and captive animal populations. Heredity (Edinb) 114: 397-403.
74. Metzger J, Karwath M, Tonda R, Beltran S, Agueda L, et al. (2015) Runs of homozygosity reveal signatures of positive selection for reproduction traits in breed and non-breed horses. BMC Genomics 16: 764.
75. Binns MM, Boehler DA, Bailey E, Lear TL, Cardwell JM, et al. (2012) Inbreeding in the Thoroughbred horse. Anim Genet 43: 340-342.
76. Coates BS, Sumerford DV, Miller NJ, Kim KS, Sappington TW, et al. (2009) Comparative Performance of Single Nucleotide Polymorphism and Microsatellite Markers for Population Genetic Analysis. Journal of Heredity 100: 556-564.
77. Varela MA, Amos W (2010) Heterogeneous distribution of SNPs in the human genome: microsatellites as predictors of nucleotide diversity and divergence. Genomics 95: 151-159.
78. Forstmeier W, Schielzeth H, Mueller JC, Ellegren H, Kempenaers B (2012) Heterozygosity-fitness correlations in zebra finches: microsatellite markers can be better than their reputation. Mol Ecol 21: 3237-3249.
79. Almeida PAR (2007) Diversidade genética e diferenciação das raças portuguesas de ovinos com base em marcadores de DNA - microssatélites: uma perspectiva de conservação [Monography]. Vila Real: Universidade de Trás-os-Montes e Alto Douro. 305 p.
80. Vega-Pla JL, Calderón J, Rodríguez-Gallardo PP, Martinez AM, Rico C (2006) Saving feral horse populations: does it really matter? A case study of wild horses from Doñana National Park in southern Spain. Animal Genetics 37: 571-578.



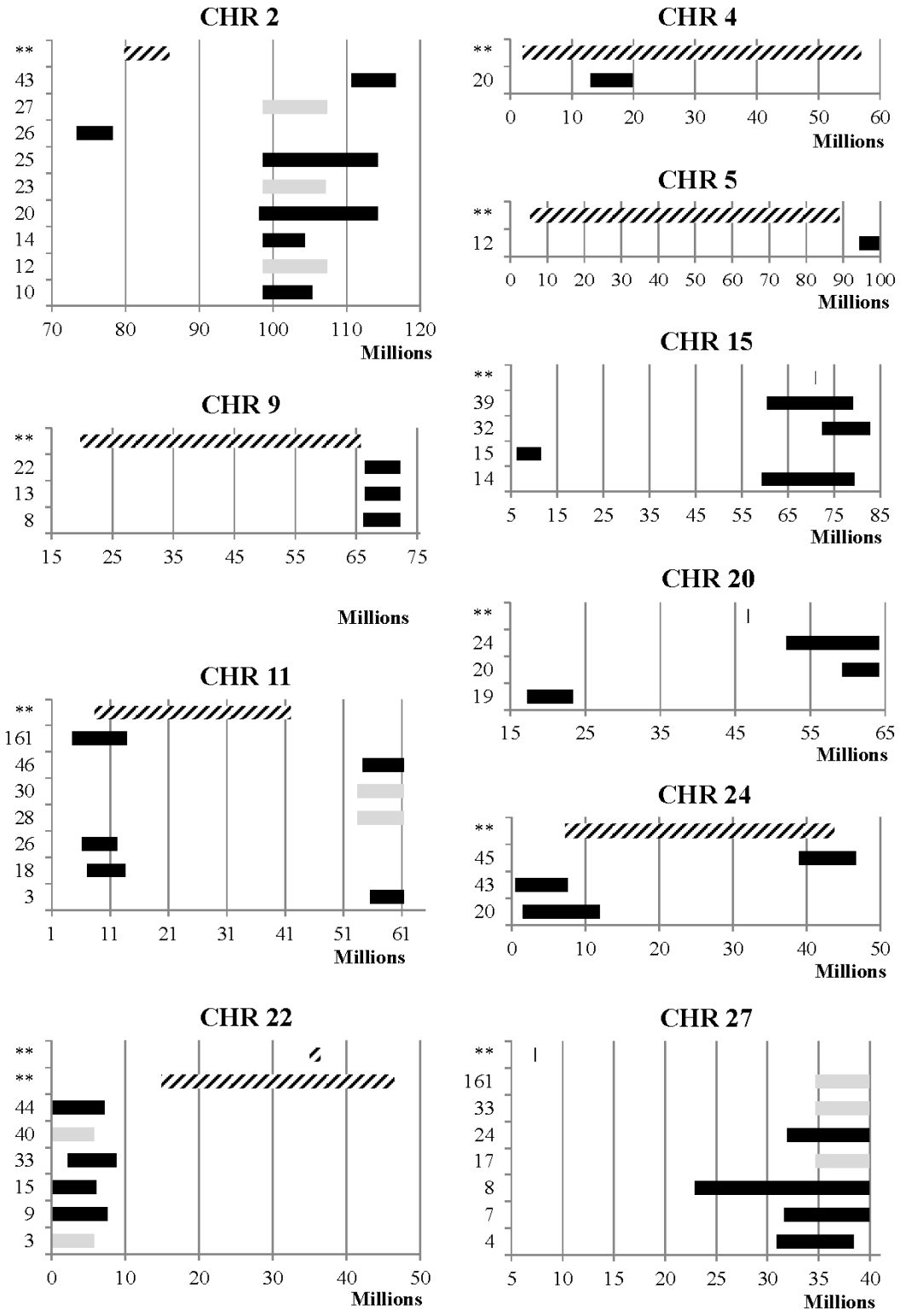
Supplementary figure 1 - Number of alleles for the 150 tested microsatellites. No amplif - loci with no amplification.



Supplementary Figure 2 - Clustering assignment obtained by STRUCTURE analysis from K=2 to K=13. Each vertical column is an individual (n=190) further divided into different coloured segments representing relative genome proportion belonging to a particular cluster. Populations are separated by black lines.



Supplementary Figure 3 - Chromosomes (CHR) with more than one animal with ROH. Individual ID is on the left side of each graph. ROH size on the bottom in millions of bp. Grey bars represent matching ROHs in different animals.



Supplementary Figure 4 - Chromosomes (CHR) with more than one animal with ROH. Individual ID is on the left side of each graph. ROH size on bottom of the graph in millions of bp. Grey bars represent matching ROHs in different animals. ** - results for non-breed horses in Metzeger *et al.* 2015, dashed bars.

Supplementary Table 1 - Microsatellites tested (N=150), chromosome location (Chr), average number of alleles, NVBI accession number, and their primer sequences

Chr	Marker	Reference	Average alleles	Accession number	Primer sequence (5' – 3')
1	ABGe105	Mittmann *et al.* (2010)	12	AM946984	GGGTCTTTTGACTGCCTGAG TTGGGAGACGAGAACAAAGG
1	ABGe001	Mittmann *et al.* (2010)	28	AM900755	TGGTGACGTAAGGGTTCTGG GAGGGGATATGTGGATGTGG
1	NVHEQ100	Roed *et al.* (1998)	10	AF056399	CCAAAGCAGAACATGTGAAGTT TGGCATAGATGTTAGCTCAGTGA
1	COR100	Tallmadge *et al.* (1999a)	9	AF154953	CCCAGAGGTTTCAGAGGG ATTCTAGGGCATATTATGACAA
1	AHT021	Swinburne *et al.* (2000)	8	503068	TCCAAGTTGCTGAATGGATC ACGGCCTGATTCTCTCTTTG
1	UMNE471	Wagner *et al.* (2004)	9	AY464495	GAGCAATCACTTCCTCTGTGG CCTTCTTCCCTCAACACAGC
1	TKY002	Sakagami *et al.* (1995)	9		TTCCCTCCCATGGTTATTTTTC TCTCTACTTTCATATACATTTGG
1	UCD493	Chowdhary *et al.* (2003)	15	U67418	ATTGGATATTTAACACCAAATGC CCCAGCTCAGTGACTCCATT
1	AHT58	Swinburne *et al.* (2003)	10	AJ507675	CAGTGATGAGCCGCAAATAG TCTACCTATAATCCGCCTCCC
1	HMS15	Guerin and Bertaud (1996)	9	U35401	ATATCTCTTGCTGTCCTACTTTCC AATGTGACACGTAAGATAGGCCTC
1	UMNe115	Wagner *et al.* (2004)	10	AY391296	TCCTCCTACACTGGCCATATC TTTCCTATCGGAGTGCTTGC
1	UM026	George *et al.* (1998)	8	AF195573	CCCAAAATCAATTAGGTCTC ATCAGTTGCTCTCTACTTTTC
2	COR065	Tallmadge *et al.* (1999b)	10	AF142602	CAAAAGCACACACAAAGTGC TCCGGAAAGTGCAAAGTTAG
2	ABGe109	Mittmann *et al.* (2010)	14	AM946988	GGGTGGCTCCTTAGAGCTTC CCCCTCCCTTGTTTATATGC
2	TKY384	Tozaki *et al.* (2004)	12	AB048290	TGCAGCAAGAAACCTAAACA CTTCAGTTGTAATCAGGCTC
2	ABGe144	Mittmann *et al.* (2010)	10	FM165574	CAAAAATGGCAAGATTTCATCC TGCCCACTGACAGATGAATG
2	UMNe323	Wagner *et al.* (2004)	16	AY391338	GATCCTGCAGGAAAGCATGT CCGCTCGGAATATTTCATTG
2	ABGe006	Mittmann *et al.* (2010)	9	AM900760	CTGAAACCAGCCAGGAAAAG TCTCCTAGCCGGGAGAAAAC
2	ASB17	Breen *et al.* (1997)	14	X93531	GAGGGCGGTACCTTTGTACC ACCAGTCAGGATCTCCACCG
2	TKY784	Tozaki *et al.* (2004)	8	AB104002	GATCAGTACTTTGCAAATGGATAAC GTAACTCCAAGGCTACGTTC
2	ABGe065	Mittmann *et al.* (2010)	16	AM940026	ATGCTCGTTTCACAAAGAGG TCTTTTTGTGCAGGGTGTG

2	A - 14	Marti *et al.* (1998)	10	Y10239	CAGCTGGGTGACACAGAGAG GTCATCACTACTCCCTACAC
3	AHT36	Swinburne *et al.* (2000)	8	AJ271521	TGCTGCTCCAGTGTCCT TAGATTTCACAGGCGGGTG
3	COR028	Murphie *et al.* (1999)	8	AF101397	TAAAGAGGAAGGCAATGGAC ACCTTTTGTGCTAGGCACTG
3	COR033	Murphie *et al.* (1999)	10	AF101402	CCTCCCCTACTTCCTCTCTG CATTTTCTTTCCAGGTTCCC
3	TKY937	Tozaki *et al.* (2004)	8	AB104155	TCCTGCGGAAATACATTAGG AGTTCAAAGTGGTCCCATAG
3	UMNe158	Wagner *et al.* (2004)	10	AY391305	AATTGAGAGCCAAGATGACACC GGCACCATTTGAGGAAGATG
3	AHT92	Swinburne *et al.* (2003)	21	AJ507709	TGAGCATCTTGAAGATGAGCA CAACAGTTGTTAGCTCAGGTGC
4	AHT43	Swinburne *et al.* (2000)	13	AJ271528	ACACAAGTGACAGGAGCGTG TGGAAGCATGCAAGAGGTC
4	AHT84	Swinburne *et al.* (2003)	19	AJ507701	TGGCAATCTGCAGGGAAC GATCTTGTGATTGTGTGTGTG
4	ABGe069	Mittmann *et al.* (2010)	10	AM940030	CATGGCAACGACAATACAAA TGGATTTACAGTGCAAGCAG
4	LEX061	Breen *et al.* (1997)	8	AF075661	TCAGTGTTCCCATCTGTA TGAAATCACACCTTTACTTTA
4	LEX050	Coogle and Bailey (1997)	7	AF075652	ATAGTCTGGGGTTAGGTAAGG TCTAGCCCAATGTAAATGC
4	TKY1451	Tozaki *et al.* (2007)	9	AB215394	CTGAGATTAACGGCCCAGTA TCAGTCATGTATTCCTGTGCAT
4	COR089	Tallmadge *et al.* (1999a)	9	AF154942	CCTGCCATAAATTTGTTTCC TCCCTACCTCATCTCCACAC
4	ABGe059	Mittmann *et al.* (2010)	8	AM919498	AGTTGCCTCTGGTCTTGCAG GCTGGCAGAATGTCTGTTTTC
4	TKY354	Tozaki *et al.* (2001)	11	AB044854	AGTGAGGTCTTCCTTGACTG TGTTAGATGGTGGTAAGTGC
4	TKY720	Tozaki *et al.* (2004)	9	AB103938	CAGGAGTATCCAGAATTGCAGA CCAGCTGTGTGTAACGCAAT
4	SGCV23	Godard *et al.* (1997)	19	U90601	GGCTTAAGATATGGGTGAGTAAGG GCCCACCCTCTTACTTTTCTCAA
5	COR062	Tallmadge *et al.* (1999b)	10	AF142599	GTCATCCAGTGACGAACACA AGGAAGTGCGCAGTAGAGAA
5	TKY456	Tozaki *et al.* (2004)	9	AB103674	GCTGGATGGATAAATGAATGG TGTTCGACTGAGCAGAGAGG
5	TKY887	Tozaki *et al.* (2004)	11	AB104105	GAGAACTAGATGCCACCC TGTTGGAGTGTGTAGGCT
5	TKY544	Tozaki *et al.* (2004)	9	AB103762	TGACCCAGTAAGCAGCCTGT TGCCAGGGGAAGAACATTTA
5	LEX069	Coogle and Bailey (1999)	10		TTTCTTTTTCCCACTTAAAGC TGGGACTTAGCAGTATGAAAC
5	ABGe012	Lampe *et al.* (2009)	8	AM905690	TCGAGTGCAACAATGTGTAGG AGTCGAAGGCTTCCCACTAC

5	ABGe141	Lampe *et al.* (2009)	12	AM992893	GAATTAGTTGTTTACTAGATTGGGAG TTGTTGCAAAAATTGATGAGTG
5	AHT107	Swinburne *et al.* (2003)	10	AJ507724	ATCTAACCAGAGCGCAACGT CCCGACACAGAAGATGGG
6	HTG31	Lindgren *et al.* (1999)	7		CTCTGTAACCCTTATATCCTTA TGTTGATTGCTCCTCCCCT
6	TKY1001	Tozaki *et al.* (2004)	9	AB104219	TCTCAGAAGCCATCTGGAG ATCGATGCAGAACACGTGG
6	NVHEQ82	Bjornstad *et al.* (2000)	7	AJ245770	TGTGGCAGCATCCCACAAAC CCTCCATTTTTGTCGGTTAGCG
6	COR070	Tallmadge *et al.* (1999b)	12	AF142607	CATCTGTTCCGTGGCATTA TTCAGGTGTGGGTTTTGAATC
6	TKY412	Tozaki *et al.* (2004)	8	AB103630	GTGTGGGACAGGAAGTTTGG ATTCTTGGGTCCCCTCATCT
7	TKY034	Hirota *et al.* (2001)	20	AB048340	GCCACTGTAATGGTCTGAAG GGAAGGTAGAGCATTCTCAG
7	ABGe102	Mittmann *et al.* (2010)	9	AM946385	AGGAAGCACGATCTGGTCTG AAGGATGCCCTGAGGAAGTC
8	COR012	Hopman *et al.* (1999)	6	AF083455	TCTAGGAAAGACCCATCACG AGTAAGTGGAGGCCAAGGAT
8	COR003	Hopman *et al.* (1999)	9	AF083446	TAGGGAAACTCCTCAAAGCC GAAACCAAAACCTTCATCCA
8	COR056	Ruth *et al.* (1999)	10	AF108373	AGATTCCAGGCATTAGGACC TCAGGGACAATCTTCCTCAAG
9	COR008	Hopman *et al.* (1999)	12	AF083451	TAAGTGCTGAGTCTGGGACC TGGTAGATAGCGTCTGGAGG
9	TKY453	Tozaki *et al.* (2004)	9	AB103671	CTTGCACCTATCAGCAGCAG GCAGAAAGGGATGAGGACAG
9	ASB4	Breen *et al.* (1997)	9	X93518	TAAATTGTAAAAGCTGGAGCCG GCAAATAGTAGTTAAGTCCTC
10	TKY601	Tozaki *et al.* (2004)	9	AB103819	CGAGGGGGAATTTTGTTTGT ATAGAGCCATGCAGGGGAAA
10	COR020	Hopman *et al.* (1999)	9	AF083463	TCTCTACCGCAAGTGAAACC CTGAATTGTAGGACATCCCG
10	NVHEQ18	Roed *et al.* (1998)	15	AF011404	GGAGGAGACAGTGGCCCCAGTC GCTGAGCTCTCCCATCCCATCG
10	UCD412	Eggleston-Stott *et al.* (1997)	10	AF000011	AGAGGAAGGCGACAGGTC CATCCGTCCATCCATCAG
10	ASB9	Breen *et al.* (1997)	9	X93523	GTGCGCATGTATGTGCGTGCC ATTTCCACAAGGGACATGAGG
10	AHT86	Swinburne *et al.* (2003)	8	AJ507703	CCCAATGAAGTCCAAGATGG GAAATCTCTAGCAAGACCCAGG
10	ABGe357	Mittmann *et al.* (2010)	9	FM179631	TGCACCAGCACTGGTAAAAG TGTACCTTTGCATTCTTTGTGG
10	COR048	Ruth *et al.* (1999)	10	AF108365	GATTGGGATGCAAAGATGAG CAAGAGGATTGGGAACAAAGG

11	ABGe099	Mittmann *et al.* (2010)	11	AM946382	TTCCTTCTGATTGCACCACTC ATTGTGGGTGACTCCCTCTG
11	SGCV24	Godard *et al.* (1997)	11	U90602	CTACCATTGAAGAGGGGTGGC GAAACGAGCAGGAAGTGAATCTCC
11	UCD457 [1]	Godard *et al.* (1997)	11	U67412	GGGGCGTGAGCATAAAGG CGCTGGATGAGTGAGGGA
12	SGCV08	Godard *et al.* (1997)	10	U90590	GAGTTCATTCTTTTTCGTGGCTG GGAAACACCCTAAGTGTCCCTTG
12	COR058	Ruth *et al.* (1999)	13	AF108375	GGGAAGGACGATGAGTGAC CACCAGGCTAAGTAGCCAAAG
13	COR069	Tallmadge *et al.* (1999b)	8	AF142606	AGCCACCAGTCTGTTCTCTG AATGTCCTTTGGTGGATGAAC
14	TKY1053	Tozaki *et al.* (2004)	7	AB104271	ATACTGGCTTTACGTCACAG ATCACCACCAGAGTTAATGG
14	UM010	Meyer *et al.* (1997)	7	AF195129	TACAGCCATTGGAAATCTAC CACCATTACATTTTCCCAG
14	TKY435	Tozaki *et al.* (2004)	8	AB103653	GTTCGTCTGTTTCTAGCCTC TATCTCCACATGGTACTCTC
14	TKY636	Tozaki *et al.* (2004)	8	AB103854	TAATCGAGGGGGCCTTAATC CGCTCTCTCTAAAGGCTCCA
15	B-8	Marti *et al.* (1998)	9	Y10240	TCCTCAGTCCTTTCTCATGC AGCTGAAGGCAATCTGTACC
15	LEX046	Coogle *et al.* (1997)	9	AF075648	ATAAGCCAATCCACTTTTCC ATTACCACCCCATTTCCTT
15	TKY1033	Penedo *et al.* (2005)	8	AB104251	AGACATGGATTTAGGGAGTG GCAGAGCCATGCTAAAACTG
15	ASB15	Breen *et al.* (1997)	10	X93529	GTCCCAAAGGGACTCAGGAAGG TGGATGCCAGTGCATAGACAG
15	ASB19	Breen *et al.* (1997)	11	X93533	GAGTTGGAGCTCAAGTCTGTC GTTTAGCAACTACAGCGTAGG
15	AHT16	Swinburne *et al.* (2006)	9		ATGTTGTGCAAATGGGATGA TGCCCATTGATTGATGATTG
15	ABGe147	Mittmann *et al.* (2010)	9	FM165577	TGGAATATCCCAGTCAAAATG CACTCCCTGAACCACAGGAG
15	ABGe114	Mittmann *et al.* (2010)	9	AM946993	AACAGTTGTGGGGAGAGTGG CCTCCTCCTAGCCTGTTTCC
15	COR014	Hopman *et al.* (1999)	12	AF083457	CTATCATGTCAGGGACCAGG CTGCCCTAGTTAGCAACCAA
15	COR075	Tallmadge *et al.* (1999b)	8	AF083457	GCCCTAGTTAGCAACCAACA AAGATTGATTCCTCAGCACG
16	ABGe094	Mittmann *et al.* (2010)	9	AM942735	AACTGCTGGCTGGATCTCTG AAGACTGCCCCATTCAATACTC
16	ABGe033	Mittmann *et al.* (2010)	6	AM919472	GGGTTTGCTTGTGAACTTCTG GTGAAGCCCTGACTTTGAGC
16	HTG013	Marklund *et al.* (1994)	10	AF169297	TTAGCACGGGGAGATCGGATCCTG GGTCTCCCTCTCCATTCACCCTGC

16	ABGe058	Mittmann *et al.* (2010)	10	AM919497	CCACACAGTATTCCCCCAAG GGAGAGAGGGTTCAGTGCAG
16	UCD505	Eggleston-Stott *et al.* (1997)	8	U67421	ATCACTCTCTTGTTGAGATAAC GGGATTTCCTTCTTTCTC
16	I-18	Marti *et al.* (1998)	9	Y10244	CAACAAAGATGTTGCAAGGG TGTGCCTCTTGTCTCTTAGG
16	AHT60	Swinburne *et al.* (2003)	10	AJ507677	GGTCAAGCTTTTGGTTTTTCC CCTAAGGAAGAGCTGTTCTTGC
16	TKY341	Tozaki *et al.* (2001)	9	AB044842	TATCCAGTCACCCATTTTAC TTGTGTCAGTACACTCTATG
17	COR105	Swinburne *et al.* (2000)	7		TTTCCTCATTGCTTCCTGAG CCCAAGGTCTGTCTTGCTCTC
17	COR007	Hopman *et al.* (1999)	9	AF083450	GTGTTGGATGAAGCGAATGA GACTTGCCTGGCTTTGAGTC
17	UMNe176	Mickelson *et al.* (2003)	11	AF536275	TTTTGAGGGGTGTGTTACAGC TTACCAGAGTTCTTACCTGGGG
17	TKY684	Tozaki *et al.* (2004)	8	AB103902	TTTGCAGGCTTTCTGTATTTTT TTCTGTTTCGTTTTCCCTGAA
18	TKY19	Kakoi *et al.* (1999)	8	AB048330	CTTCTGCTGATTCCTGAATG GGATCTCCTTAAATGGAACA
18	ABGe151	Mittmann *et al.* (2010)	8	FM177589	CTCACTCTGGGCCCACTATC CGGAGTGAGAAGACAGTCCAG
18	ABGe152	Mittmann *et al.* (2010)	9	FM177590	CCCTAGGTCCCCCACTTTAG CCATCCCTTCAGGAATACCAC
18	COR096	Tallmadge *et al.* (1999a)	9	AF154949	CCCCTCTTTTGCTTGAGAAT GCGTGTATGTGAGGATTGAAG
18	ABGe155	Mittmann *et al.* (2010)	9	FM177593	GGTCAGAAGACAGTCAAGAGTCC CCTCTCAGGCCTCTTACCAC
18	TKY741	Tozaki *et al.* (2004)	9	AB103959	CCTTCCTTCTCCTAACTCAGTCC TGGAAACCAGGAATAGGTGTG
18	TKY101	Mashima *et al.* (1999)	9		TCTGAAATACCGTGTGCCT TTCTGCCTCCCTCCAACTTT
18	ABGe157	Mittmann *et al.* (2010)	10	FM177595	GAGGGAGTCATTCCTGTACCC CCTCAGCCATGAATCTACCAG
18	ABGe159	Mittmann *et al.* (2010)	8	FM177597	TCGGCTCTTTTCTTCTATTTGC TCGGGCTCTGAATGAGAAAC
19	TKY448	Tozaki *et al.* (2004)	10	AB103666	TTCTTTTGCTCTCCCTCTCG GGAAAGACAGAGTAAGTGCGTGT
19	LEX073	Bailey *et al.* (2000)	9	AF213359	CCAGCCATCCACTGGTAGAG GGGAAAAGGGGAACCTTCTA
19	AHT55	Swinburne *et al.* (2003)	8	AJ507672	TGAAAATACACCCAGCTACGC GGGAGATATTTCTTGGCTTGC
20	UM011	Meyer *et al.* (1997)	13	AF195130	TGAAAGTAGAAAGGGATGTGG TCTCAGAGCAGAAGTCCCTG
20	TKY477	Tozaki *et al.* (2004)	9	AB103695	CCCCTCTCTCTCTCAAGTGC CTCCTGGGTGGGAGAACTTT

20	TKY321	Tozaki *et al.* (2000)	8	AB034629	CACTGTGTAACACTAACACC TGTGACTTCAAGAACAGACG
21	SGCV16	Godard *et al.* (1997)	7	U90594	AATTCTCAAATGGTTCAGTGA CTCCCTCCCTTCCTTCTA
21	TKY806	Tozaki *et al.* (2004)	9	AB104024	TGGAACTGTGATGATGTTGC TCTTTCTTCCCTTCCGAGAG
21	COR073	Tallmadge *et al.* (1999b)	7	AF142610	GCCAAGACATGGAAACAATC GTTCTCAAGGTGCATCCCTA
21	TKY296	Tozaki *et al.* (2000)	11	AB034605	CTCTCACTTCCAAGACACTC ATCAAACGTACAGGAAGAGC
21	TKY623	Tozaki *et al.* (2004)	7	AB103841	CAGTGTGGGTGGGCTTTATC ACCACTAGGGTGTGCATGTG
22	HTG14	Marklund *et al.* (1994)	8	AF169298	CCAGTCTAAGTTTGTTGGCTAGAA CAAAGGTGAGTGATGGATGGAAGC
22	HTG21	Lindgren *et al.* (1999)	7		ATTACTTCCTCCAGGTATCTCAG AGGCAGGGCTGGGAGACGT
22	COR016	Chowdhary *et al.* (2003)	7	AF083459	CAGCTCAGTAGATGATTGTCCA GCAAAGACAAGGAGGTTAAGTT
22	ABGe121	Mittmann *et al.* (2010)	12	AM947000	AGGAGCTGGAACTGACACAG GCTTCTCAGGGCAGTATTCC
23	COR055	Chowdhary *et al.* (2003)	9	AF108372	TAGTGACGCCTACGGATTTC CCCAAGAGGGCTTAGAAAGAG
23	TKY568	Tozaki *et al.* (2004)	11	AB103786	TTCCTGACGTGAAGGCATTA TGCCCTTCCTGCCTAGTAGA
24	Lex074	Bailey *et al.* (2000)	10	AF213360	AAGAGTGCTCCCGTGTG GACAATGCAGAACTGGGTAA
24	TKY524	Tozaki *et al.* (2004)	7	AB103742	AGTTGTGGCTTGCTTTCTAC TTGCACTTGAGCACTTAGTC
25	COR018	Hopman *et al.* (1999)	7	AF083461	AGTCTGGCAATATTGAGGATGT AGCAGCTACCCTTTGAATACTG
25	NVHEQ43	Roed *et al.* (1998)	8	AF056396	TGACACAAGATAAAAGCCCCAGG GATTGGGAAAAGAGCACAGCC
26	ABGe126	Mittmann *et al.* (2010)	10	AM947005	AACCCTGAAATAACCAAAGTGC CGCTTTGAAAGAGCTTTTACTCC
26	ABGe124	Mittmann *et al.* (2010)	11	AM947003	TAACACAAAGCCCCCAGTTG GCCAAACCACACATGAGAGC
26	UMNe547	Wagner *et al.* (2004)	11		CTGCCATTCATCAGAAAATCTC TCAATCATTTTGTACTCATGGC
26	TKY523	Tozaki *et al.* (2004)	9	AB103741	TGCACACCCATTCTAGCTCA GTGGCTCACTCCTCGCTTAC
27	TKY315	Tozaki *et al.* (2000)	8	AB034624	GATGCCTCGAACTAGCTTG GATCTTCCATGTTTTTGTTTGG
27	COR017	Hopman *et al.* (1999)	7	AF083460	GAAGGCCTGAAGCATTTACA CGTAATGTTGACCAAACTTCA
28	TKY333	Tozaki *et al.* (2001)	9	AB044834	CCTTCACTAGCCTTCAAATG TTGTGTTTAGACAGTGCTGC

28	UCDEQ425	Eggleston-Stott *et al.* (1997)	10	U67406	AGCTGCCTCGTTAATTCA CTCATGTCCGCTTGTCTC
29	TKY715	Tozaki *et al.* (2004)	8	AB103933	CAGTTTCACAGGAGAGAGAGTCC CTGGAGTCCCACCTCCAAC
29	COR027	Murphie *et al.* (1999)	9	AF101396	CAGCTCTGCAATTTCTCCTC AATGACCAAGGCATTGAAAG
29	TKY478	Tozaki *et al.* (2004)	9	AB103696	GCCTGGGTACCTTTGTTGAA GGAACAGAATGGGAGTCCAG
30	LEX025	Coogle *et al.* (1996b)	8	AF075627	CAATCGTGGCCCGGTAAC TTCACTCCAATCCTCAGTCA
30	LEX075	Chowdhary *et al.* (2003)	7	AF213361	TCTGAAAAGTTGCAGTTTGAGAA TACAGTGTATTGGGGGCACA
31	AHT33	Swinburne *et al.* (2000)	12	AJ271518	CTGAGGGCGTAAGTCGAGTC GTTAATAGGAGCGGTTGTTTGG
31	ABGe241	Mittmann *et al.* (2010)	13	FM179521	AAAACCAGTCATGCGGAATC TGAGCTTGTTCCTGCTAGGG
X	LEX027	Coogle *et al.* (1996b)	7	AF075629	ACCACTGGGAAACTGTGTAA GCCCAGAATCCGAACC
X	LEX010	Coogle *et al.* (1996a)	9	AF075613	TGGGCTAAAATTTAATTTGGG ACCAAAACATATGCAAATTAA
X	AHT28	Swinburne *et al.* (2000)	15	AJ271513	CCTGGCTTATAGATGGCTGC ATTTGGAGATGGGGGTCTTT
X	TKY598	Tozaki *et al.* (2004)	8	TKY598	CGCTCAGGACCCTCCTACTT CGGCACCAACAAGCTTTAAT
X	TKY754	Tozaki *et al.* (2004)	9	TKY754	GTCCAAATCTGCCCAACAAG GCATCTGAGCTAGGAAGCTG

References

Bailey E., Skow L., Bernoco D., DelValle A., Scavone M.D., Bowling A.T. & Murray J.D. (2000) Equine dinucleotide repeat loci LEX071 through LEX078. *Anim Genet* **31**, 286-7.

Bjornstad G., Midthjell L. & Roed K.H. (2000) Characterization of ten equine dinucleotide microsatellite loci: NVHEQ21, NVHEQ54, NVHEQ67, NVHEQ70, NVHEQ75, NVHEQ77, NVHEQ79, NVHEQ81, NVHEQ82 and NVHEQ83. *Anim Genet* **31**, 78-9.

Breen M., Lindgren G., Binns M.M., Norman J., Irvin Z., Bell K., Sandberg K. & Ellegren H. (1997) Genetical and physical assignments of equine microsatellites--first integration of anchored markers in horse genome mapping. *Mamm Genome* **8**, 267-73.

Chowdhary B.P., Raudsepp T., Kata S.R., Goh G., Millon L.V., Allan V., Piumi F., Guerin G., Swinburne J., Binns M., Lear T.L., Mickelson J., Murray J., Antczak D.F., Womack J.E. & Skow L.C. (2003) The first-generation whole-genome radiation hybrid map in the horse identifies conserved segments in human and mouse genomes. *Genome Res* **13**, 742-51.

Coogle L. & Bailey E. (1997) Equine dinucleotide repeat loci LEX049-LEX063. *Anim Genet* **28**, 378.

Coogle L. & Bailey E. (1999) Equine dinucleotide repeat loci LEX064 through LEX070. *Anim Genet* **30**, 71-2.

Coogle L., Bailey E., Reid R. & Russ M. (1996a) Equine dinucleotide repeat polymorphisms at loci LEX002, -003, -004, -005, -007, -008, -009, -010, -011, -013 and -014. *Anim Genet* **27**, 126-7.

Coogle L., Reid R. & Bailey E. (1996b) Equine dinucleotide repeat loci from LEX025 to LEX033. *Anim Genet* **27**, 289-90.

Coogle L., Reid R. & Bailey E. (1997) Equine dinucleotide repeat loci LEX034-LEX048. *Anim Genet* **28**, 309.

Eggleston-Stott M.L., DelValle A., Bautista M., Dileanis S., Wictum E. & Bowling A.T. (1997) Nine equine dinucleotide repeats at microsatellite loci UCDEQ136, UCDEQ405, UCDEQ412, UCDEQ425, UCDEQ437, UCDEQ467, UCDEQ487, UCDEQ502 and UCDEQ505. *Anim Genet* **28**, 370-1.

George L.A., Miller L.M., Valberg S.J. & Mickelson J.R. (1998) Fourteen new polymorphic equine microsatellites. *Anim Genet* **29**, 469-70.

Godard S., Vaiman D., Oustry A., Nocart M., Bertaud M., Guzylack S., Meriaux J.C., Cribiu E.P. & Guerin G. (1997) Characterization, genetic and physical mapping analysis of 36 horse plasmid and cosmid-derived microsatellites. *Mamm Genome* **8**, 745-50.

Guerin G. & Bertaud M. (1996) Characterization of two polymorphic horse microsatellites: HMS15 and HMS20. *Anim Genet* **27**, 123.

Hirota K., Tozaki T., Mashima S. & Miura N. (2001) Cytogenetic assignment and genetic characterization of the horse microsatellites, TKY4-18, TKY20, TKY22-24, TKY30-41 derived from a cosmid library. *Anim Genet* **32**, 160-2.

Hopman T.J., Han E.B., Story M.R., Schug M.D., Aquadro C.F., Bowling A.T., Murray J.D., Caetano A.R. & Antczak D.F. (1999) Equine dinucleotide repeat loci COR001-COR020. *Anim Genet* **30**, 225-6.

Kakoi H., Tozaki T., Hirota K. & Mashima S. (1999) Genetic polymorphisms of equine microsatellite loci: TKY16, TKY19 and TKY21. *Anim Genet* **30**, 68-9.

Lampe V., Dierks C. & Distl O. (2009) Refinement of a quantitative trait locus on equine chromosome 5 responsible for fetlock osteochondrosis in Hanoverian warmblood horses. *Anim Genet* **40**, 553-5.

Lindgren G., Persson H. & Ellegren H. (1999) Five equine dinucleotide microsatellite loci HTG17, HTG20, HTG21, HTG28 and HTG31. *Anim Genet* **30**, 70-1.

Marklund S., Ellegren H., Eriksson S., Sandberg K. & Andersson L. (1994) Parentage testing and linkage analysis in the horse using a set of highly polymorphic microsatellites. *Anim Genet* **25**, 19-23.

Marti E., Breen M., Fischer P., Swinburne J. & Binns M.M. (1998) Six new cosmid derived and physically mapped equine dinucleotide repeat microsatellites. *Anim Genet* **29**, 236-8.

Mashima S., Kakoi H. & Tozaki T. (1999) TKY101: a highly polymorphic equine dinucleotide repeat locus. *Anim Genet* **30**, 163.

Meyer A.H., Valberg S.J., Hillers K.R., Schweitzer J.K. & Mickelson J.R. (1997) Sixteen new polymorphic equine microsatellites. *Anim Genet* **28**, 69-70.

Mickelson J.R., Wu J.T., Morrison L.Y., Swinburne J.E., Binns M.M., Reed K.M. & Alexander L.J. (2003) Eighty-three previously unreported equine microsatellite loci. *Anim Genet* **34**, 71-4.

Mittmann E.H., Lampe V., Momke S., Zeitz A. & Distl O. (2010) Characterization of a minimal microsatellite set for whole genome scans informative in warmblood and coldblood horse breeds. *J Hered* **101**, 246-50.

Murphie A.M., Hopman T.J., Schug M.D., Aquadro C.F., Bowling A.T., Murray J.D., Caetano A.R. & Antczak D.F. (1999) Equine dinucleotide repeat loci COR021-COR040. *Anim Genet* **30**, 235-7.

Penedo M.C., Millon L.V., Bernoco D., Bailey E., Binns M., Cholewinski G., Ellis N., Flynn J., Gralak B., Guthrie A., Hasegawa T., Lindgren G., Lyons L.A., Roed K.H., Swinburne J.E. & Tozaki T. (2005) International Equine Gene Mapping Workshop Report: a comprehensive linkage map constructed with data from new markers and by merging four mapping resources. *Cytogenetic and genome research* **111**, 5-15.

Roed K.H., Midthjell L. & Bjornstad G. (1998) Eight new equine dinucleotide repeat microsatellites at the NVHEQ26, NVHEQ29, NVHEQ31, NVHEQ40, NVHEQ43, NVHEQ90, NVHEQ98 and NVHEQ100 loci. *Anim Genet* **29**, 470.

Ruth L.S., Hopman T.J., Schug M.D., Aquadro C.F., Bowling A.T., Murray J.D., Caetano A.R. & Antczak D.F. (1999) Equine dinucleotide repeat loci COR041-COR060. *Anim Genet* **30**, 320-1.

Sakagami M., Tozaki T., Mashima S., Hirota K. & Mukoyama H. (1995) Equine parentage testing by microsatellite locus at chromosome 1q2.1. *Anim Genet* **26**, 123-4.

Swinburne J., Gerstenberg C., Breen M., Aldridge V., Lockhart L., Marti E., Antczak D., Eggleston-Stott M., Bailey E., Mickelson J., Roed K., Lindgren G., von Haeringen W., Guerin G., Bjarnason J., Allen T. & Binns M. (2000) First comprehensive low-density horse linkage map based on two 3-generation, full-sibling, cross-bred horse reference families. *Genomics* **66**, 123-34.

Swinburne J.E., Boursnell M., Hill G., Pettitt L., Allen T., Chowdhary B., Hasegawa T., Kurosawa M., Leeb T., Mashima S., Mickelson J.R., Raudsepp T., Tozaki T. & Binns M. (2006) Single linkage group per chromosome genetic linkage map for the horse, based on two three-generation, full-sibling, crossbred horse reference families. *Genomics* **87**, 1-29.

Swinburne J.E., Turner A., Alexander L.J., Mickleson J.R. & Binns M.M. (2003) Characterization and linkage map assignments for 61 new horse microsatellite loci (AHT49-109). *Anim Genet* **34**, 65-8.

Tallmadge R.L., Evans K.G., Hopman T.J., Schug M.D., Aquadro C.F., Bowling A.T., Murray J.D., Caetano A.R. & Antczak D.F. (1999a) Equine dinucleotide repeat loci COR081-COR100. *Anim Genet* **30**, 470-1.

Tallmadge R.L., Hopman T.J., Schug M.D., Aquadro C.F., Bowling A.T., Murray J.D., Caetano A.R. & Antczak D.F. (1999b) Equine dinucleotide repeat loci COR061-COR080. *Anim Genet* **30**, 462-3.

Tozaki T., Kakoi H., Mashima S., Hirota K., Hasegawa T., Ishida N., Miura N. & Tomita M. (2000) The isolation and characterization of 34 equine microsatellite loci, TKY290-TKY323. *Anim Genet* **31**, 234-6.

Tozaki T., Mashima S., Hirota K., Miura N., Choi-Miura N.H. & Tomita M. (2001) Characterization of equine microsatellites and microsatellite-linked repetitive elements (eMLREs) by efficient cloning and genotyping methods. *DNA Res* **8**, 33-45.

Tozaki T., Penedo M.C., Oliveira R.P., Katz J.P., Millon L.V., Ward T., Pettigrew D.C., Brault L.S., Tomita M., Kurosawa M., Hasegawa T. & Hirota K. (2004) Isolation, characterization and chromosome assignment of 341 newly isolated equine TKY microsatellite markers. *Anim Genet* **35**, 487-96.

Tozaki T., Swinburne J., Hirota K., Hasegawa T., Ishida N. & Tobe T. (2007) Improved resolution of the comparative horse-human map: investigating markers with in silico and linkage mapping approaches. *Gene* **392**, 181-6.

Wagner M.L., Goh G., Wu J.T., Raudsepp T., Morrison L.Y., Alexander L.J., Skow L.S., Chowdhary B.P. & Mickelson J.R. (2004) Radiation hybrid mapping of 75 previously unreported equine microsatellite loci. *Anim Genet* **35**, 68-71.

Supplementary Table 2 - Used microsatellite panel (N=50), chromosome location (Chr), accession number, and their primer sequences. * Microsatellites used in parentage testing at the Applied Genetics Laboratory (cE3c, Lisbon)

Chr	Marker	Reference	Accession number	Primer sequence (5' – 3')
1	AHT058	Swinburne *et al.* (2003)	AJ507675	CAGTGATGAGCCGCAAATAG TCTACCTATAATCCGCCTCCC
1	NVHEQ100	Roed *et al.* (1998)	AF056399	CCAAAGCAGAACATGTGAAGTT TGGCATAGATGTTAGCTCAGTGA
1	HMS007*	Swinburne *et al.* (2000)	X74636	CAGGAAACTCATGTTGATACCATC TGTTGTTGAAACATACCTTGACTGT
1	LEX020*	Coogle *et al.* (1996b)	AF075622	GGAATAGGTGGGGGTCTGTT AGGGTACTAGCCAAGTGACTGC
2	COR065	Tallmadge *et al.* (1999b)	AF142602	CAAAAGCACACACAAAGTGC TCCGGAAAGTGCAAAGTTAG
2	TKY384	Tozaki *et al.* (2004)	AB048290	TGCAGCAAGAAACCTAAACA CTTCAGTTGTAATCAGGCTC
3	UMNe158	Wagner *et al.* (2004)	AY391305	AATTGAGAGCCAAGATGACACC GGCACCATTTGAGGAAGATG
4	COR089	Tallmadge *et al.* (1999a)	AF154942	CCTGCCATAAATTTGTTTCC TCCCTACCTCATCTCCACAC
4	HMS006*	Breen *et al.* (1997)	X74635	GAAGCTGCCAGTATTCAACCATTG CTCCATCTTGTGAAGTGTAACTCA
5	AHT107	Swinburne *et al.* (2003)	AJ507724	ATCTAACCAGAGCGCAACGT CCCGACACAGAAGATGGG
6	TKY1001	Tozaki *et al.* (2004)	AB104219	TCTCAGAAGCCATCTGGAG ATCGATGCAGAACACGTGG
6	TKY412	Tozaki *et al.* (2004)	AB103630	GTGTGGGACAGGAAGTTTGG ATTCTTGGGTCCCCTCATCT
7	TKY034	Hirota *et al.* (2001)	AB048340	GCCACTGTAATGGTCTGAAG GGAAGGTAGAGCATTCTCAG
8	COR012	Hopman *et al.* (1999)	AF083455	TCTAGGAAAGACCCATCACG AGTAAGTGGAGGCCAAGGAT
8	COR003	Hopman *et al.* (1999)	AF083446	TAGGGAAACTCCTCAAAGCC GAAACCAAAACCTTCATCCA
8	AHT005*	Swinburne *et al.* (2000)		ACGGACACATCCCTGCCTGC GCAGGCTAAGGGGGCTCAGC
8	LEX023*	Coogle *et al.* (1996b)	AF075625	GGATGAAACAGGGAAGGAAA CCAACGGATTCATGAAAGCTA
9	HMS003*	Swinburne *et al.* (2000)	FJ915131	CCAACTCTTTGTCACATAACAAGA CCATCCTCACTTTTTCACTTTGTT
9	HTG004*	Ellegren *et al.* (1992)	AF169165	CTATCTCAGTCTTCATTGCAGGAC CTCCCTCCCTCCCTCTGTTCTC
10	ASB009	Breen *et al.* (1997)	X93523	GTGCGCATGTATGTGCGTGCC ATTTCCACAAGGGACATGAGG
11	ABGe099	Mittmann *et al.* (2010)	AM946382	TTCCTTCTGATTGCACCACTC ATTGTGGGTGACTCCCTCTG

11	SGCV24	Godard *et al.* (1997)	U90602	CTACCATTGAAGAGGGGTGGC GAAACGAGCAGGAAGTGAATCTCC
12	COR058	Ruth *et al.* (1999)	AF108375	GGGAAGGACGATGAGTGAC CACCAGGCTAAGTAGCCAAAG
15	ASB002*	Breen *et al.* (1997)	X93516	CCTTCCGTAGTTTAAGCTTCTG CACAACTGAGTTCTCTGATAGG
15	HTG006*	Breen *et al.* (1997)		CCTGCTTGGAGGCTGTGATAAGAT GTTCACTGAATGTCAAATTCTGCT
17	COR105	Swinburne *et al.* (2000)		TTTCCTCATTGCTTCCTGAG CCCAAGGTCTGTCTTGCTCTC
17	COR007	Hopman *et al.* (1999)	AF083450	GTGTTGGATGAAGCGAATGA GACTTGCCTGGCTTTGAGTC
18	ABGe151	Mittmann *et al.* (2010)	FM177589	CTCACTCTGGGCCCACTATC CGGAGTGAGAAGACAGTCCAG
18	TKY741	Tozaki *et al.* (2004)	AB103959	CCTTCCTTCTCCTAACTCAGTCC TGGAAACCAGGAATAGGTGTG
19	TKY448	Tozaki *et al.* (2004)	AB103666	TTCTTTTGCTCTCCCTCTCG GGAAAGACAGAGTAAGTGCGTGT
19	LEX036*	Coogle *et al.* (1997)	AF075638	ATCAGCCCAGCCTCTTCA AACAACCGGCNAAATAGTGC
19	COR062	Tallmadge *et al.* (1999b)	AF142599	GTCATCCAGTGACGAACACA AGGAAGTGCGCAGTAGAGAA
20	UM011	Meyer *et al.* (1997)	AF195130	TGAAAGTAGAAAGGGATGTGG TCTCAGAGCAGAAGTCCCTG
20	TKY477	Tozaki *et al.* (2004)	AB103695	CCCCTCTCTCTCTCAAGTGC CTCCTGGGTGGGAGAACTTT
20	TKY321	Tozaki *et al.* (2000)	AB034629	CACTGTGTAACACTAACACC TGTGACTTCAAGAACAGACG
21	TKY806	Tozaki *et al.* (2004)	AB104024	TGGAACTGTGATGATGTTGC TCTTTCTTCCCTTCCGAGAG
21	COR073	Tallmadge *et al.* (1999b)	AF142610	GCCAAGACATGGAAACAATC GTTCTCAAGGTGCATCCCTA
21	HTG10*	Marklund *et al.* (1994)	AF169294	CAATTCCCGCCCCACCCCCGGCA TTTTTATTCTGATCTGTCACATTT
21	TKY623	Tozaki *et al.* (2004)	AB103841	CAGTGTGGGTGGGCTTTATC ACCACTAGGGTGTGCATGTG
22	ABGe121	Mittmann *et al.* (2010)	AM947000	AGGAGCTGGAACTGACACAG GCTTCTCAGGGCAGTATTCC
23	TKY568	Tozaki *et al.* (2004)	AB103786	TTCCTGACGTGAAGGCATTA TGCCCTTCCTGCCTAGTAGA
24	AHT004*	Binns *et al.* (1995)	Y07733	AACCGCCTGAGCAAGGAAGT CCCAGAGAGTTTACCCT
25	NVHEQ43	Roed *et al.* (1998)	AF056396	TGACACAAGATAAAAGCCCCAGG GATTGGGAAAAGAGCACAGCC
25	UCDEQ405 *	Eggleston-Stott *et al.* (1997)	AF000010	ACCTCGTCTGGCTGTTGTAAG ACTTGCTGTGCGACTCTG

26	TKY523	Tozaki *et al.* (2004)	AB103741	TGCACACCCATTCTAGCTCA GTGGCTCACTCCTCGCTTAC
27	TKY315	Tozaki *et al.* (2000)	AB034624	GATGCCTCGAACTAGCTTG GATCTTCCATGTTTTTGTTTGG
27	UCDEQ005*	Eggleston-Stott *et al.* (1996)	U35423	AGCGGAAGTGCTGCGAAAG CCAGCATCTCTGGGCAGG
29	TKY478	Tozaki *et al.* (2004)	AB103696	GCCTGGGTACCTTTGTTGAA GGAACAGAATGGGAGTCCAG
30	VHL20*	van Haeringen *et al.* (1994)	X75970	CAAGTCCTCTTACTTGAAGACTAG AACTCAGGGAGAATCTTCCTCAG
31	ΛBGe241	Mittmann *et al.* (2010)	FM179521	AAAACCAGTCATGCGGAATC TGAGCTTGTTCCTGCTAGGG
X	LEX003*	Kakoi *et al.* (2005)	AF075607	ACATCTAACCAGTGCTGAGACT AAGAACTAGAACCTACAACTAGG
X	LEX024*	Coogle *et al.* (1996b)	AF075626	GGGGGTAGAGGGAAAAAGAG TTGTTGGCAGATCCCAGG
X	LEX027*	Coogle *et al.* (1996a)	AF075629	ACCACTGGGAAACTGTGTAA GCCCAGAATCCGAACC
X	TKY038*	Kakoi *et al.* (2005)	AB048344	TAAGTATTCTCATAAACGGG GGAATAATAACAGCATCCTC
Y	YA16*	Kakoi *et al.* (2005)		TGACTGGAAATTGAAGATG TTGTAGCAACAAAGTAACAC
Y	YM2*	Kakoi *et al.* (2005)		TGGTTCAGATGGTGTATTTTGTT TTTGCAGCCAGTACCTACCTT

References

Binns M.M., Holmes N.G., Holliman A. & Scott A.M. (1995) The identification of polymorphic microsatellite loci in the horse and their use in thoroughbred parentage testing. *Br Vet J* **151**, 9-15.

Breen M., Lindgren G., Binns M.M., Norman J., Irvin Z., Bell K., Sandberg K. & Ellegren H. (1997) Genetical and physical assignments of equine microsatellites--first integration of anchored markers in horse genome mapping. *Mamm Genome* **8**, 267-73.

Coogle L., Reid R. & Bailey E. (1996a) Equine dinucleotide repeat loci from LEX025 to LEX033. *Anim Genet* **27**, 289-90.

Coogle L., Reid R. & Bailey E. (1996b) Equine dinucleotide repeat loci LEX015-LEX024. *Anim Genet* **27**, 217-8.

Coogle L., Reid R. & Bailey E. (1997) Equine dinucleotide repeat loci LEX034-LEX048. *Anim Genet* **28**, 309.

Eggleston-Stott M.L., DelValle A., Bautista M., Dileanis S., Wictum E. & Bowling A.T. (1997) Nine equine dinucleotide repeats at microsatellite loci UCDEQ136, UCDEQ405, UCDEQ412, UCDEQ425, UCDEQ437, UCDEQ467, UCDEQ487, UCDEQ502 and UCDEQ505. *Anim Genet* **28**, 370-1.

Eggleston-Stott M.L., DelValle A., Bowling A.T., Bautista M., Zahorchak R. & Malyj W. (1996) Four equine dinucleotide repeats at microsatellite loci UCDEQ5, UCDEQ14, UCDEQ46 and UCDEQ62. *Anim Genet* **27**, 129.

Ellegren H., Johansson M., Sandberg K. & Andersson L. (1992) Cloning of highly polymorphic microsatellites in the horse. *Anim Genet* **23**, 133-42.

Godard S., Vaiman D., Oustry A., Nocart M., Bertaud M., Guzylack S., Meriaux J.C., Cribiu E.P. & Guerin G. (1997) Characterization, genetic and physical mapping analysis of 36 horse plasmid and cosmid-derived microsatellites. *Mamm Genome* **8**, 745-50.

Hirota K., Tozaki T., Mashima S. & Miura N. (2001) Cytogenetic assignment and genetic characterization of the horse microsatellites, TKY4-18, TKY20, TKY22-24, TKY30-41 derived from a cosmid library. *Anim Genet* **32**, 160-2.

Hopman T.J., Han E.B., Story M.R., Schug M.D., Aquadro C.F., Bowling A.T., Murray J.D., Caetano A.R. & Antczak D.F. (1999) Equine dinucleotide repeat loci COR001-COR020. *Anim Genet* **30**, 225-6.

Kakoi H., Hirota K., Gawahara H., Kurosawa M. & Kuwajima M. (2005) Genetic diagnosis of sex chromosome aberrations in horses based on parentage test by microsatellite DNA and analysis of X- and Y-linked markers. *Equine Vet J* **37**, 143-7.

Marklund S., Ellegren H., Eriksson S., Sandberg K. & Andersson L. (1994) Parentage testing and linkage analysis in the horse using a set of highly polymorphic microsatellites. *Anim Genet* **25**, 19-23.

Meyer A.H., Valberg S.J., Hillers K.R., Schweitzer J.K. & Mickelson J.R. (1997) Sixteen new polymorphic equine microsatellites. *Anim Genet* **28**, 69-70.

Mittmann E.H., Lampe V., Momke S., Zeitz A. & Distl O. (2010) Characterization of a minimal microsatellite set for whole genome scans informative in warmblood and coldblood horse breeds. *J Hered* **101**, 246-50.

Roed K.H., Midthjell L. & Bjornstad G. (1998) Eight new equine dinucleotide repeat microsatellites at the NVHEQ26, NVHEQ29, NVHEQ31, NVHEQ40, NVHEQ43, NVHEQ90, NVHEQ98 and NVHEQ100 loci. *Anim Genet* **29**, 470.

Ruth L.S., Hopman T.J., Schug M.D., Aquadro C.F., Bowling A.T., Murray J.D., Caetano A.R. & Antczak D.F. (1999) Equine dinucleotide repeat loci COR041-COR060. *Anim Genet* **30**, 320-1.

Swinburne J., Gerstenberg C., Breen M., Aldridge V., Lockhart L., Marti E., Antczak D., Eggleston-Stott M., Bailey E., Mickelson J., Roed K., Lindgren G., von Haeringen W., Guerin G., Bjarnason J., Allen T. & Binns M. (2000) First comprehensive low-density horse linkage map based on two 3-generation, full-sibling, cross-bred horse reference families. *Genomics* **66**, 123-34.

Swinburne J.E., Turner A., Alexander L.J., Mickleson J.R. & Binns M.M. (2003) Characterization and linkage map assignments for 61 new horse microsatellite loci (AHT49-109). *Anim Genet* **34**, 65-8.

Tallmadge R.L., Evans K.G., Hopman T.J., Schug M.D., Aquadro C.F., Bowling A.T., Murray J.D., Caetano A.R. & Antczak D.F. (1999a) Equine dinucleotide repeat loci COR081-COR100. *Anim Genet* **30**, 470-1.

Tallmadge R.L., Hopman T.J., Schug M.D., Aquadro C.F., Bowling A.T., Murray J.D., Caetano A.R. & Antczak D.F. (1999b) Equine dinucleotide repeat loci COR061-COR080. *Anim Genet* **30**, 462-3.

Tozaki T., Kakoi H., Mashima S., Hirota K., Hasegawa T., Ishida N., Miura N. & Tomita M. (2000) The isolation and characterization of 34 equine microsatellite loci, TKY290-TKY323. *Anim Genet* **31**, 234-6.

Tozaki T., Penedo M.C., Oliveira R.P., Katz J.P., Millon L.V., Ward T., Pettigrew D.C., Brault L.S., Tomita M., Kurosawa M., Hasegawa T. & Hirota K. (2004) Isolation, characterization and chromosome assignment of 341 newly isolated equine TKY microsatellite markers. *Anim Genet* **35**, 487-96.

van Haeringen H., Bowling A.T., Stott M.L., Lenstra J.A. & Zwaagstra K.A. (1994) A highly polymorphic horse microsatellite locus: VHL20. *Anim Genet* **25**, 207.

Wagner M.L., Goh G., Wu J.T., Raudsepp T., Morrison L.Y., Alexander L.J., Skow L.S., Chowdhary B.P. & Mickelson J.R. (2004) Radiation hybrid mapping of 75 previously unreported equine microsatellite loci. *Anim Genet* **35**, 68-71.

Supplementary Table 3 - Inbreeding (by pedigree), d^2 and heterozygosity average ± standard deviation for 50 autosomal mSats in the whole population (TOTAL) and by subpopulation (Portugal = PT; Germany = GER) (N=190).

	Inbreeding	d^2	Heterozygosity
TOTAL	0.3816 ± 0.066	60.6449 ± 17.556	0.5708 ± 0.097
PT	0.3858 ± 0.070	58.7353 ± 17.549	0.5678 ± 0.101
GER	0.3671 ± 0.043	67.3742 ± 16.032	0.5812 ± 0.081

Supplementary Table 4 - Inbreeding (by pedigree), inbreeding (by SNP analysis) and individual heterozygosity average ± standard deviation for SNPs in the whole population (TOTAL) and by subpopulation (Portugal = PT; Germany = GER) (N=50).

	Inbreeding	Inbreeding by SNPs	Individual Heterozygosity
TOTAL	0.3393 ± 0.065	-0.0684 ± 0.084	0.3352 ± 0.026
PT	0.3401 ± 0.070	-0.0569 ± 0.079	0.3316 ± 0.025
GER	0.3355 ± 0.334	-0.1210 ± 0.089	0.3516 ± 0.028

Supplementary Table 5 - ROH description. ID - individual ID; CHR- Chromosome; SNP1- SNP at the start of region; SNP2 - SNP at the end of region; POS1- position (bp) of SNP1; POS2 - position (bp) of SNP2; bp - Region size (in base-pairs); NSNP - Number of SNPs in run.

ID	CHR	SNP1	SNP2	POS1	POS2	bp	NSNP
10	2	BIEC2_528259	BIEC2_504708	98670916	105334216	6,663,301	146
12	2	BIEC2_528259	BIEC2_530754	98670916	107289833	8,618,918	175
14	2	BIEC2_502494	BIEC2_504614	98631308	104315986	5,684,679	126
20	2	BIEC2_502173	BIEC2_533681	98151103	114208001	16,056,899	324
23	2	BIEC2_528259	BIEC2_504953	98670916	107182288	8,511,373	174
25	2	BIEC2_502494	BIEC2_533681	98631308	114208001	15,576,694	317
26	2	BIEC2_488944	BIEC2_491394	73338834	78264598	4,925,765	105
27	2	BIEC2_528259	BIEC2_530754	98670916	107289833	8,618,918	175
43	2	BIEC2_505832	BIEC2_508817	110669319	116608714	5,939,396	121
20	4	BIEC2_849249	BIEC2_896170	13029056	19850553	6,821,498	145
12	5	BIEC2_930374	BIEC2-934471	94482164	99631010	5,148,847	106
17	8	BIEC2_1035218	BIEC2_1095227	19006526	25207057	6,200,532	125
26	8	BIEC2_1050839	BIEC2_1059911	50183825	66728267	16,544,443	336
36	8	BIEC2_1055953	BIEC2_1061033	60064433	68451759	8,387,327	174
8	9	BIEC2_1102000	BIEC2_1165121	66152662	72133579	5,980,918	124
13	9	BIEC2_1102292	BIEC2_1165121	66493146	72133579	5,640,434	118
22	9	BIEC2_1162514	BIEC2_1165121	66397968	72133579	5,735,612	121
3	11	BIEC2_169263	BIEC2_165604	55547022	61282299	5,735,278	117
18	11	BIEC2_136373	BIEC2_139925	7097967	13564825	6,466,859	130
26	11	BIEC2-136211	BIEC2_146484	6145250	12190664	6,045,415	125
28	11	BIEC2_159977	BIEC2_165604	53420235	61282299	7,862,065	160
30	11	BIEC2_159977	BIEC2_165604	53420235	61282299	7,862,065	160
46	11	BIEC2-160382	BIEC2_165604	54257970	61282299	7,024,330	141
161	11	BIEC2-135170	BIEC2_140172	4551910	13888096	9,336,187	187
14	15	BIEC2_312377	BIEC2_336734	59448498	79362010	19,913,513	403
15	15	BIEC2_283521	UKUL2785	6253234	11451399	5,198,166	107
32	15	BIEC2_333776	BIEC2_339339	72389304	82706171	10,316,868	224
39	15	BIEC2_312831	BIEC2_321535	60498665	79116298	18,617,634	383
19	20	BIEC2_522389	BIEC2_524835	17280513	23353751	6,073,239	123
20	20	BIEC2_570540	BIEC2_545557	59235745	64165294	4,929,550	106
24	20	BIEC2_539123	BIEC2_545557	51863399	64165294	12,301,896	257
3	22	BIEC2-574399	BIEC2_577869	157780	5787206	5,629,427	120
9	22	BIEC2-574399	BIEC2-579150	157780	7530450	7,372,671	155
15	22	BIEC2-574399	BIEC2_578203	157780	6049945	5,892,166	126
33	22	BIEC2_604728	BIEC2_579959	2265074	8819939	6,554,866	134
40	22	BIEC2-574399	BIEC2_577869	157780	5787206	5,629,427	120
44	22	BIEC2-574399	BIEC2_578956	157780	7176739	7,018,960	148
20	24	BIEC2_629168	BIEC2_633071	1445929	11925619	10,479,691	218
43	24	BIEC2_628800	BIEC2_631224	510741	7520156	7,009,416	144
45	24	BIEC2_649920	BIEC2_652935	39045681	46706798	7,661,118	156
4	27	BIEC2_751382	BIEC2-721363	30896790	38384610	7,487,821	160
7	27	BIEC2_715750	BIEC2_723001	31633888	39932455	8,298,568	180

Supplementary Table 6 - Comparison of microsatellite (STRs) genetic variability in different horse breeds. Mean number of alleles (MNA), observed heterozygosity (H_O), unbiased expected heterozygosity (H_E), polymorphic information content (PIC), probability of paternity exclusion (PE), heterozygote deficiency coefficient (F_{IS}), mean individual heterozygosity (Het), mean d^2 (in bp^2), average inbreeding coefficients (F) number of autosomal STRs used (#STRs), number of samples genotyped (#samples).

Breed	MNA	H_o	H_e	PIC	PE	F_{IS}	Het	mean d^2	F	#STRs	# samples	Reference
Sorraia	3.7	0.573	0.586	0.526	1	0.023	0.57	60.6	0.382	50	190	*This study*
Sorraia	3.3	0.450	0.465	0.415	0.99	-0.0215	0.464	36.9	0.325	22	131	(Luis *et al.* 2007a)
Sorraia	3.83	0.506	0.487								60	
Garrano	7.17	0.779	0.745								37	
Lusitano	6.33	0.721	0.677								70	
Andalusian	5.67	0.727	0.693							12	33	(Luis *et al.* 2007b)
Friesian	4.50	0.466	0.454								159	
Lipizzan	6.33	0.724	0.700								40	
Thoroughbred	6.25	0.695	0.674								175	
Terceira Pony	7	0.700	0.674	0.630	0.99	−0.047				15	30	(Lopes *et al.* 2015)
Garrano	10,05	0,758	0.767	0,752		0,025			0.0065	22	60	(Cipriano 2007)
Sorraia	3.43	0.529	0.525							31	23	(Aberle *et al.* 2004)
Przewalski's Horse	3.83	0.468	0.526								18	
Andalusian	7.7	0.748	0.752			0.0060					60	
Retuertas pony	5.91	0.726	0.686			-0.0587				22	55	(Vega-Pla et al. 2006)
Asturcón pony	7.18	0.758	0.736			-0.0307					39	
Thoroughbred	4.7	0.628	0.646						0.13	13	211	(Cunningham *et al.* 2001)
Hanoverian Warmblood	7.3	0.71		0.65						322	311	(Mittmann *et al.* 2010)
German cold-blood	8	0.71		0.67							267	
Asturcón Pony	8.7	0.737				0.060				15	297	(Álvarez *et al.* 2011)

Zanskari pony	8.52	0.676	0.764	0.732		0.115				48	52	(Gupta *et al.* 2012)
Lipizzan	6.82						0.67	43.16	0.10	17	360	(Curik 2003)
Noriker draught	6.83	0.64								12	134	(Aurich *et al.* 2003)
Brazilian Criollo	13.6	0.470	0.821	0.798	0.99	0.432				8	2542	(Costa *et al.* 2010)
Saudi Arabian	5.13	0.68	0.68			0.008					33	
Syrian registered	6.47	0.70	0.69			-0.007					138	
Syrian non-reg	8.47	0.72	0.75			0.037					114	
Iranian Arabian	5.93	0.70	0.71			0.017					40	
Davenport Arabian	3.00	0.40	0.46			0.132				15	23	(Khanshour *et al.* 2013a)
Egyptian-Saudi mix	3.53	0.58	0.55			-0.066					28	
USA-Egyptian	4.00	0.53	0.56			0.047					47	
USA-Saudi	4.40	0.66	0.65			-0.015					57	
Shagya Arabian	4.93	0.68	0.69			0.005					21	
Polish Arabian	5.67	0.69	0.68			-0.015					36	
Pantaneiro	9.10	0.628	0.740	0.706	0.993	0.116				10	227	(Giacomoni *et al.* 2008)
Arabian from Syria-reg	5.69	0.694	0.712	0.657	0.9999					16	45	(Khanshour *et al.* 2013b)
Hucul	9.92	0.71	0.72			0.013			0.042	12	1627	(Mackowski et al. 2015)
Polish primitive	9.92	0.68	0.70			0.020			0.077		3865	
Andalusian									0.082			(Gómez *et al.* 2009)
Lusitano									0.099			(Vicente *et al.* 2014)
Friesian									0.157			(Sevinga *et al.* 2004)
Przewalskii									0.14			(Der Sarkissian *et al.* 2015)

Supplementary Table 7 - CNV overlaps (x) between Sorraia horses and other published horse CNVs. CHR- Chromosome; Start - start of region; Stop - end of region; Size - region size (in base-pairs).

Chr	Start	Stop	Size	Doan et al. (2012a)	Doan et al. (2012b)	Dupuis et al. (2013)	Metzger et al. (2013)	Wang at al. (2014)
chr4	23,686,504	23,809,130	122,626				x	
chr4	96,112,893	96,120,426	7,533					x
chr7	73,341,169	73,689,153	347,984	x		x		
chr8	85,933,859	86,060,633	126,774	x			x	
chr13	1,016,356	1,040,180	23,824	x			x	
chr20	27,715,495	27,726,675	11,180	x			x	x
chr27	60,642	290,954	230,312			x		
chr29	890,859	1,231,957	341,098		x	x		
chr30	12,461,123	12,467,768	6,645	x	x			
chr30	28,193,772	28,211,474	17,702	x			x	
chrUn	432,940	438,527	5,587			x		

Doan R., Cohen N., Harrington J., Veazey K., Juras R., Cothran G., McCue M.E., Skow L. & Dindot S.V. (2012a) Identification of copy number variants in horses. Genome Res 22, 899-907.

Doan R., Cohen N.D., Sawyer J., Ghaffari N., Johnson C.D. & Dindot S.V. (2012b) Whole-genome sequencing and genetic variant analysis of a Quarter Horse mare. BMC Genomics 13, 78.

Dupuis M.C., Zhang Z., Durkin K., Charlier C., Lekeux P. & Georges M. (2013) Detection of copy number variants in the horse genome and examination of their association with recurrent laryngeal neuropathy. Anim Genet 44, 206-8.

Metzger J., Philipp U., Lopes M.S., da Camara Machado A., Felicetti M., Silvestrelli M. & Distl O. (2013) Analysis of copy number variants by three detection algorithms and their association with body size in horses. BMC Genomics 14, 487.

Wang W., Wang S., Hou C., Xing Y., Cao J., Wu K., Liu C., Zhang D., Zhang L., Zhang Y. & Zhou H. (2014) Genome-Wide Detection of Copy Number Variations among Diverse Horse Breeds by Array CGH. PloS one 9, e86860.

CHAPTER 3 - REPRODUCTIVE FITNESS

Paper 4 - Kjöllerström H.J., Gama L.T. & Oom M.M. (2015) Impact of inbreeding on fitness-related traits in the highly threatened Sorraia horse breed. Livestock Science 180, 84-9.

Paper 5 - Kjöllerström H.J., Collares-Pereira M.J. & Oom M.M. (2011) First evidence of sex chromosome mosaicism in the endangered Sorraia Horse breed. Livestock Science 136, 273-6.

Paper 6 - Kjöllerström H.J., Oom M.M., Chowdhary, B.P & Radsepp, T. (2016) Fertility assessment in Sorraia stallions by sperm-FISH and *FKBP6* genotyping. Reproduction in Domestic Animals.

CHAPTER 3 - PAPER 4

Impact of inbreeding on fitness-related traits in the highly threatened Sorraia horse breed.

Kjöllerström H.J., Gama L.T. & Oom M.M.

Livestock Science (2015) 180, 84-9

Livestock Science 180 (2015) 84–89

Contents lists available at ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

Impact of inbreeding on fitness-related traits in the highly threatened Sorraia horse breed

H.J. Kjöllerström [a], L.T. Gama [b], M.M. Oom [a,*]

[a] cE3c – Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal
[b] CIISA-Faculdade de Medicina Veterinária, Universidade de Lisboa, Avenida da Universidade Técnica, 1300-477 Lisboa, Portugal

ARTICLE INFO

Article history:
Received 13 March 2014
Received in revised form
31 July 2015
Accepted 1 August 2015

Keywords:
Sorraia horse
Inbreeding
Viability
Fertility
Conservation

ABSTRACT

The Sorraia horse population can be regarded as a universal equine genetic resource, most likely the representative of the ancestor of Iberian saddle horses and probably the ancestor of several New World horse breeds. The breed was recovered in 1937 and managed without further additions to the initial founder group of 12 horses, originating an extant population with extremely high inbreeding levels (mean F=0.38). There are only about 300 animals representing the Sorraia horse population worldwide, which places it in a critical risk status and strongly supports the need to establish a conservation-breeding plan aiming at a long-term self-sustaining population. Data on all registered horses in the Sorraia Studbook were used to study the impact of inbreeding on offspring's viability at birth and at 6 months of age, and a sub-sample was used to determine the influence of inbreeding on stallion and mare fertility rates, foaling intervals and age at first parturition. The effect of inbreeding on the analysed traits was only significant for the relationship of mare fertility with mare inbreeding coefficient (P=0.003). The influence of age of the mare was quadratic, with a reduction in foal mortality (both at birth and at six months of age) and an increase in foaling interval as age of mare increased. Stud farms had a statistically significant influence on age at first parturition. Decisive management-breeding plans must be taken to control inbreeding levels in Sorraia horses, and contribute to the conservation of this breed.

1. Introduction

Inbreeding is the mating of individuals that share one or more common ancestors. It is commonly represented by the inbreeding coefficient (F) which measures the probability that two alleles at a locus are identical by descent. As it increases homozygosity, inbreeding exposes rare deleterious alleles and is unavoidable in small closed populations, such as the Sorraia horse, as all individuals become related by descent over time (e.g. Lacy, 2000; Frankham et al., 2002) and may, in many cases, give rise to inbreeding depression (Charlesworth and Willis, 2009).

Inbreeding depression reflects the loss of fitness due to matings between relatives and is known to affect different aspects of biological systems, revealing important consequences for the conservation of small isolated populations, either natural or captive ones. The traits more affected by inbreeding depression are often those more closely related to fitness, such as reproductive characteristics and survival, conformation, growth and weight traits (Leroy, 2014).

The pernicious effects of inbreeding on juvenile survival have long been known in various species (Darwin, 1859). Furthermore, inbreeding depression has long been associated with decreased fitness in human-bred animals, for both domestic and wild species, and significantly corroborated since it was documented that inbreeding lowered juvenile survival in 41 out of 44 populations of captive ungulates (Ralls et al., 1979). Even though most evidence on the effects of inbreeding on fitness components in horses is on juvenile survival, inbred individuals that survive until adulthood may still suffer from inbreeding depression, expressed as lowered survival, growth rate, fertility, inability to mate, reduced fecundity, and insufficient parental care (e.g. Ryan et al., 2003; Charlesworth and Willis, 2009; Collins et al., 2012).

The Sorraia horse is most likely the representative of the ancestor of Iberian saddle horses (Andrade, 1945; Oom et al., 2004) and is strongly related to several New World horse breeds (Andrade, 1945; Bouman, 1989; Luís et al., 2006). This breed was recovered from 4 stallions and 7 mares in 1937 and a male brought in 1948 from Argentina (Oom et al., 2004). In 1976, a subpopulation was exported to Germany, with no further immigration until recently (Luís et al., 2007). Traditionally, Sorraia mares are

* Corresponding author. Fax: +351 217500028.
E-mail address: mmoom@fc.ul.pt (M.M. Oom).

http://dx.doi.org/10.1016/j.livsci.2015.08.001
1871-1413/

managed extensively and the stallions are turned into the pasture to cover the mares during the breeding season (usually February to June) (Oom et al., 2004), as they have shown lower breeding performance with in-hand mating.

The worldwide population of the Sorraia horse consists of only about 300 animals which, according to FAO criteria, correspond to a critical maintained risk status (FAO, 1998). This status of the population requires a conservation-breeding plan to establish a long-term self-sustaining population by adopting appropriate breeding strategies. The small number of founders, reduced effective population size and complete genetic isolation of the Sorraia breed, has led to extremely high levels of inbreeding, with a mean F of 0.38 in the current population (F ranging from 0.22 to 0.60) and an effective population size of 11.59 in the last 8.67 generations (Pinheiro et al., 2013). This makes the Sorraia horse a very interesting population to assess the effects of inbreeding in horses, as it exceeds the levels of inbreeding reported, e.g. for the Lipizzan (mean F=0.10) (Curik, 2003), the Lusitano (mean F=0.11) (Vicente et al., 2012), Finnish Standardbred Trotters (mean F=0.10) and Finnhorses (mean F=0.04) (Sairanen et al., 2009) and even the endangered Przewalski horse (mean F=0.21) (Buisman and van Weeren, 1982). Moreover, the effective number of founders is 7.46, two of the founders are no longer represented in the living population and genetic contribution of underrepresented founders is at great risk of loss (Luís et al., 2007; Pinheiro et al., 2013).

Little is known about the reproductive performance of Sorraia horses, but one report suggests a fertility rate of about 57% (Oom et al., 1991), which is not that low when compared with the Lusitano (61%) and the Garrano (44%) (Gomes and Oom, 2000), the other two autochthonous Portuguese horse breeds. For horses, in general, the age at first parturition is usually between 5.5 and 6 years of age (Valera et al., 2000). In the Sorraia breed, the mares first reproduce at an average age of 4 years and 11 months and the stallions at 4 years and 6 months (Kjöllerström, 2005). Fecundity rates, according to some authors, are higher in horses between 5 and 15 years, with values ranging from 0.80 to 0.90 (Cothran et al., 1984) and the foaling rates tend to decrease as mares get older (Sairanen et al., 2009). The fertility peak of Sorraia mares (between 4 and 16 years) is much wider than that observed for stallions (from 10 to 14 years) (Kjöllerström, 2005) and of that described for horses in general, i.e., 4.2 years for Lusitano mares, 5.2 years for Pura Raza Española mares, and 4.2 years for Lusitano stallions (Valera et al., 2000).

Despite the extremely high levels of inbreeding in Sorraia horses, little is known about inbreeding depression in this breed. It is, thus, of relevant importance to determine the effect that such inbreeding levels might have on the fitness of the population, namely by a comprehensive knowledge of the likely causes of the low levels of fertility and viability observed. In the present study we investigate the impact of inbreeding depression in Sorraia horses, by studying the relationships existing between offspring and parental inbreeding and the viability of newborns, fertility in both sexes, foaling intervals and age at first parturition. The results of our analyses should allow a better understanding of the consequences of inbreeding in Sorraia horses and, in the long run, provide the bases for definition of a management-breeding programme in a way best suited to establish a long-term self-sustaining population for this highly endangered horse breed.

2. Material and methods

All Sorraia horses are registered in the Sorraia Studbook, with complete pedigrees tracing back to the founder's population, started in 1937 (Oom et al., 2004). Since 1992, a mandatory paternity test was required before registration in order to increase the validation of the genealogies. All Studbook data were entered into SPARKS v.1.6 (ISIS, 2011), a software developed for the compilation, edition, analysis and production of reports of captive populations.

2.1. Data

For the analysis of the effect of inbreeding in Sorraia horses on mortality at birth and at six months of age, data from 749 animals registered from 1937 until the end of 2010 were used (385 females, 364 males). For better consistency in data analyses, we did not include data from 2010 onwards, as they may be incomplete, thus underestimating fertility rates.

Mortality at birth grouped abortions, stillbirths and dead until 30 days of age, and mortality at six months included those animals dead between 31 days and six months of age. Animals dead in one of these categories were coded as 1, and as 0 if they survived, and these codes were the response variables.

Individual inbreeding coefficients were calculated with SPARKS v.1.6 (ISIS, 2011) based on the additive relationship matrix (Ballou, 1983) and considering complete pedigrees traced back to the founders (base population).

For the evaluation of the effect of inbreeding on fertility, only two stud farms considered to be more representative (with higher census) were included and these were coded as A and B (not named for ethical reasons), both in Portugal. These analyses were also restricted to the years for which information was more consistent (1980–2010). Fertility rate for mares was calculated as the number of foals produced relative to the number of years in reproduction, whereas for stallions it was the total number of foals sired relative to the number of mares available in the same stud. Information about mares (N=96) and stallions (N=24) was analysed separately.

Foaling intervals were calculated as the difference, in days, between birth dates of two consecutive foals produced by the same mare. These analyses were carried-out using all registered births in the same subset of data, from 1980 to 2010 in stud farms A and B (N=232). Age at first parturition was analysed in the subset used before (1980–2010, stud farms A and B), totalizing 91 mares.

2.2. Statistical analysis

Preliminary statistical analyses were performed with the packages JMP® Pro 10.0.2 (JMP®, 2013) and Microsoft Excel® 2010, with α set at 0.05.

The effect of inbreeding on mortality (at birth and six months), fertility of mares and stallions, foaling interval and age at first parturition in mares, was analysed with mixed model procedures, using a Best Linear Unbiased Prediction (BLUP) – Animal Model and the Wombat package (Meyer, 2007). Univariate analyses were carried out, considering repeated records in the analyses of foaling interval and single records in the remaining traits. The random effects considered were the additive genetic effect in all traits analysed, plus the permanent environmental effect of the mare in the analysis of foaling interval.

The fixed effects considered in the analyses of the various traits were chosen according to their biological expression, and are summarised in Table 1. Briefly, the effect of stud was considered in the analyses of all traits, and possible time trends were accounted for by including in the model year of foaling (in the case of foaling interval) or year of birth (in all the other traits). However, considering the highly unbalanced distribution of records among years, these were grouped in five-year intervals, and considered as such in the fixed factors. Sex of the foal was included as a fixed factor in the analysis of foal mortality and month of parturition in

Table 1
Fixed effects included in the mixed linear model for the various traits analysed.

Fixed effect	Trait analyzed					
	Foaling interval	Age at first parturition	Foal mortality at birth	Foal mortality up to 6 months	Mare fertility	Stallion fertility
Stud	x	x	x	x	x	x
Five-years period[a]	x	x	x	x	x	x
Sex of foal			x	x		
Month of foaling	x					
Age of mare (quadratic)	x		x	x		
F_{mare} (linear)	x	x	x	x	x	
F_{foal} (linear)			x	x		
$F_{stallion}$ (linear)						x

[a] Period of five years when the mare foaled (analyses of foaling interval) or when the animal was born (remaining traits).

the analyses of foaling interval, while the linear and quadratic effects of age of the mare were included in the analyses of foaling interval and foal mortality. The linear effect of inbreeding was assessed by including as a linear covariate the inbreeding coefficient of the mare in the analyses of all traits except sire fertility, plus the effect of inbreeding of the foal in the analysis of foal mortality. In the analysis of sire fertility, the only inbreeding considered was that of the stallion. The estimates of fixed effects and the corresponding standard error obtained in Wombat, in particular those referring to inbreeding depression, were used to compute approximate levels of significance, assuming a *t*-distribution, to allow an assessment of the magnitude and importance of inbreeding for the traits analysed.

3. Results

In the Sorraia horse population the average inbreeding coefficient per year of birth has increased steadily since the breed's foundation (Fig. 1A), and from 1960 onward all births are inbred. In 2012, the average inbreeding coefficient of all foals born was 0.42, and in the current active population the inbreeding coefficients are extremely high, with an average of 0.38 (minimum=0.22, maximum=0.60 and SD=0.07) (Fig. 1B).

The proportion of registered non-viable animals in the Sorraia breed is not alarming (Fig. 2), with the incidence of abortions and stillborns accounting for only about 2% of the total births (0.26% and 1.83%, respectively) during the period studied. Nevertheless, it is not easy to assess the importance of unreported cases of neonatal mortality and abortions, but these have certainly existed. On the other hand, the mean fertility rates observed in this study were 46.74% for stallions and 35.79% for mares (results not shown).

The heritability estimates obtained in our study (Table S1) were not very reliable, given the reduced number of observations included in the different analyses. Still, the estimated heritability was highest for fertility (0.78 ± 0.18 and 0.57 ± 0.46 for mares and stallions, respectively), close to zero for survival at birth and six months, and 0.04 ± 0.01 for foaling interval.

Our results indicate that there was some tendency for an increase in the inbreeding coefficient of a foal to cause an increase in the probability of death at birth and up to six months of age, but this tendency was not significant ($P > 0.05$). Somewhat surprisingly, an increase in mare inbreeding was associated with an increase in juvenile survival (Table 2), but again these results are not statistically significant ($P > 0.05$). Age of mare influenced foal survival, with a reduction in foal mortality at birth and up to 6 months as mares aged (Fig. 3).

In the A and B studs, the average female fertility was 0.40 ± 0.21 and 0.28 ± 0.19, whereas average mare inbreeding was

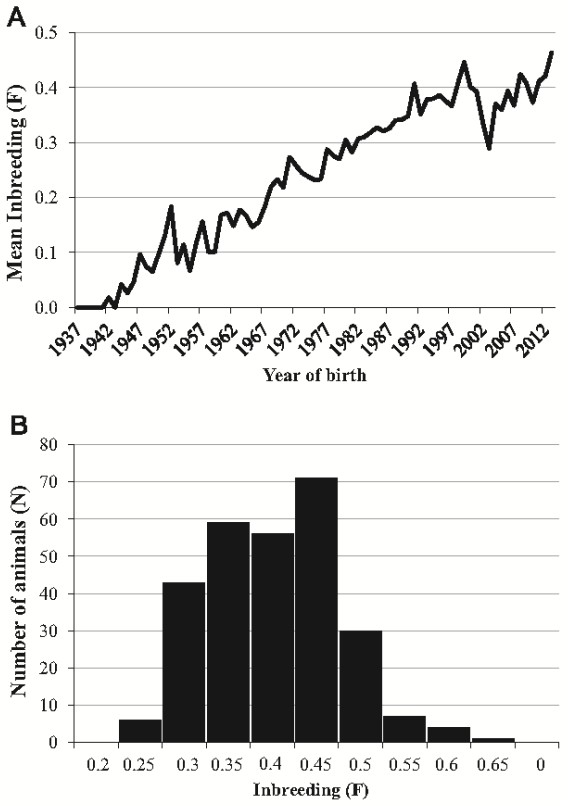


Fig. 1. (A) Evolution of the mean inbreeding coefficient (F) by year of birth, from the population's foundation until 2012 ($N=798$). (B) Distribution of the inbreeding coefficients of the current Sorraia horse population (mean=0.38 ± 0.07; $N=278$).

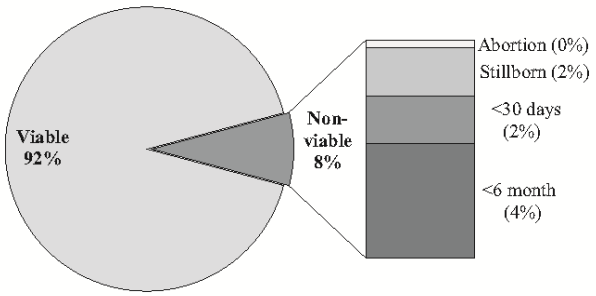


Fig. 2. Distribution (%) of offspring viability in the Sorraia horse population since its foundation until 2010 ($N=765$). The non-viable class is further partitioned into subclasses.

Table 2
Regression coefficient on inbreeding (per 1% F) ± standard error, and significance assessed by *t*-test [*P*(t)], for the traits analyzed in Sorraia horses.

	Inbreeding of mare		Inbreeding of foal		Inbreeding of stallion	
	Estimate	*P*(t)	Estimate	*P*(t)	Estimate	*P*(t)
Foaling interval (d)	2.68 ± 4.36	0.269	–	–	–	–
Age at first parturition (m)	0.34 ± 0.35	0.164	–	–	–	–
Foal mortality (birth) %	−0.149 ± 0.120	0.108	0.016 ± 0.113	0.444	–	–
Foal mortality (6 M) %	−0.083 ± 0.170	0.313	0.069 ± 0.162	0.335	–	–
Mare fertility (%)	−0.678 ± 0.365	0.032	–	–	–	–
Stallion fertility	–	–	–	–	−0.086 ± 0.437	0.422

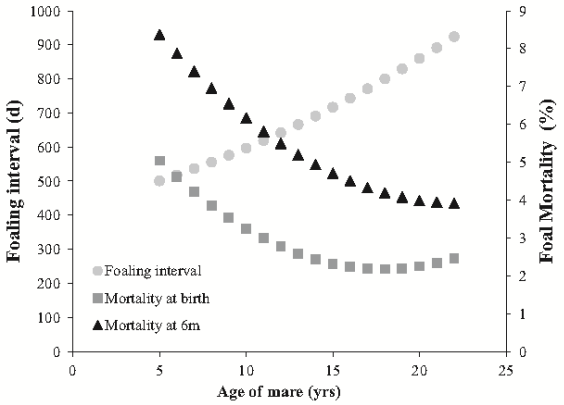


Fig. 3. Quadratic effect of age of the mare in: foal mortality (both at birth and at six months of age, in %) (N=749), and in parturition interval (in days) (N=232).

0.33 ± 0.09 and 0.34 ± 0.08, respectively. For A and B stallions, the average fertility was 0.51 ± 0.14 and 0.40 ± 0.20, with an average inbreeding in stallions of 0.33 ± 0.07 and 0.32 ± 0.06, respectively. Sorraia stallions and mares both show a negative influence of inbreeding on fertility (Table 2), which was only statistically significant in mares, with a decline of almost 0.8% in fertility when inbreeding increased by 1%.

Mare inbreeding had no significant effect on foaling interval ($P > 0.05$, Table 2), even though there was an increase of nearly 2 days in foaling interval as inbreeding increased by 1%. On the other hand, age of the mare had a statistically significant influence on foaling interval, which increased steadily as mares aged (Fig. 3).

The effect of inbreeding of the mare on age at first parturition was not statistically significant (P=0.176), but stud farm had an influence on this trait (P=0.0083). Animals from stud farm B have, on average, a higher age at first parturition than those from stud farm A (means of 72.51 ± 3.24 and 61.37 ± 2.51 months, respectively).

4. Discussion

The consequences of inbreeding over generations have been demonstrated in other equid small populations with increasing kinship between breeding animals, which may cause a reduction in fitness. An increase in juvenile mortality and decrease in life expectancy throughout successive generations of inbreeding have been reported (Ryder and Wedemeyer, 1982) and in the Przewalski horse it has been shown that animals with inbreeding coefficients of 0.25 or higher originated fewer descendants, possibly because inbreeding may affect ovarian function (Collins et al., 2012), and had a shorter life expectancy than those with lower inbreeding coefficients (Bouman and Bos, 1979). In the Lusitano breed, Oom (1992) showed that animals with higher inbreeding coefficients are more likely to produce non-viable offspring. Somewhat unexpectedly, the proportion of stillborn foals reported in our study with the Sorraia breed seems to be very low (about 2%), but it is important to refer that abortions and stillbirths are more difficult to detect in pasture breeding, which is the system typically used in Sorraia. Comparatively, Barbosa and Abreu (1986) observed 1.5% of abortions and 2.7% of stillborns in Lusitano mares, and Monfort et al. (1994) registered only 2% of abortions in the Przewalski horse.

Although the avoidance of inbreeding in Sorraia horses is nearly impossible, as the number of breeding animals is rather small and there are no longer unrelated animals, it is imperative that a conservation-breeding plan is adopted to keep inbreeding under control and allow the extremely high levels of inbreeding to decrease. Selecting adequate breeding pairs for that purpose should be a short-term goal, and efforts have been implemented since the last decade with promising results.

Horses tend to have lower reproductive performance than other domestic animals. The low fertility of horses has been attributed to the long period of oestrous behaviour of mares, the practise of reproducing them before the natural breeding season, and to inbreeding (Cothran et al., 1984). Some relevant contribution to this low reproductive performance should also be attributed to behavioural aspects, particularly under extensive management, namely those related with mare´s social dominance (Heitor et al., 2006), that are often not considered in these studies, as referred by Margulis (1998).

The number of offspring produced per mare is obviously limited by its own reproductive cycle (Cothran et al., 1984), whereas the number of offspring per stallion is a function of the number of mares covered per year, the duration of his reproductive life (time span during which the animal originates offspring), and his own fertility rate (Oliveira, 1999). The mean annual fertility rate of 46.74% obtained for Sorraia stallions in our study is lower than that obtained by Davies Morel and Gunnarsson (2000) for Icelandic stallions (67.7%), by Rho et al. (2004) for the Jeju pony stallions (60%), and by Barbosa and Abreu (1986) for Lusitano and Arab stallions (62.4% and 61.6%, respectively). This result might be due to the extremely high levels of inbreeding of Sorraia stallions that have been shown to lower semen quality (Gamboa et al., 2009). The fertility of mares (35.79%) is slightly lower than that of stallions and is also lower when compared to 65% in Jeju pony mares (Rho et al., 2004), 72.6% in Finnish Standardbred Trotters and 66.3% Finnhorses (Sairanen et al., 2009), 62% in Lusitano mares and 61.56% in Arabian mares (Barbosa and Abreu, 1986). The same explanation may be given to the difference between breeds, because Sorraia mares have the highest levels of inbreeding. However, we must be aware that fertility rates of Sorraia mares may be underrepresented, because the occurrence of abortions and foetus reabsorptions are often undetected in extensive management and a mare can thus be considered as to have failed to conceive when it indeed foaled a stillborn that was taken by predators before being reckoned, and this is virtually impossible to overcome in data collection. With respect to stallions, because breeding is done extensively and stallions stay up to four months with the herd having

several opportunities to cover the mares during that period and produce offspring, this could result in the overestimation of the stallions' fertility rate due to the effect of compensatory mating. A study done in Sorraia stallions by Gamboa et al. (2009) has revealed overall very poor semen quality, sperm motility and vitality, when compared to other horse breeds, as well as poor preservation ability, possibly due to inbreeding. In our study, there was a significant negative relationship ($P=0.003$) between inbreeding and fertility rates in Sorraia mares. These results suggest that in this population the effect of inbreeding depression for this component of fitness is indeed a concern, especially considering the aforementioned under- and overestimation calculations of fertility rates. In Standardbred horses, a negative relationship ($P<0.05$) exists between fertility and inbreeding (Cothran et al., 1984), but this relationship explains only about 2% of the variation in fertility. Sairanen et al. (2009) reported that foaling rates in horses decreased with increased inbreeding, despite the fact that the most inbred animals did not have the lowest foaling rates and, in the Finnhorse, they even had the highest performance regarding this variable.

Foaling intervals were influenced by the age of the mare ($P<0.0001$) but not by inbreeding. These findings are contrary to those of Panetto et al. (2012), where the age of Zebu cows did not cause an increase of calving intervals but inbreeding did. Age at first parturition is also not influenced by the inbreeding coefficient of the mare but is influenced by the stud farm ($P=0.0083$), and mares from stud farm A foal for the first time at an age almost 1 year younger than mares from stud farm B (average of 61.37 and 72.51 months, respectively).

Although previous studies have indicated that inbreeding does not have a significant effect on viability of Sorraia foals (Oom et al., 1991; Matos, 1996), the present results are based on a larger data set, collected over a longer period than the previous ones, but still indicate a weak relationship between foal inbreeding and viability. In this type of studies, the discrepancy of results may be caused by the disguising effects of husbandry conditions on inbreeding depression (Kalinowski et al., 2000; Kalinowski and Hedrick, 2001), as management conditions may ameliorate or enhance the effects of inbreeding depression, if the animals are well- or poorly kept, respectively.

The results of our study are not very conclusive regarding the effects of inbreeding in Sorraia horses. A general tendency was observed indicating that inbreeding is detrimental for the majority of the traits analysed, but residual variability was quite important and probably overshadowed the role of inbreeding depression in this highly endangered population. It is also possible that natural selection and genetic drift carried out over the years may have purged some of the negative effects of inbreeding, as suggested by Lacy (2000).

5. Conclusion

It is obviously necessary to establish a conservation-breeding plan for the Sorraia breed, as it is essential to promote the control of the extremely high levels of inbreeding currently observed, if the long-term self-sustaining maintenance of the population is intended. With this purpose, stallions should be chosen yearly for each group of mares according to their genetic relationship and the rotation of stallions among different stud farms should be promoted. These measures, which have already given positive results, will result in the genetic and demographic improvement of the Sorraia horse and help prevent the additional genetic erosion in the future. It is of paramount importance to preserve this breed of worldwide interest. Since it has not been subject of directional selection to enhance particular traits, the primitive characteristics and its hardiness will only be preserved if this breed is allowed to reproduce extensively. However, human intervention can never be excluded, as the preservation of the breed and the required increase of its census will only benefit from a management and conservation strategy, giving continuity to the multidisciplinary approach that is being implemented. To achieve conservation genetic goals we must not neglect the hypothesis of trying, in some particular cases, the use of some of the reproductive technologies currently available, to maximise the use of breeding animals in the population, both mares and stallions, even with low fertility rates.

Acknowledgements

The authors wish to thank the Associação Internacional de Criadores do Cavalo Ibérico de Tipo Primitivo – Sorraia for supplying the data analysed in the present study. This work was supported by Centro de Biologia Ambiental [PEst-OE/BIA/UI0329/2011] and by PRODER (Contract number 57882) (EU)/Action 2.2.3 (2011–2013) [Conservation and improvement of genetic resources. Sub-action 2.2.3.2. Animal component (Sorraia breed)]. H.J. Kjöllerström was supported by a PhD grant (SFRH/BD/81502/2011) from the Foundation for Science and Technology (FCT/MEC).

Appendix A. Supplementary information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.livsci.2015.08.001.

References

Andrade, R., 1945. O Cavalo do Sorraia. Boletim Pecuário XIII, pp. 1–13.
Ballou, J., 1983. Calculating inbreeding coefficients from pedigrees. In: Schonewald-Cox, C.M., Chambers, S.M., MacBryde, B., Thomas, L. (Eds.), Benjamin/Cummings Publ. Co., Menlo Park, California-USA, pp. 509–520.
Barbosa, M., Abreu, J.V., 1986. Alguns parâmetros reprodutivos em cavalos lusitanos e árabes. Rev. Port. Ciênc. Vet. 478, 147–169.
Bouman, I., 1989. Origins of the domestic horse. Przewalski Horse, pp. 32–43.
Bouman, J.G., Bos, H., 1979. Two symptoms of inbreeding depression in Przewalski horses living in captivity. In: Boer, L.E.M., Bouman, J., Bouman, I. (Eds.), Genetics and Hereditary Diseases of the Przewalski Horse. Foundation for the Preservation and Protection of the Przewalski Horse, Rotterdam – The Netherlands, pp. 111–117.
Buisman, A.K., van Weeren, R., 1982. Breeding and management of Przewalski horses in captivity. In: Bouman, J., Bouman, I., Groeneveld, A. (Eds.), Breeding Przewalski Horses in Captivity for Release in the Wild. Foundation for the Preservation and Protection of the Przewalski Horse, Rotterdam, The Netherlands, pp. 76–111.
Charlesworth, D., Willis, J.H., 2009. The genetics of inbreeding depression. Nat. Rev. Genet. 10, 783–796.
Collins, C.W., Songsasen, N.S., Vick, M.M., Wolfe, B.A., Weiss, R.B., Keefer, C.L., Monfort, S.L., 2012. Abnormal reproductive patterns in Przewalski's mares are associated with a loss in gene diversity. Biol. Rep. 86 (2), 1–10 28.
Cothran, E.G., MacCluer, J.W., Weitkamp, L.R., Pfennig, D.W., Boyce, A.J., 1984. Inbreeding and reproductive performance in standardbred horses. J. Hered. 75, 220–224.
Curik, I., 2003. Inbreeding, microsatellite heterozygosity, and morphological traits in Lipizzan horses. J. Hered. 94, 125–132.
Darwin, C., 1859. On the Origin of Species by Means of Natural Selection. Murray, London, England.
Davies Morel, M.C.G., Gunnarsson, V., 2000. A survey of the fertility of Icelandic stallions. Anim. Reprod. Sci. 64, 49–64.
Frankham, R., Briscoe, D.A., Ballou, J.D., 2002. Introduction to Conservation Genetics. Cambridge University Press, Cambridge, UK.
Gamboa, S., Machado-Faria, M., Ramalho-Santos, J., 2009. Seminal traits, suitability for semen preservation and fertility in the native Portuguese horse breeds Puro Sangue Lusitano and Sorraia: Implications for stallion classification and assisted reproduction. Anim. Reprod. Sci. 113, 102–113.
Gomes, J.C., Oom, M.M., 2000. Caracterização demográfica do núcleo de garranos bravios no PNPG (in Portuguese). Rev. Port. Zootec. 1, 67–78.
Heitor, F., Oom, M.M., Vicente, L., 2006. Social relationships in a herd of Sorraia horses Part II. Factors affecting affiliative relationships and sexual behaviours. Behav. Process. 73, 231–239.
ISIS, 2011. SPARKS (Single Population Analysis and Records Keeping System) ver 1.6. International Species Information System, Eagan, MN.

JMP®, 2013. Version Pro 10.0.2. SAS Institute Inc., Cary, NC.
Kalinowski, S.T., Hedrick, P.W., 2001. Inbreeding depression in captive bighorn sheep. Anim. Conserv. 4, 319–324.
Kalinowski, S.T., Hedrick, P.W., Miller, P.S., 2000. Inbreeding depression in the Speke's gazelle captive breeding program. Conserv. Biol. 14, 1375–1384.
Kjöllerström, H.J., 2005. O Studbook como instrumento na análise genética e demográfica de populações ameaçadas: o caso do Cavalo do Sorraia (in Portuguese). Animal Biology Department. Faculty of Sciences, University of Lisbon, Portugal, p. 50.
Lacy, R.C., 2000. Considering threats to the viability of small populations. Ecol. Bull. 48, 39–51.
Leroy, G., 2014. Inbreeding depression in livestock species: review and meta-analysis. Anim. Genet. 45, 618–628.
Luís, C., Bastos-Silveira, C., Cothran, E.G., Oom, M.M., 2006. Iberian origins of new world horse breeds. J. Hered. 97, 107–113.
Luís, C., Cothran, E.G., Oom, M.M., 2007. Inbreeding and genetic structure in the endangered Sorraia horse breed: implications for its conservation and management. J. Hered. 98, 232–237.
Margulis, S.W., 1998. Relationships among parental inbreeding, parental behaviour and offspring viability in oldfield mice. Anim. Behav. 55, 427–438.
Matos, S.M.S., 1996. Contributo para a gestão genética do Cavalo do Sorraia (in Portuguese). Animal Biology Department, University of Lisbon, Portugal, p. 44.
Meyer, K.W., 2007. WOMBAT-A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J. Zhejiang Univ. Sci. B. 8, 815–821.
Monfort, S.L., Arthur, N.P., Wildt, D.E., 1994. Reproduction in the Przewalski's horse. In: Boyd, L., Houpt, K.A. (Eds.), Suny Press, New York, pp. 173–194.
Oliveira, T.S.F., 1999. Análise de Parâmetros Demográficos e Reprodutivos na Coudelaria de Alter (in Portuguese). University of Lisbon, Portugal.
Oom, M.M., 1992. O cavalo Lusitano. Uma raça em Recuperação (in Portuguese). University of Lisbon, Portugal, p. 299.
Oom, M.M., Costa-Ferreira, J., Cothran, E.G., 1991. Inbreeding, reproductive success and genetic variation in the Sorraia horse of Portugal. Anim. Genet. 11, 22–23.
Oom, M.M., d'Andrade, J.L., Costa-Ferreira, J., 2004. Studbook da raça Sorraia (in Portuguese). Garrido Artes Gráficas, Alpiarça, Portugal.
Panetto, J.C.C., Val, J.E., Marcondes, C.R., Peixoto, M.G.C.D., Verneque, R.S., Ferraz, J.B. S., Golden, B.L., 2012. Female fertility in a Guzerat dairy subpopulation: heterogeneity of variance components for calving intervals. Livest. Sci. 145, 87–94.
Pinheiro, M., Kjöllerström, H.J., Oom, M.M., 2013. Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis. Livest. Sci. 152, 1–10.
Ralls, K., Brugger, K., Ballou, J., 1979. Inbreeding and juvenile mortality in small populations of ungulates. Science 206, 1101–1103.
Rho, J., Srygley, R., Choe, J., 2004. Age-specific fertility rates of the Jeju pony (Equus caballus). J. Ethol. 22, 113–117.
Ryan, K.K., Robert, C.L., Margulis, S.W., 2003. Impacts of inbreeding on components of reproductive success. In: Holt, W.V., Pickard, A.R., Rodger, J.C., Wildt, J.C. (Eds.), Reproduction Science and Integrated Conservation. Cambridge University Press, Cambridge, United Kingdom, pp. 82–95.
Ryder, O.A., Wedemeyer, E.A., 1982. A cooperative breeding programme for the Mongolian wild horse Equus przewalskii in the United States. Biol. Conserv. 22, 259–271.
Sairanen, J., Nivola, K., Katila, T., Virtala, A.M., Ojala, M., 2009. Effects of inbreeding and other genetic components on equine fertility. Animal 3, 1662–1672.
Valera, M., Esteves, L., Oom, M.M., Molina, A., 2000. La raça equina autóctone Puro Sangue Lusitano: estudo genético dos parâmetros reproductivos de importância nos esquemas de conservação e melhoramento. Arch. Zootec. 49, 147–156.
Vicente, A.A., Carolino, N., Gama, L.T., 2012. Genetic variability in the Lusitano horse breed assessed by pedigree analysis. Livest. Sci. 148, 16–25.

Table S1. Genetic parameter estimates

Trait	Heritability	Repeatibility
Foaling interval	0.12 ± 0.01	0.12 ± 0.06
Age at first parturition	0.48 ± 0.24	-
Foal survival (birth)	0.00 ± 0.09	-
Foal survival (6M)	0.00 ± 0.07	-
Mare fertility	0.92 ± 0.13	-
Stallion fertility	0.89 ± 0.64	-

CHAPTER 3 - PAPER 5

First evidence of sex chromosome mosaicism in the endangered Sorraia Horse breed.

Kjöllerström H.J., Collares-Pereira M.J. & Oom M.M.

Livestock Science (2011) 136, 273-6

Livestock Science 136 (2011) 273–276

Contents lists available at ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

Short Communication

First evidence of sex chromosome mosaicism in the endangered Sorraia Horse breed

H.J. Kjöllerström, M.J. Collares-Pereira, M.M. Oom *

Universidade de Lisboa, Faculdade de Ciências, Centro de Biologia Animal, Campo Grande, 1749-016 Lisboa, Portugal

ARTICLE INFO

Article history:
Received 29 January 2010
Received in revised form 31 July 2010
Accepted 11 August 2010

Keywords:
X monosomy
Mosaic
Subfertility
Equus caballus
Sorraia Horse

ABSTRACT

The Sorraia Horse is a critically endangered Portuguese breed with an extremely reduced effective size, having reached unusual high levels of inbreeding. Fertility is of crucial importance for the long-term survival of the extant population and it has been shown that chromosomal abnormalities, especially on sex chromosomes, in horses are associated with infertility or subfertility. To date, no cytogenetic studies were performed in the Sorraia breed to assess the extent of chromosome abnormalities in animals with fertility problems. We now report the results of the first studied case – a subfertile mare with a stallion-like behaviour. A mosaic 63,X0/64,XX karyotype (with 10.45% and 89.55% frequency, respectively) was found. Further cytogenetic screenings in Sorraia horses with low breeding performance and ambiguous sexual phenotypes are needed, to determine the extent of chromosomal abnormalities, since early detection of these animals is of paramount importance for the management and conservation of the breed.

1. Introduction

The Sorraia horse is one of the three native Portuguese horse breeds and it is believed to represent a primitive equine type with a continuous presence in the Iberian Peninsula since early Pleistocene (Luís et al., 2006; Oom et al., 2004). Recovered in 1937 from only 12 founders in the Sorraia river valley, the breed has been managed as a closed population since (e.g. Oom et al., 1991). As an important animal genetic resource (AnGR) with less than 150 extant breeding mares, this breed was considered in "critical maintained risk status" by FAO (FAO/UNEP, 2000), and is the only equine breed recognized as "rare/particularly endangered" by national authorities (MADRP, 2007).

The small number of founders, the reduced effective population size (N_e), the complete genetic isolation and the breeding management adopted, led inbreeding to steadily increase, reaching extremely high levels (average F = 0.37), and mating of closely related individuals is now unavoidable (e.g. Oom et al., 2004). Complete pedigrees are available and were crucial to correctly estimate inbreeding coefficients. All these data strongly indicate that the Sorraia breed is a unique biological model to investigate the effects of inbreeding on different traits in horses.

Previous studies provided general information about the Sorraia breed through genetic, pedigree, ethological and morphological analyses. Molecular markers evidenced a decreased genetic variation, indicating that the genome has been largely affected by founder effect, genetic drift and inbreeding (Luís et al., 2007a,b).

Since reproductive success is of paramount importance for the conservation and management of this endangered breed, a comprehensive knowledge of the likely causes of reduced fertility is required. Inbreeding is unavoidable in the Sorraia breed and is commonly associated with decreased fitness, such as reproductive characteristics and offspring viability. The inbreeding effects vary between populations and traits and phenotypes often lack accuracy and give little information to evaluate the extent of the resulting events. In some preliminary studies (Kjöllerström, 2005; Oom et al., 1991), inbreeding was significantly negatively correlated with

* Corresponding author. Tel.: +351 217500000; fax: +351 217500028.
E-mail address: mmoom@fc.ul.pt (M.M. Oom).

1871-1413/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.livsci.2010.08.005

juvenile survival and some suggestions exist that it may be related with both male and female fertility problems, though more data are required for an accurate evaluation. This is of particular concern as a small N_e makes the population even more sensitive to demographic stochasticity, thus special attention must be given to study and improve reproductive performance.

Extensive cytogenetic investigations in horses showed that chromosomal abnormalities, especially of sex chromosomes, are commonly associated to horse infertility or subfertility and to repeated early embryonic death, abortion and stillbirth, although occasional pregnancies may occur (e.g. Breen et al., 1997; Bugno et al., 2008, 2009; Chowdhary and Raudsepp, 2000; Lear and Bailey, 2008). Cytogenetic analyses were never performed in the Sorraia breed, despite several known cases of infertility and subfertility.

Most sex chromosomes abnormalities correspond to normal phenotypes in mares and karyotyping is needed for definitive diagnosis. As it is possible, although not common, that mares with some chromosomal abnormalities produce live foals, those with a small number of life-time foals are potential carriers of chromosomal aberrations and should be checked (Lear et al., 2008; Vanderwall, 2008).

The nuclear genome of the horse comprises 64 chromosomes (31 pairs of autosomes plus the sex chromosomes, X and Y) which are well characterized (ISCNH, 1997). The most commonly reported sex chromosome abnormality is X monosomy (63,X0), first described by Payne et al. (1968), representing more then 50% of all identified chromosome abnormalities in horses (Bugno et al., 2001). Also reported are the XY sex reversal (second most common anomaly found in infertile mares), the XX sex reversal, different types of mosaicism and chromosomal rearrangements and, more rarely, the XXX trisomy. Chromosomal mosaicism (around 30% of horse chromosomal abnormalities according to Lear and Bailey, 2008) is the presence of two or more chromosomally distinct cell lines in an individual.

In the frame of an ongoing cytogenetic survey of the extant Sorraia horses' population, particularly focused in animals with reduced breeding performance and ambiguous sexual phenotypes, we now report the results of the first studied case – a subfertile mare with a stallion-like behaviour – that evidenced a karyotype with a mosaic pattern.

2. Material and methods

2.1. Case study

A subfertile mare, born in 1993, with a phenotypic normal external genitalia but exhibiting stallion-like behaviour and a masculinised body conformation (Fig. 1), was cytogenetically analysed. Ultrasonography showed normal developed ovaries and uterus, although no natural cyclic pattern of oestrus behaviour was detected (P. Bravo, personal communication).

With a high value of inbreeding coefficient ($F = 0.33$), the mare has a low fertility value (0.18) when compared to the average fertility achieved at the same breeding farm (0.38) (Fig. 2): only two foals were produced during her lifetime, one male in 1998 and one female in 2007 (the former by natural covering in the pasture, the latter after an intensive hormonal treatment and artificial insemination, AI). From March to July of 2009 she was submitted to another intensive program of assisted reproduction at the ESAC Animal Reproduction Laboratory Dr. França Martins (Coimbra, Portugal) for induced oestrus and ovulation with Dinolytic® (Pfizer) and Chorulon® (Intervet Schering–Plough). Although developed follicles and corpus luteum were successively detected by ultrasonography, negative gestation diagnosis was confirmed after AI with fresh diluted semen (P. Bravo, personal communication). Sperm concentration, motility, vitality and morphology from the selected stallion were considered above average as regards the Sorraia breed semen characteristics (S. Gamboa, personal communication).

2.2. Cytogenetic analysis

We used standard methods for obtaining chromosome preparations from peripheral blood lymphocytes cultures (Raudsepp and Chowdhary, 2008). Slides were first stained with Giemsa and metaphase spreads assessed by karyotyping following ISCNH (1997) (data not shown). As X chromosomes have a conspicuous interstitial heterochromatic band and the Y chromosome is also well recognized by the length of a heterochromatic block, both in the q arm, we primordially used CBG-banding technique, adjusting the classic protocol (Sumner, 1972): 25 min HCl (0.2 N) at room temperature, 5 min $Ba(OH)_2$ (2.5%) at 60 °C, 10 min SSC 2X

Fig. 1. Phenotype of the studied animal showing male-like characteristics unusual in normal mares (thicker and crested neck, higher bone mass and muscle condition).

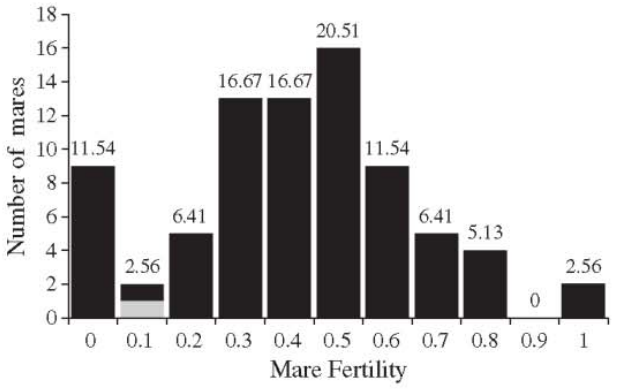


Fig. 2. Distribution of mare fertility in the Sorraia breed calculated as the number of foals per number of years in reproduction. The lighter area corresponds to the studied case. Percentages of mares in each class are presented at the top of the column (N = 78; data from one breeding farm between 1967 and 2007).

at 60 °C, and 10 min in 5% Giemsa staining at room temperature.

Metaphase spreads were screened under an Olympus BX60 microscope equipped with a DP50 Olympus digital camera and further analysed for karyotype description.

3. Results

A total of 174 metaphase spreads were analysed using standard Giemsa staining (data not shown) and CBG-banding techniques and allowed to clearly find a mosaic pattern, i.e. two lymphocyte cell lines, 63,X0/64,XX. Although 89.6% of the observed plates had the normal karyotype pattern (2n = 64, XX), the remaining (10.4%) had 2n = 63 and only one X chromosome. This X monosomy was confirmed by CBG-banding (Fig. 3).

Moreover, due to the stallion-behaviour of the studied mare, the eventual existence of a Y chromosome and of another cell line was easily discarded also by CBG-banding.

4. Discussion

The mosaicism herein found and characterized by the co-occurrence of cells with normal female sex chromosomes (64, XX) and cells with X monosomy (63,X0), was firstly described by Chandley et al. (1975), being the second most common aberration involving sex chromosomes, after excluding XY sex-reversal cases (Chowdhary and Raudsepp, 2000).

Such mosaic-type is usually associated with fertility problems, smaller body conformation, normal female external genitalia and gonadal dysgenesis, and an irregular or absent oestrus cycle, in similar clinical manifestation of 63,X0 mares (e.g. Bugno et al., 2009; Chowdhary and Raudsepp, 2000; Hughes and Trommershausen-Smith, 1977). Impact on fertility may depend on the relative proportion of the two cell lines in the gonads: individuals may be only subfertile, as it is the case here reported, or exhibit an infertile condition (Bugno et al., 2001; Chandley et al., 1975; Reid et al., 1987; Wieczorek et al., 2001).

The results obtained are consistent with the reduced fertility of the studied mare. Also relevant is the fact that, despite being a mosaic, she produce normal gametes and had two viable offspring to date, confirming that some of these mares can reproduce, as pointed out in three previous cases (Bugno et al., 2001; Halnan, 1985; Wieczorek et al., 2001), contradicting the frequently assumed full sterility of mosaic individuals.

The fact that this mosaic-type was never associated with a masculine body conformation and a stallion-like behaviour, as in the studied mare, corroborates the importance of carrying out a systematic cytogenetic survey of all specimens as the same chromosome aberration may indeed correspond to different phenotypes, including normal ones. Thus looking for a hypothetic mosaicism involving a Y chromosome by screening other cell lines, such as fibroblasts, or testing nuclear markers and microsatellites linked to sex chromosomes is needed to disentangle this question.

Inheritance of aneuploidies of sex chromosomes in the few reported cases of subfertility hasn't been described so far. In the follow up cytogenetic study we will determine whether

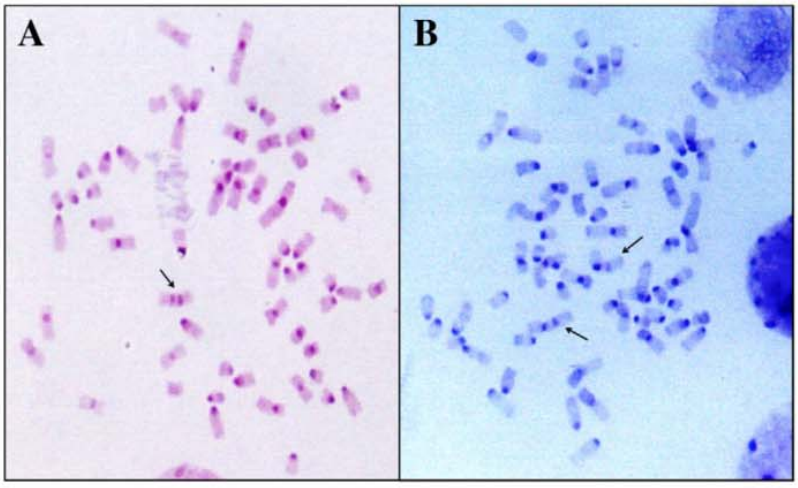


Fig. 3. CBG-banding metaphases from the studied mare representing both cell lines found in lymphocyte cell cultures: (A) 2n = 63,X0 and (B) 2n = 64,XX. Arrows indicate the X chromosomes.

or not the progeny of this mare has chromosomal abnormalities, by analysing the karyotypes of the two offspring, one stallion and one filly, also highly inbred (F = 0.49 and F = 0.27, respectively). The stallion has never been able to reproduce or to give a semen sample, despite having achieved to cover a mare in a pasture breeding system, but with no further positive pregnancy diagnostic.

Moreover, the ongoing study will look for chromosomal abnormalities that might be contributing to induce infertility or subfertility in a representative sample of the Sorraia population, with special focus on animals with reported fertility problems. Molecular cytogenetic analyses, using painting probes specific for particular chromosome pairs by fluorescent *in situ* hybridization (FISH) technique, are rapidly expanding in horses (Bugno et al., 2009) and will be of great help for karyotyping. The possibility of using X or/and Y whole chromosome painting probes (WCPP), also on interphase nuclei, will increase the efficiency and accuracy of karyotyping tests, reducing the time needed to detect subfertile and infertile horses due to sex chromosomes aneuploidies, as first showed by Breen et al. (1997) and further pointed out by other authors (Bugno et al., 2006, 2008, 2009; Wieczorek et al., 2001).

Since most sex chromosomes abnormalities correspond to normal phenotypes in mares, a higher than thought percentage of particularly difficult breeding cases may be attributed to genetically abnormal animals that remain barren no matter the best efforts from veterinaries and breeders. Thus, karyotyping will ultimately be of great help in providing definitive diagnosis for cases of reduced fertility in this critically endangered horse breed, and also in the early detection of animals that will have none or poor breeding performance due to chromosomal abnormalities will save time, efforts and breeders' resources.

Acknowledgments

We thank B. Chowdhary and T. Raudsepp (Texas A&M University, USA) for suggestions and help with banding and culture techniques and the mare's breeder, Andrade de Oliveira e Sousa Stud Farm (Coruche, Portugal), as well as J. Serrano for supplying the mare's photos. From ESAC (Escola Superior Agrária de Coimbra, Portugal) Reproduction Laboratory Dr. França Martins, we are grateful to P. Bravo and R. Rebordão for ultrasonographies and reproductive evaluation of the mare, and S. Gamboa for the analyses and preparation of semen samples and AI program. This reproductive evaluation was partially supported by funds from PRODER (Conservação e Melhoramento de Recursos Genéticos, Componente Animal), MADRP, Portugal.

References

Breen, M., Langford, C.F., Carter, N.P., Fischer, P.E., Marti, E., Gerstenberg, C., Allen, W.R., Lear, T.L., Binns, M.M., 1997. Detection of equine X chromosome abnormalities in equids using a horse X Whole Chromosome Paint Probe (WCPP). Vet. J. 153, 235–238.

Bugno, M., Slota, E., Zabek, M., 2001. Two cases of subfertile mares with 64, XX/63,X mosaic karyotype. Ann. Anim. Sci. 1 (2), 7–11.

Bugno, M., Pienkowska-Schelling, A., Schelling, C., Wlodarczak, N., Slota, E., 2006. Utilization of the probe generated by chromosome microdissecation, for detection of equine X chromosome aneuploidy. Ann. Anim. Sci. 6 (2), 205–210.

Bugno, M., Zabek, T., Golonka, P., Pienkowska-Schelling, A., Schelling, C., Slota, E., 2008. A case of an intersex horse with 63, X/64, XX/65, XX, del (Y)(q?) karyotype. Cytogenet. Genome Res. 120, 123–126.

Bugno, M., Slota, E., Pienkowska-Schelling, A., Schelling, C., 2009. Identification of chromosome abnormalities in the horse using a panel of chromosome-specific painting probes generated by microdissection. Acta Vet. Hung. 57 (3), 369–381.

Chandley, A.C., Fletcher, J.M., Rossdale, P.D., Peace, C.K., Ricketts, S.W., McEnery, R.J., Thorn, J.P., Short, R.V., Allen, W.R., 1975. Chromosome abnormalities as a cause of infertility in mares. J. Reprod. Fertil. Suppl. 23, 377–383.

Chowdhary, B.P., Raudsepp, T., 2000. Cytogenetics and physical gene maps. In: Bowling, A.T., Ruvinsky, A. (Eds.), The Genetics of the Horse. CABI, Wallingford, pp. 171–242.

FAO/UNEP, 2000. In: Scherf, B.D. (Ed.), World Watch List for Domestic Animal Diversity, 3rd edition. Rome, Italy.

Halnan, C.R.E., 1985. Sex chromosome mosaicism and infertility in mares. Vet. Rec. 116, 542–543.

Hughes, J.P., Trommershausen-Smith, A., 1977. Infertility in the horse associated with chromosomal abnormalities. Aust. Vet. J. 53, 253–257.

ISCNH, 1997 (International System for Cytogenetics Nomenclature of the Domestic Horse). Bowling, A.T., Breen, M., Chowdhary, B.P., Hirota, K., Lear, T., Millon, L.V., Ponce de Leon, F.A., Raudsepp, T., Stranzinger, G. (Committee). Chromosome Res. 5, 433–443.

Kjöllerström, H.J., 2005. The studbook as an instrument in the genetic and demographic analysis of endangered populations: the Sorraia horse case. Final Graduation Report. Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal. 50 pp. (in Portuguese).

Lear, T., Bailey, E., 2008. Equine clinical cytogenetics: the past and future. Cytogenet. Genome Res. 120, 42–49.

Lear, T.L., Lundquist, J., Zent, W.W., Fishback Jr., W.D., Clark, A., 2008. Three autosomal chromosome translocations associated with repeated early embryonic loss (REEL) in the domestic horse (*Equus caballus*). Cytogenet. Genome Res. 120, 117–122.

Luís, C., Bastos-Silveira, C., Cothran, E.G., Oom, M.M., 2006. Iberian origins of New World horse breeds. J. Hered. 97, 107–113.

Luís, C., Cothran, E.G., Oom, M.M., 2007a. Inbreeding and genetic structure in the endangered Sorraia horse breed. Implications for its conservation and management. J. Hered. 98, 232–237.

Luís, C., Juras, R., Oom, M.M., Cothran, E.G., 2007b. Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. Anim. Genet. 38 (1), 20–27.

MADRP, 2007. Programa de Desenvolvimento Rural — Continente 2007–2013. Lisboa, Portugal. (in Portuguese).

Oom, M.M., Costa-Ferreira, J., Cothran, E.G., 1991. Genetic variation in the Sorraia Horse of Portugal. Anim. Genet. 22 (Supl.1), 22–23.

Oom, M.M., d'Andrade, J.L., Costa-Ferreira, J., 2004. The Sorraia horse breed Studbook. Garrido Artes Gráficas, Alpiarça, Portugal. 357 pp. (in Portuguese).

Payne, H.W., Ellsworth, K., DeGroot, A., 1968. Aneuploidy in an infertile mare. J. Am. Vet. Med. Assoc. 153, 1293–1299.

Raudsepp, T., Chowdhary, B.P., 2008. FISH for mapping single copy genes. In: Murphy, W. (Ed.), Phylogenomics: Methods in Molecular Biology. Humana Press, New Jersey, USA, pp. 31–49.

Reid, S.W., Weatherston, J.F., Robinson, B.N., 1987. Chromosome mosaicism and infertility in a mare. Can. Vet. J. 28, 533–534.

Sumner, A.T., 1972. A simple technique for demonstrating centromeric heterochromatine. Exp. Cell Res. 75, 303–306.

Vanderwall, D.K., 2008. Early embryonic loss in the mare. J. Equine Vet. Sci. 28 (11), 691–702.

Wieczorek, M., Switoński, M., Yang, F., 2001. A low-level X chromosome mosaicism in mares, detected by chromosome painting. J. Appl. Genet. 42 (2), 205–209.

CHAPTER 3 - PAPER 6

Fertility assessment in Sorraia stallions by sperm-FISH and FKBP6 genotyping.

Kjöllerström H.J., Oom M.M., Chowdhary, B.P & Radsepp, T.

Reproduction in Domestic Animals (2016) 51, 351-9

Reproduction in Domestic Animals

Reprod Dom Anim doi: 10.1111/rda.12686
ISSN 0936–6768

Fertility Assessment in Sorraia Stallions by Sperm-Fish and *Fkbp6* Genotyping

HJ Kjöllerström[1,2], M do Mar Oom[1], BP Chowdhary[3] and T Raudsepp[2]

[1]*cE3c – Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal;* [2]*Department of Veterinary Integrative Biosciences, College of Veterinary Medicine & Biomedical Sciences, Texas A&M University, College Station, Texas, USA;* [3]*New Research Complex, Qatar University, Doha, Qatar*

Contents
The Sorraia, a critically endangered indigenous Iberian horse breed, is characterized by low genetic variability, high rate of inbreeding, bad sperm quality and subfertility. Here, we studied 11 phenotypically normal but subfertile Sorraia stallions by karyotyping, sex chromosome sperm-FISH and molecular analysis of *FKBP6* – a susceptibility locus for impaired acrosome reaction (IAR). The stallions had normal sperm concentration (>300 million cells/ml), but the numbers of progressively motile sperm (21%) and morphologically normal sperm (28%) were invariably low. All stallions had a normal 64,XY karyotype. The majority of sperm (89%) had normal haploid sex chromosome content, although 11% of sperm carried various sex chromosome aneuploidies. No correlation was found between the percentage of sperm sex chromosome abnormalities and inbreeding, sperm morphology or stallion age. Direct sequencing of *FKBP6* exon 4 for SNPs g.11040315G>A and g.11040379C>A revealed that none of the stallions had the susceptibility genotype (A/A-A/A) for IAR. Instead, all animals had a G/G-A/A genotype – a testimony of low genetic variability. The findings ruled out chromosomal abnormalities and genetic predisposition for IAR as contributing factors for subfertility. However, low fertility of the Sorraia stallions could be partly attributed to relatively higher rate of sex chromosome aneuploidies in the sperm.

Introduction

The Sorraia horse is a critically endangered indigenous Portuguese horse breed with primitive characteristics and ability to survive in the harsh environment of the Iberian Peninsula (Oom et al. 2004; Pinheiro et al. 2013). Genetic isolation, small number of founders (n = 12), small population size (roughly 300 animals) and breeding management have led to a decrease in genetic variation among Sorraias over time (Luis et al. 2007a,b). The high rate of inbreeding (average 37.8%; Pinheiro et al. 2013; Kjöllerström et al. 2015), in turn, has resulted in low fertility (mean fertility rate 46.7%) and bad sperm quality of Sorraia stallions compared to other Portuguese, as well as non-Iberian horse breeds (Gamboa and Ramalho-Santos 2005; Gamboa et al. 2009; Kjöllerström et al. 2015).

Stallion fertility is a complex trait regulated by environmental, physiological, behavioural, epigenetic and genetic factors (Raudsepp et al. 2013). Knowledge about the latter is sparse and mainly limited to chromosome aberrations. For example, balanced translocations can cause subfertility by affecting spermatogenesis and the viability of embryos, while the overall phenotype of the carrier stallion remains normal (Raudsepp et al. 2013; Raudsepp and Chowdhary 2016). That is why conventional karyotyping is typically one of the first approaches to evaluate the genetic health of subfertile stallions (Durkin et al. 2011). In addition, chromosomal constitution of the sperm can be directly evaluated by fluorescence *in situ* hybridization (FISH) on decondensed sperm heads. The method, known as sperm-FISH, was initially developed for humans (Wyrobek et al. 1990) but has become a state-of-art technique for the detection of meiotic chromosome defects in the sperm of domestic species (Raudsepp and Chowdhary 2016). Sperm-FISH has also been optimized for stallions (Bugno-Poniewierska et al. 2009) and used for the study of sex chromosome aneuploidies in the sperm of reproductively normal stallions (Bugno et al. 2010) and for evaluating correlation between stallion age and the rate of sperm chromosome abnormalities (Bugno-Poniewierska et al. 2011, 2014). No sperm-FISH studies have, as yet, been conducted in subfertile or infertile stallions.

Besides cytogenetic evaluation of stallions, recent genomewide association studies have revealed genes and genomic regions that might be associated with stallion fertility (Schrimpf et al. 2014, 2015) and conditions underlying subfertility (Raudsepp et al. 2012). One of the genes of interest is *FK506-binding protein 6* (*FKBP6*). The gene has been proposed as a susceptibility locus for subfertility due to impaired acrosome reaction (IAR) in Thoroughbred stallions (Raudsepp et al. 2012). Specifically, strong association was found between IAR and a double homozygous A/A-A/A genotype consisting of SNPs g.11040315G>A and g.11040379C>A in *FKBP6* exon 4. Association between *FKBP6* and stallion fertility was also reported in a recent study of Hanoverian stallions (Schrimpf et al. 2015). However, in contrast to the IAR study (Raudsepp et al. 2012), here the *FKBP6* SNP g.11040379C>A conferred higher conception rates in A/A homozygous and lower conception rates in C/C homozygous stallions, whereas the double homozygous A/A-A/A genotype in exon 4 was not associated with fertility determined as pregnancy rate per oestrous cycle (Schrimpf et al. 2015). It appears that *FKBP6* sequence variants show contrasting associations with stallion fertility in different breeds, and the gene requires further investigation.

In this study, we carry out karyotyping, sperm-FISH and *FKBP6* exon 4 genotyping in subfertile Sorraia stallions to determine the prevalence of chromosome abnormalities, meiotic segregation errors of the sex chromosomes and the frequency of the IAR susceptibility genotype, respectively.

Material and Methods

Procurement of stallion semen and peripheral blood was performed by veterinarians according to the protocols of Gamboa et al. (2009) and Raudsepp and Chowdhary (2008).

Animals and phenotypes

We used 11 phenotypically normal 4- to 11-year-old Sorraia stallions (Table 1) available at the National Stud Farm in Alter do Chão, Portugal. Inbreeding values were calculated based on the additive relationship matrix (Ballou 1983) in SPARKS (ISIS 2011) with complete pedigrees traced back to the base population (founders). Stallion fertility was calculated as total number of offspring divided by total number of available mares (Kjöllerström et al. 2015). A stallion was considered subfertile if the fertility rate was lower than 40%.

Semen collection and evaluation of seminal characteristics

Semen samples were collected using an oestrous mare and a INRA model artificial vagina. Immediately after collection, raw ejaculates were filtered through a sterile gauze to remove the gel and any large debris. Seminal characteristics such as semen volume and concentration, sperm motion characteristics, morphological features and vitality were assessed in the gel-free semen using standard techniques described in Gamboa et al. (2009). The remaining gel-free semen samples were purified through 40% EquiPure™ Top Layer (Nidacon International, Mölndal, Sweden) and stored in RNAlater® (Ambion, Applied BioSystems) at −80°C (Das et al. 2010) for further analysis.

Karyotyping

All 11 stallions involved in this study were karyotyped using short-term blood lymphocyte cultures according to standard protocols (Raudsepp and Chowdhary 2008). Images were captured and analysed with Zeiss Axioplan 2 fluorescence microscope and Ikaros (MetaSystems GmbH) software. Karyotypes were arranged following the international standard chromosome nomenclature for the horse (ISCNH 1997).

Sperm-FISH

Semen samples of 6 stallions, all with sperm count over 300×10^6 cells/ml (n = 6; Table 1), were used for sperm-FISH.

Preparation of slides

Slides for sperm-FISH were prepared according to standard protocols with our modifications (Pellestor et al. 2003). Briefly, purified semen samples were washed three times in PBS; 5 µl of sperm suspension was placed on one end of a microscope slide, spread with a coverslip, air-dried and checked under phase contrast microscope for quality and concentration. Next, the slides were fixed in methanol/acetic acid (3 : 1) for 20 min at −20°C and air-dried. Sperm heads were decondensed by immersion the slides in 3M NaOH for 5 min and air-dried. Finally, the slides were immersed in methanol/acetic acid (3 : 1) for 10 min at −20°C and air-dried.

Probes and labelling

The probes for sperm-FISH were 2 clones from horse CHORI-241 BAC library (http://bacpac.chori.org/equine241.htm): BAC 99M16 containing X chromosome

Table 1. Sperm characteristics, inbreeding rates and fertility of the 11 Sorraia stallions used in this study

Stallion ID	Age	Concentration (x10^6/ml)	Progressive motile sperm (%)	Sperm with normal morphology (%)	Inbreeding (%)	Fertility(%)
23667*	11	479	35.0	37	26.9	27.3
E22*	7	409	25.0	34	29.5	NU
E23*	7	437	15.0	11	29.7	0
E59*	8	381	42.5	44	26.8	12.9
E133*	5	338	12.5	20	42.8	0
E135*	4	719	45.0	37	43.2	0
14 837	12	0	0	0	34.8	NU
E13	13	1	0	0	34.8	NU
E35	6	151	23.6	NA	26.5	40.0
E197	15	300	15.0	38	33.4	12.5
E223	14	163	27.5	63	44.6	20.0
Mean	9.3	307	21.9	28	33.9	10.2

Animals with an asterisk were used for sperm-FISH analysis.
Age, age at the time of collection; NA, data not available; NU, never used as stallion.

sequences including the androgen receptor gene (AR) and BAC 152E2 containing single-copy sequences from the Y chromosome (our unpublished data). The X chromosome BAC was labelled with biotin and the Y chromosome BAC with digoxigenin using Biotin- and DIG-Nick Translation Mix (Roche), respectively.

FISH

Before using the probes for sperm-FISH, they were tested for hybridization quality on chromosome slides from a male and a female horse. Both chromosome and sperm hybridizations and signal detection were conducted as described by Raudsepp and Chowdhary (2008). The only difference was that sperm slides were denatured in 70% formamide/2xSSC at 70°C for 5 min, while chromosome slides for 2 min.

Signal scoring and image analysis

FISH results were analysed and images captured with a Zeiss Axioplan 2 fluorescence microscope equipped with Isis V 5.2 (MetaSystems GmbH) software. At least 1300 sperm cells per individual were scored for the X and Y chromosome signals. In case the same chromosome (X or Y) produced two or more signals on one sperm head, they were scored as separate signals if the distance between them was larger than one signal diameter. Otherwise, they were scored as a single signal. At least 5 images were captured for control hybridizations on metaphase spreads and at least 20 images for each sperm-FISH experiment.

Statistical analysis

A t-test was used to evaluate whether the prevalence of X- and Y-bearing sperm deviated from the expected 1 : 1 ratio. Pearson's correlation (r) of sperm sex chromosome abnormalities with stallion age, inbreeding rate and sperm morphology was tested using Statistica 12 (StatSoft 2013) software.

FKBP6 exon 4 genotyping

DNA samples

DNA was extracted from whole blood of 11 Sorraia stallions (Table 1) according to standard protocols (Montgomery and Sise 1990), with minor modifications. DNA from one fertile Thoroughbred stallion heterozygous for both *FKBP6* exon 4 SNPs (Raudsepp et al. 2012) and one stallion with the IAR phenotype and the susceptibility genotype A/A-A/A were used as controls.

Sequencing

Genotyping of *FKBP6* exon 4 SNPs g.11040315G>A and g.11040379C>A was carried out by bidirectional Sanger sequencing of exon 4 PCR amplicon as described by Raudsepp et al. (2012). Sequencing reactions were performed with BigDye (Applied Biosystems) chemistry, resolved on an ABI 3730 capillary sequencer (Applied Biosystems) and analysed with Sequencher V 4.7 software (GeneCodes Co).

Results

Karyotyping

Chromosome analysis showed that all stallions had 64, XY karyotypes with normal chromosome number and morphology (data not shown).

Seminal characteristics

Sperm concentration of all Sorraia stallions ranged from 0×10^6 to 719×10^6 cells/ml with an average of 307×10^6 cells/ml. Percentage of progressive motile sperm, however, was low, with an average of 21.9% and not exceeding 50% in any of the stallions. Morphological characteristics of the sperm of all stallions were poor. Stallion E23 had only 11% of morphologically normal sperm, and on average, only 28% of sperm of all stallions had normal morphology. A summary of all seminal characteristics together with fertility rates and inbreeding coefficients for individual stallions is presented in Table 1. We concluded that the overall sperm quality of the Sorraia stallions involved in this study was poor.

Sperm-FISH

Before applying to stallion sperm preparations, hybridization quality of the X and Y chromosome-specific BAC probes was tested in normal male and female horse somatic cell chromosomes. Both probes gave strong and clean FISH signals in metaphase and interphase chromosomes without confounding background. The BAC 99M16 mapped to Xq12 (Fig. 1a), which is consistent with the location of *AR* (Raudsepp et al. 2008), and BAC 152E2 mapped to Yq14-q15 (Fig. 1b), consistent with the location of male specific region of Y (MSY) in horses (Raudsepp et al. 2004).

After FISH to decondensed sperm heads, the X- and Y-specific signals were scored, on average in 1355 sperm per stallion, and the results are summarized in Table 2. The majority of the sperm (average 89%) in all stallions were normal haploid for the sex chromosome content, carrying either a single X chromosome (average 44%) or a single Y chromosome (average 45%) (Figs 1c; 2). On average, we counted more Y-bearing than X-bearing sperm, although this was not significantly ($p = 0.478$) different from the expected (50 : 50) ratio. The remaining 11% of sperm had abnormal content of sex chromosomes. Of these, the most prevalent were sperm without sex chromosomes (average 5.7%), followed by XY (average 2%), YY (average 1.9%) and XX (average 1.3%) sperm (Figs. 1c, 1d and 2). A small fraction of

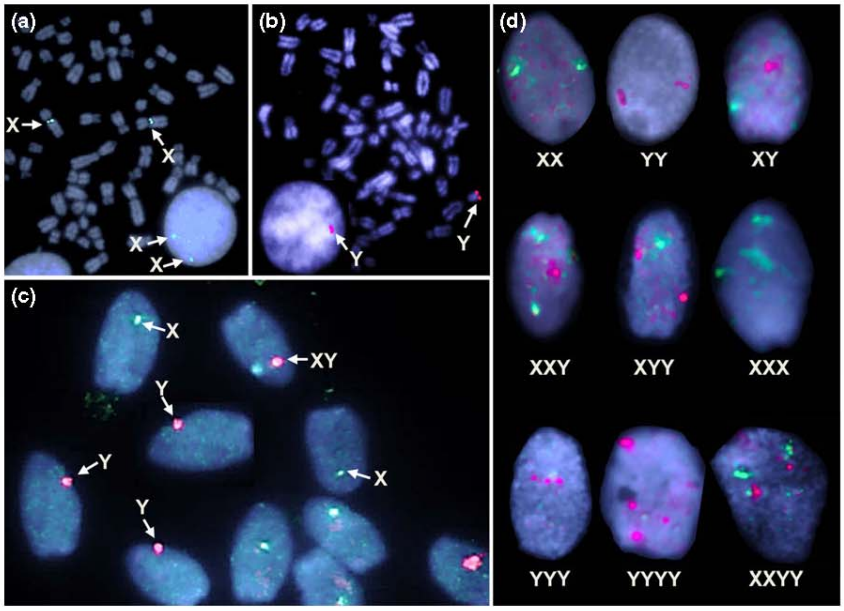


Fig. 1. FISH with X (BAC 99M16; androgen receptor – green) and Y (BAC 152E2 – red) probes. (a–b) Metaphase and interphase somatic cells of control horses: female (a) and male (b) showing signals (arrows) on the two X chromosomes and the Y chromosome, respectively; (c) sperm with X, Y and XY sex chromosomes (arrows) and (d) sperm with various sex chromosomal aneuploidies

Table 2. Sex chromosome content of Sorraia stallion sperm as revealed by sperm-FISH

	Chromosome content													
Stallion ID	X	Y	XX	XY	YY	XXY	XXX	XYY	YYY	YYYY	XXYY	No signal	Total Abnormal sperm (%)	Total # of cells analysed
23667	45.60	49.00	0.30	0.40	0.70	0.10	–	–	–	–	–	3.90	1.5	1342
E22	41.30	40.60	3.60	3.90	2.20	0.30	0.20	0.30	0.20	0.20	–	7.30	10.8	1317
E23	45.90	44.90	1.00	2.20	1.90	0.20	–	0.10	–	–	0.10	3.70	5.5	1362
E59	40.90	44.20	1.60	2.00	3.50	0.30	–	0.10	–	–	0.20	7.30	7.7	1324
E133	44.20	45.90	1.20	2.80	1.40	0.40	–	0.10	0.10	0.10	–	3.80	6.0	1406
E135	43.70	44.70	0.40	0.90	2.20	–	–	–	–	–	–	8.20	3.5	1381
Mean	43.58	44.89	1.35	2.05	1.99	0.21	0.03	0.10	0.04	0.04	0.04	5.70	5.8	1355

All values are in percentage; – no cells found.

sperm (average 0.46%) contained XXX, XYY, YYY, YYYY or XXYY sex chromosome complement (Fig. 1d and 2). On average, almost 6% of sperm of these Sorraia stallions contained cells with abnormal number of sex chromosomes. The prevalence of such sperm was as high as 10.8% in stallion E22 and the lowest 1.5% in stallion 23667 (Table 2). Notably, there was no correlation between the percentage of sex chromosome abnormalities and inbreeding ($r = -0.186$, $p = 0.724$), sperm morphology ($r = -0.255$, $p = 0.626$) or stallion age ($r = -0.223$, $p = 0.671$) (Tables 1 and 2).

FKBP6 genotypes

Direct sequencing of a PCR amplicon of *FKBP6* exon 4 revealed that all 11 Sorraia stallions had the same genotype being homozygous G/G for SNPs g.11040315G>A and homozygous A/A for SNP g.11040379C>A (Fig. 3). Sequencing the two control horses confirmed their known genotypes: the normal Thoroughbred control was heterozygous G/A-C/A, and the Thoroughbred with the IAR phenotype had the double homozygous A/A-A/A genotype – the susceptibility genotype for IAR (Raudsepp et al. 2012) (Fig. 3).

Discussion

Here, we present a detailed evaluation of Sorraia stallions by aligning data on sperm characteristics, inbreeding and fertility rates with chromosome analysis, sex chromosome sperm-FISH and molecular testing for a susceptibility genotype for impaired acrosome reaction. The study is the first of its kind in Sorraias and represents an important step for the conservation of this rare and endangered horse breed.

It is well established that the Sorraia horse breed has a critically small effective population size and high rate of inbreeding (Pinheiro et al. 2013; Kjöllerström et al. 2015). Inbreeding depression is probably the main reason why almost all Sorraia stallions analysed in this study had poor semen quality with suboptimal motility

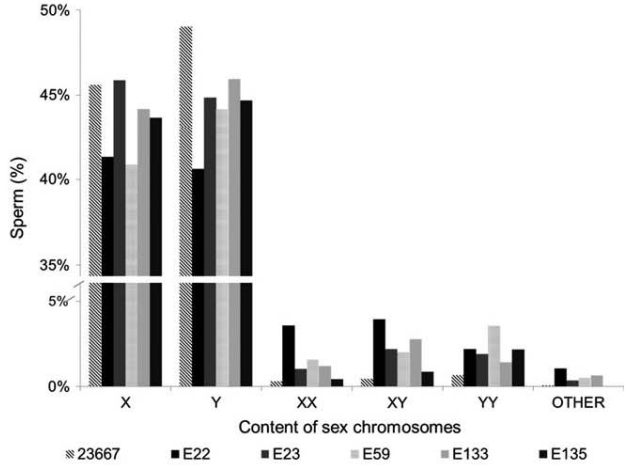


Fig. 2. Percentage of sex chromosome content of sperm from 6 stallions: 23667, E22, E23, E59, E133 and E135; 'Other' refers to sperm with XXY, XXX, XYY, YYY, YYYY or XXYY sex chromosomes

and morphological characteristics (Table 1). The rates of progressively motile and morphologically normal sperm in Sorraia stallions were just half of those reported for the Lusitano, another indigenous Iberian breed (Gamboa and Ramalho-Santos 2005; Gamboa et al. 2009), as well as for other horse breeds (Table S1). So was semen vitality which in Sorraias was 33.16% compared to, for example the 78.44% in the Thoroughbred (Table S1). It is noteworthy that almost all Sorraia semen parameters were inferior to those of the breeds with similar closed breeding structure, low genetic diversity and high rate of inbreeding, such as the Thoroughbred and the Friesian (Ducro 2011; McCue et al. 2012; Table S1). The only exception was Sorraia semen concentration which was one of the highest among all breeds (Table S1).

The genetic component of stallion fertility is complex and thought to involve interactions of multiple genes, details of which are, as yet, poorly understood (Raudsepp et al. 2013). Therefore, only a limited number of approaches are currently available to evaluate the genetic component of reproductive health of stallions. Among these, perhaps the best established is chromosome analysis by karyotyping to determine the possible presence of balanced chromosome abnormalities (Durkin et al. 2011; Ghosh et al. 2015; Raudsepp and Chowdhary 2016). Balanced chromosome rearrangements, such as translocations, do not typically affect the overall viability or health of the carrier and can remain unnoticed in non-breeding population. Yet, they can give rise to genetically unbalanced sperm which, if fertilized, will result in subfertility due to embryonic loss (Ghosh et al. 2015; Raudsepp and Chowdhary 2016). The overall frequency of such chromosome abnormalities in horses is very low, although a few cases have been described in Thoroughbred breeding stallions (Durkin et al. 2011). All Sorraia stallions in this study, however, were chromosomally normal, excluding chromosome abnormalities as a possible cause for their subfertility.

While karyotyping evaluates chromosomes in diploid somatic cells, sperm-FISH allows determining the chromosomal constitution of mature sperm and is potentially more informative for fertility evaluation. To date, the majority of sperm-FISH studies in domestic species, including horses, have analysed sperm of chromosomally and reproductively normal males to determine rates of diploidy and sex chromosome aneuploidy in normal sperm (Raudsepp and Chowdhary 2016). Rates of X and Y aneuploidy vary between species and even between different studies within a species but are overall the highest in horses and the lowest in pigs (Table 3).

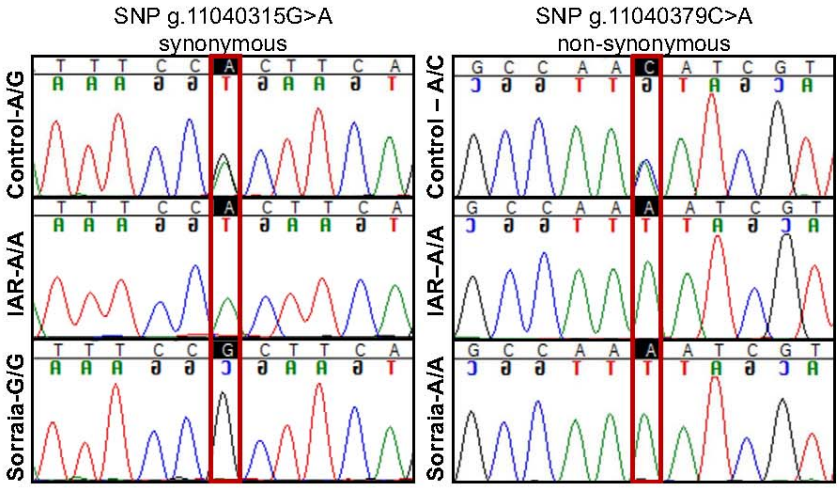


Fig. 3. Sequencing chromatograms showing genotypes for *FKBP6* exon 4 SNPs in a heterozygous control horse (A/G-A/C), a stallion with the IAR phenotype (A/A-A/A) and in a Sorraia stallion (G/G-A/A)

Table 3. Summary data on the rate of sperm sex chromosome aneuploidies in domestic species and men based on sperm-FISH

Species	Sperm phenotype	Sperm sex chr. aneuploidy; %	References
Sorraia horse	low motility; structurally abnormal sperm (teratozoospermia)	5.83	This study
Horse	normal	0.32–1.14	(Bugno et al. 2010; Bugno-Poniewierska et al. 2011, 2014)
Cattle	normal	0.10–3.09	(Hassanane et al. 1999; Di Berardino et al. 2004; Bugno et al. 2008; Nicodemo et al. 2009; Pauciullo et al. 2011)
Pig	normal	0.04–0.12	(Rubes et al. 1999; Bonnet-Garnier et al. 2009; Orsztynowicz et al. 2011)
Goat	normal	0.39	(Di Berardino et al. 2004)
River buffalo	normal	0.22	(Di Berardino et al. 2004)
Sheep	normal	0.03	(Di Berardino et al. 2004)
Dog	normal	0.13	(Komaki et al. 2014)
Human	normal	0.14–1.65	(Wyrobek et al. 1994; Martin et al. 1995; Robbins et al. 1995; Rubes et al. 2002; Kirkpatrick et al. 2008; Sarrate et al. 2010; Perry et al. 2011; Rodrigo et al. 2011)
Human	asthenoteratozoospermia	1.91	(Sarrate et al. 2010)
Human	asthenozoospermia	0.14–0.93	(Sarrate et al. 2010; Piomboni et al. 2014)
Human	azoospermia	0.35–1.13	(Rodrigo et al. 2011)
Human	globozoospermia	0.65	(Piomboni et al. 2014)
Human	oligoasthenoteratozoospermia	0.03–1.71; 14.50*; 18.63*; 22.00*	(In't Veld et al. 1997; Çinar et al. 2008; Kirkpatrick et al. 2008; Sarrate et al. 2010; Piomboni et al. 2014)
Human	oligoasthenozoospermia	2.35	(Sarrate et al. 2010)
Human	oligoteratozoospermia	0.47	(Piomboni et al. 2014)
Human	oligozoospermia	0.20–1.49	(Sarrate et al. 2010; Piomboni et al. 2014)
Human	teratoasthenozoospermia	0.40	(Piomboni et al. 2014)
Human	teratozoospermia	1.06	(Piomboni et al. 2014)

*Extremely high rates of sex chromosome abnormalities in three men with oligoasthenoteratozoospermia are shown separately.

In the present study, for the first time sperm-FISH was applied to the sperm of chromosomally normal stallions with poor semen quality. Similar sperm-FISH data are available only for men, although the sperm phenotypes in men have been more refined and range from oligospermia to severe oligoasthenoteratozoospermia (Piomboni et al. 2014; Table 3). Compared to subfertile men, the Sorraias had an overall higher rate of sex chromosome aneuploidies (5.83% in Sorraias vs 0.02–2.35% in men; Table 3), although in some human patients with oligoasthenoteratozoospermia, the rate of XY sperm ranged from 14% to 22% (In't Veld et al. 1997; Çinar et al. 2008; Kirkpatrick et al. 2008; Table 3), suggesting considerable individual variation within the same phenotype group.

Compared with normal stallions and with normal males of other domestic species, the rates of X and Y aneuploidies in Sorraia stallions were essentially higher (Table 3). The most frequent aneuploidy observed in this study was XY sperm (Table 2), suggesting that errors in meiosis I (MI) prevail over errors in meiosis II (MII), the latter leading to XX and YY sperm. Tendency for higher rates of MI errors has also been observed in normal males as reported for stallions (Bugno et al. 2010; Bugno-Poniewierska et al. 2011, 2014), bulls (Hassanane et al. 1999; Bugno et al. 2008), river buffalos, rams, billies (Di Berardino et al. 2004), dogs (Komaki et al. 2014) and men (Kirkpatrick et al. 2008; Piomboni et al. 2014).

Studies in men and stallions have found positive correlation between the rate of sperm aneuploidies and male age. For example, the rate of diploidy and disomy of chromosomes X, Y, 1, 9 and 21 shows significant increase in the sperm of older men (Egozcue et al. 2003; Templado et al. 2011), and older stallions have significantly more sperm with disomies XY, XX, YY and trisomy XXY (Bugno-Poniewierska et al. 2011). In contrast to these studies, no correlation between sperm aneuploidies and age was observed in Sorraia stallions. This was probably because our study cohort was small and relatively young (4–11 years) and did not allow estimation of age-related changes in the sperm.

Human studies also show that the rate of sex chromosome aneuploidies in the sperm of chromosomally normal but subfertile men is significantly correlated with the severity of sperm alterations (Piomboni et al. 2014). Similar tendency was observed in subfertile Sorraia stallions, that is the rate of sex chromosome aneuploidies increased with the reduction in the percentage of morphologically normal sperm (Tables 1 and 2). However, this correlation was not significant, most likely due to the small number of individuals analysed, and will require further study.

It is well established that inbreeding reduces fitness by exposing rare, deleterious alleles (Charlesworth and Willis 2009) and has been shown to negatively affect fertility in Sorraia stallions (Kjöllerström et al. 2015).

Therefore, it was somewhat surprising that we did not observe any correlation between sperm aneuploidies and the rate of inbreeding (Table 1). On the other hand, our findings are similar to those in Noriker stallions where sperm motility was not correlated with inbreeding (Aurich et al. 2003).

Lastly, even though sperm-FISH provides data rapidly for thousands of cells, its ability to detect aneuploidies is limited to the availability of chromosome-specific probes and the number of fluorochromes that can be simultaneously visualized under the microscope. Due to this, the majority of sperm-FISH in domestic species, including this study, are restricted to sex chromosomes (Raudsepp and Chowdhary 2016). Also, in this study, we simultaneously analysed just two fluorochromes – one for the X and another for the Y, and did not include an autosomal probe as internal control. Therefore, we were not able to distinguish between disomy and diploidy or between nullisomy and the lack of hybridization.

While traditional and molecular cytogenetic methods are readily available to assess the chromosomal component of stallion fertility, only a few molecular markers have been associated with fertility traits (Giesecke et al. 2010a,b; Raudsepp et al. 2012; Schrimpf et al. 2014, 2015). Of these, the *FKBP6* gene, a susceptibility locus for impaired acrosome reaction (IAR) in Thoroughbred stallions (Raudsepp et al. 2012), is the only one that has been associated with stallion sperm characteristics such as membrane fusion and acrosome reaction during fertilization. Besides, *FKBP6* is also involved in synaptonemal complex, meiotic pairing and segregation (Crackower et al. 2003). As Sorraia stallions with fertility problems are known to have poor quality acrosomal membrane (Gamboa and Ramalho-Santos 2005) and the stallions involved in this study had low fertility rate, poor semen quality and high rate of sex chromosome aneuploidies (Tables 1–3), it was justified to test them for the IAR susceptibility genotype. Stallions with IAR are homozygous A/A-A/A for a synonymous SNPs g.11040315G>A and a non-synonymous SNP g.11040379C>A in *FKBP6* exon 4 (Raudsepp et al. 2012). Yet, *FKBP6* exon 4 genotypes in all Sorraia stallions were invariably G/G-A/A, a testimony for their low genetic variability. The G/G-A/A genotype has been found in approximately 22% of reproductively normal stallions of various breeds and has not been associated with fertility traits (Raudsepp et al. 2012). Though, the non-synonymous SNP g.11040379C>A, alone, was recently associated with per-cycle pregnancy rate (PCPR) in Hanoverian stallions showing that stallions with A/A genotype have higher PCPR when compared to stallions with C/C genotype (Schrimpf et al. 2015). Given that all Sorraia stallions in this study were subfertile with poor sperm characteristics, connection between the A/A genotype and improved fertility remains obscure. It is possible that the *FKBP6* variants associated with IAR in Thoroughbreds (Raudsepp et al. 2012) or improved PCPR in Hanoverians (Schrimpf et al. 2015) are tagging underlying functional variants of other genes via regulatory pathways or synergistic interactions. Therefore, phenotypic effects of these regulatory variants may depend on the genetic make-up (haplotype structure) of the particular equine population or breed. Overall, *FKBP6* genotyping results in Sorraia stallions are intriguing, insisting that further research on *FKBP6* in stallion fertility is needed.

In summary, studying the Sorraia stallions with the available cytogenetic and molecular tools ruled out gross chromosomal abnormalities and genetic predisposition for IAR as possible genetic causes for subfertility. However, low fertility of these stallions could be partly attributed to relatively higher rate of sex chromosome aneuploidies in the sperm. Although the scope of these approaches is limited, they could be recommended for routine genetic evaluation of subfertile stallions until a more informative and complete set of molecular fertility markers becomes available.

Acknowledgements

This work was supported by FCT/MEC (Fundação para a Ciência e Tecnologia) PhD Fellowship for HJK (SFRH/BD/81502/2011). The authors also acknowledge support by LINK Equine Research Endowment, American Quarter Horse Foundation, USDA (grant #2012-67015-19632), cE3c FCT Unit funding (Ref. UID/BIA/00329/2013) and PRODER (Contract number 57882)(EU)/Action 2.2.3 (2011-2013) [Conservation and improvement of genetic resources. Sub-action 2.2.3.2. Animal component (Sorraia breed)]. The authors thank the Associação Internacional de Criadores do Cavalo Ibérico de Tipo Primitivo – Sorraia and the National Stud Farm in Alter do Chão, Portugal, for supplying the samples analysed in the present study. We also thank Dr. Sandra Gamboa from CERNAS/IIA/ESAC/IPC, Portugal, and Dr. Vitor Grácio and Dr. Maria José Correia from for the National Stud Farm in Alter do Chão, Portugal, for assistance in sample collection and Raquel Silva in sample processing.

Conflict of interest

None of the authors have any conflict of interest to declare.

Author contributions

The study was designed by TR, BPC, MM and HJK; experiments were conducted by HJK and data analysed by HJK and TR. The manuscript was drafted by TR, HJK, BPC and MM.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1 Seminal characteristics of stallions from different breeds

References

Aurich C, Achmann R, Aurich JE, 2003: Semen parameters and level of microsatellite heterozygosity in Noriker draught horse stallions. Theriogenology **60**, 371–378.

Ballou J, 1983: Calculating inbreeding coefficients from pedigrees. In: Schonewald-Cox CM, Chambers SM, MacBryde B, Thomas L(eds), Genetics and Conservation: a Reference for Managing Wild Animal and Plant Populations. Benjamin/Cummings, Menlo Park, California, pp. 509–520.

Bonnet-Garnier A, Guardia S, Pinton A, Ducos A, Yerle M, 2009: Analysis using sperm-FISH of a putative interchromosomal effect in boars carrying reciprocal translocations. Cytogenet Genome Res **126**, 194–201.

Bugno M, Jablonska Z, Slota E, 2008: Double Colour Fluorescence *in Situ* Hybridization with X and Y Painting Probes in Bovine Spermatozoa. Ann Anim Sci **8**, 303–309.

Bugno M, Jablonska Z, Tischner M, Klukowska-Rotzler J, Pienkowska-Schelling A, Schelling C, Slota E, 2010: Detection of sex chromosome aneuploidy in equine spermatozoa using fluorescence *in situ* hybridization. Reprod Domest Anim **45**, 1015–1019.

Bugno-Poniewierska M, Jablonska Z, Slota E, 2009: Modification of equine sperm chromatin decondensation method to use fluorescence *in situ* hybridization (FISH). Folia Histochem Cytobiol **47**, 663–666.

Bugno-Poniewierska M, Kozub D, Pawlina K, Tischner M Jr, Tischner M, Slota E, Wnuk M, 2011: Determination of the correlation between stallion's age and number of sex chromosome aberrations in spermatozoa. Reprod Domest Anim **46**, 787–792.

Bugno-Poniewierska M, Pawlina K, Tischner M, Tischner M, 2014: Age-related effects on sex chromosome aberrations in equine spermatozoa. J Equine Vet Sci **34**, 34.

Charlesworth D, Willis JH, 2009: The genetics of inbreeding depression. Nat Rev Genet **10**, 783–796.

Çinar Ç, Yazici C, Ergünsu Ş, Beyazyürek Ç, Javadova D, Sağlam Y, Tarcan T, Güney AI, 2008: Genetic diagnosis in infertile men with numerical and constitutional sperm abnormalities. Genet Test **12**, 195–202.

Crackower MA, Kolas NK, Noguchi J, Sarao R, Kikuchi K, Kaneko H, Kobayashi E, Kawai Y, Kozieradzki I, Landers R, Mo R, Hui CC, Nieves E, Cohen PE, Osborne LR, Wada T, Kunieda T, Moens PB, Penninger JM, 2003: Essential role of Fkbp6 in male fertility and homologous chromosome pairing in meiosis. Science **300**, 1291–1295.

Das P.J., Paria N., Gustafson-Seabury A., Vishnoi M., Chaki S.P., Love C.C., Varner D.D., Chowdhary B.P., Raudsepp T., 2010: Total RNA isolation from stallion sperm and testis biopsies. Theriogenology, **74**, 1099–1106, 106e1-2

Di Berardino D, Vozdova M, Kubickova S, Cernohorska H, Coppola G, Coppola G, Enne G, Rubes J, 2004: Sexing river buffalo (Bubalus bubalis L.), sheep (Ovis aries L.), goat (Capra hircus L.), and cattle spermatozoa by double color FISH using bovine (Bos taurus L.) X- and Y-painting probes. Mol Reprod Dev **67**, 108–115.

Ducro BJ, 2011: Relevance of Test Information in Horse Breeding. Wageningen University, Wageningen, The Netherlands. p. 170.

Durkin KW, Raudsepp T, Chowdhary BP, 2011: Cytogenetic evaluation. In: McKinnon AO, Squires EL, Vaala WE, Varner DD (eds), Equine Reproduction. Wiley-Blackwell, Chichester, pp. 1462–1468.

Egozcue J, Blanco J, Anton E, Egozcue S, Sarrate Z, Vidal F, 2003: Genetic analysis of sperm and implications of severe male infertility–a review. Placenta **24**(Suppl B), S62–S65.

Gamboa S, Ramalho-Santos J, 2005: SNARE proteins and caveolin-1 in stallion spermatozoa: possible implications for fertility. Theriogenology **64**, 275–291.

Gamboa S, Machado-Faria M, Ramalho-Santos J, 2009: Seminal traits, suitability for semen preservation and fertility in the native Portuguese horse breeds Puro Sangue Lusitano and Sorraia: implications for stallion classification and assisted reproduction. Anim Reprod Sci **113**, 102–113.

Ghosh S, Das PJ, Avila F, Thwaits BK, Chowdhary BP, Raudsepp T, 2015: A Non-Reciprocal Autosomal Translocation 64,XX, t(4;10)(q21;p15) in an Arabian Mare with Repeated Early Embryonic Loss. Reprod Domest Anim **51**, 171–174.

Giesecke K, Hamann H, Sieme H, Distl O, 2010a: INHBA-associated markers as candidates for stallion fertility. Reprod Domest Anim **45**, 342–347.

Giesecke K, Sieme H, Distl O, 2010b: Infertility and candidate gene markers for fertility in stallions: a review. Vet J **185**, 265–271.

Hassanane M, Kovacs A, Laurent P, Lindblad K, Gustavsson I, 1999: Simultaneous detection of X- and Y-bearing bull spermatozoa by double colour fluorescence *in situ* hybridization. Mol Reprod Dev **53**, 407–412.

In't Veld PA, Broekmans FJ, de France HF, Pearson PL, Pieters MH, van Kooij RJ, 1997: Intracytoplasmic sperm injection (ICSI) and chromosomally abnormal spermatozoa. Hum Reprod **12**, 752–754.

ISCNH, 1997: International system for cytogenetic nomenclature of the domestic horse. Chromosome Res **5**, 433–443.

ISIS, 2011: SPARKS (Single Population Analysis and Records Keeping System). International Species Information System, Eagan, MN.

Kirkpatrick G, Ferguson KA, Gao H, Tang S, Chow V, Yuen BH, Ma S, 2008: A comparison of sperm aneuploidy rates between infertile men with normal and abnormal karyotypes. Hum Reprod **23**, 1679–1683.

Kjöllerström HJ, Gama LT, Oom MM, 2015: Impact of inbreeding on fitness-related traits in the highly threatened Sorraia horse breed. Livestock Sci **180**, 84–89.

Komaki H, Oi M, Suzuki H 2014: Detection of sex chromosome aneuploidy in dog spermatozoa by triple color fluorescence *in situ* hybridization. Theriogenology **82**, 652–656.

Luis C, Cothran EG, Oom MM, 2007a: Inbreeding and genetic structure in the endangered Sorraia horse breed: implications for its conservation and management. J Hered **98**, 232–237.

Luis C, Juras R, Oom MM, Cothran EG, 2007b: Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. Anim Genet **38**, 20–27.

Martin RH, Spriggs E, Ko E, Rademaker AW, 1995: The relationship between paternal age, sex ratios, and aneuploidy frequencies in human sperm, as assessed by multicolor FISH. Am J Hum Genet **57**, 1395–1399.

McCue ME, Bannasch DL, Petersen JL, Gurr J, Bailey E, Binns MM, Distl O, Guerin G, Hasegawa T, Hill EW, Leeb T, Lindgren G, Penedo MC, Roed KH, Ryder OA, Swinburne JE, Tozaki T, Valberg SJ, Vaudin M, Lindblad-Toh K, Wade CM, Mickelson JR, 2012: A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. PLoS Genet **8**, e1002451.

Montgomery GW, Sise JA, 1990: Extraction of DNA from sheep white blood cells. N Z J Agri Res **33**, 437–441.

Nicodemo D, Pauciullo A, Castello A, Roldan E, Gomendio M, Cosenza G, Peretti V, Perucatti A, Di Meo GP, Ramunno L, Iannuzzi L, Rubes J, Di Berardino D, 2009: X-Y sperm aneuploidy in 2 cattle (Bos taurus) breeds as determined by dual color fluorescent *in situ* hybridization (FISH). Cytogenet Genome Res **126**, 217–225.

Oom MM, d'Andrade JL & Costa-Ferreira J, 2004: Studbook da raça Sorraia (in Portuguese). Garrido Artes Gráficas, Alpiarça, Portugal.

Orsztynowicz M, Pawlak P, Oles D, Kubickova S, Lechniak D, 2011: Low incidence of chromosome aberrations in spermatozoa of fertile boars. Reprod Biol **11**, 224–235.

Pauciullo A, Cosenza G, Peretti V, Iannuzzi A, Di Meo GP, Ramunno L, Iannuzzi L, Rubes J, Di Berardino D, 2011: Incidence of X-Y aneuploidy in sperm of two indigenous cattle breeds by using dual

color fluorescent *in situ* hybridization (FISH). Theriogenology **76**, 328–333.
Pellestor F, Andreo B, Taneja K, Williams B, 2003: PNA on human sperm: a new approach for *in situ* aneuploidy estimation. Eur J Hum Genet **11**, 337–341.
Perry MJ, Chen X, McAuliffe ME, Maity A, Deloid GM, 2011: Semi-automated scoring of triple-probe FISH in human sperm: methods and further validation. Cytometry A **79**, 661–666.
Pinheiro M, Kjöllerström HJ, Oom MM, 2013: Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis. Livestock Sci **152**, 1–10.
Piomboni P, Stendardi A, Gambera L, 2014: Chromosomal aberrations and aneuploidies of spermatozoa. Adv Exp Med Biol **791**, 27–52.
Raudsepp T, Chowdhary BP, 2008: FISH for mapping single copy genes. Methods Mol Biol **422**, 31–49.
Raudsepp T, Chowdhary B 2016: Chromosome Aberrations and Fertility Disorders in Domestic Animals. In: Lewin A (ed.) Annual Review of Animal Biosciences. Annual Reviews, Palo Alto, CA. in press
Raudsepp T, Santani A, Wallner B, Kata SR, Ren C, Zhang HB, Womack JE, Skow LC, Chowdhary BP, 2004: A detailed physical map of the horse Y chromosome. Proc Natl Acad Sci USA **101**, 9321–9326.
Raudsepp T, Gustafson-Seabury A, Durkin K, Wagner ML, Goh G, Seabury CM, Brinkmeyer-Langford C, Lee EJ, Agarwala R, Stallknecht-Rice E, Schaffer AA, Skow LC, Tozaki T, Yasue H, Penedo MC, Lyons LA, Khazanehdari KA, Binns MM, MacLeod JN, Distl O, Guerin G, Leeb T, Mickelson JR, Chowdhary BP, 2008: A 4,103 marker integrated physical and comparative map of the horse genome. Cytogenet Genome Res **122**, 28–36.
Raudsepp T, McCue ME, Das PJ, Dobson L, Vishnoi M, Fritz KL, Schaefer R, Rendahl AK, Derr JN, Love CC, Varner DD, Chowdhary BP, 2012: Genome-wide association study implicates testis-sperm specific FKBP6 as a susceptibility locus for impaired acrosome reaction in stallions. PLoS Genet **8**, e1003139.
Raudsepp T, Das PJ, Chowdhary BP, 2013: Genomics of Reproduction and Fertility. In: Chowdhary BP (ed.) Equine Genomics. Blackwell Publishing Ltd, Chichester, pp. 199–215.
Robbins WA, Baulch JE, Moore D 2nd, Weier HU, Blakey D, Wyrobek AJ, 1995: Three-probe fluorescence *in situ* hybridization to assess chromosome X, Y, and 8 aneuploidy in sperm of 14 men from two healthy groups: evidence for a paternal age effect on sperm aneuploidy. Reprod Fertil Dev **7**, 799–809.
Rodrigo L, Rubio C, Peinado V, Villamon R, Al-Asmar N, Remohi J, Pellicer A, Simon C, Gil-Salom M, 2011: Testicular sperm from patients with obstructive and nonobstructive azoospermia: aneuploidy risk and reproductive prognosis using testicular sperm from fertile donors as control samples. Fertil Steril **95**, 1005–1012.
Rubes J, Vozdova M, Kubickova S, 1999: Aneuploidy in pig sperm: multicolor fluorescence *in situ* hybridization using probes for chromosomes 1, 10, and Y. Cytogenet Cell Genet **85**, 200–204.
Rubes J, Vozdova M, Robbins WA, Rezacova O, Perreault SD, Wyrobek AJ, 2002: Stable variants of sperm aneuploidy among healthy men show associations between germinal and somatic aneuploidy. Am J Hum Genet **70**, 1507–1519.
Sarrate Z, Vidal F, Blanco J, 2010: Role of sperm fluorescent *in situ* hybridization studies in infertile patients: indications, study approach, and clinical relevance. Fertil Steril **93**, 1892–1902.
Schrimpf R, Dierks C, Martinsson G, Sieme H, Distl O, 2014: Genome-wide association study identifies phospholipase C zeta 1 (PLCz1) as a stallion fertility locus in Hanoverian warmblood horses. PLoS ONE **9**, e109675.
Schrimpf R, Metzger J, Martinsson G, Sieme H, Distl O, 2015: Implication of FKBP6 for male fertility in horses. Reprod Domest Anim **50**, 195–199.
StatSoft 2013: Statistica Version 12.
Templado C, Donate A, Giraldo J, Bosch M, Estop A, 2011: Advanced age increases chromosome structural abnormalities in human spermatozoa. Eur J Hum Genet **19**, 145–151.
Wyrobek AJ, Alhborn T, Balhorn R, Stanker L, Pinkel D, 1990: Fluorescence *in situ* hybridization to Y chromosomes in decondensed human sperm nuclei. Mol Reprod Dev **27**, 200–208.
Wyrobek AJ, Robbins WA, Mehraein Y, Pinkel D, Weier HU, 1994: Detection of sex chromosomal aneuploidies X-X, Y-Y, and X-Y in human sperm using two-chromosome fluorescence *in situ* hybridization. Am J Med Genet **53**, 1–7.

Submitted: 24 Jan 2016; Accepted: 27 Feb 2016

Author's address (for correspondence): Terje Raudsepp, Department of Veterinary Integrative Biosciences, College of Veterinary Medicine & Biomedical Sciences, Texas A&M University, College Station, Texas, USA. E-mail: traudsepp@cvm.tamu.edu

Supplementary Table 1. Seminal characteristics of stallions from different breeds; AC – after collection; NA - not available

Breed	Semen volume ± SE mL	Sperm concentration $\times 10^6$/mL	Motility AC %	Vitality %	Sperm with normal morphology %	Reference
Sorraia	**26.91±37.19**	**307.14±215.89**	**21.92±15.22**	**33.16±22.47**	**28.39±20.21**	**This Study**
Sorraia	21.64±14.44	232.15±140.68	20.71±9.34	41.41±16.75	51.47±16.57	Gamboa *et al.* 2009
Lusitano	46.78±28.91	217.07±133.03	46.70±11.91	68.15±19.97	77.78±15.64	Gamboa *et al.* 2009
Thoroughbred	28.33±0.09	114.29±0.53	70.73±1.06	78.44±1.07	NA	Dowsett & Knott 1996
Standardbred	30.20±0.09	97.24±0.64	78.82±1.07	84.64±1.09	NA	Dowsett & Knott 1996
Arabian	36.20±0.10	286.82±0.86	85.02±1.08	89.90±1.10	NA	Dowsett & Knott 1996
Australian Stock Horse	33.19±0.10	116.06±0.73	77.79±1.08	82.88±1.09	NA	Dowsett & Knott 1996
Quarter Horse	22.80±0.14	171.66±1.29	73.92±1.11	76.19±1.13	NA	Dowsett & Knott 1996
Palomino	23.78±0.13	138.48±1.13	72.85±1.10	78.68±1.12	NA	Dowsett & Knott 1996
Pony	20.78±0.23	104.01±1.09	69.65±1.10	75.21±1.11	NA	Dowsett & Knott 1996
Shetland	44.39±0.49	101.25±4.37	70.06±1.21	61.53±1.25	NA	Dowsett & Knott 1996
Appaloosa	23.28±0.19	90.42±1.10	73.20±1.14	84.19±1.16	NA	Dowsett & Knott 1996
Miniature	24.18±0.76	233.60±11.43	63.80±0.70	NA	54.28±1.05	Paccamonti *et al.* 1999
Dutch Warmblood	65±26	20.61±0.17	68±9	NA	66±15	Parlevliet *et al.* 1994
Pura Raza Española	40–50	>200	60-70	70–80	NA	Gamboa et al 2009
Thoroughbred and Standardbred	NA	NA	NA	NA	52.5±20.1	Jasko *et al.* 1990
19 breeds	45±30	335±232	53±15	NA	51±15	Pickett *et al.* 1988

7 breeds	45±31	178±168	72±16	NA	NA	Dowsett & Pattie 1982
NA	NA	NA	50±2.5	NA	46±2.2	Brinsko *et al.* 2000
Quarter Horse	20.9±2.6	161.0±20.5	26.6±4.5	NA	44.0±2.8	Naden *et al.* 1990
Arabian, Spanish, Cross-breed	49.2±0.2	304.8±1.9	88.9±0.1	NA	9.1±0.06	Quintero-Moreno *et al.* 2003
Estonian	NA	NA	NA	NA	74.4±3.8	Kavak *et al.* 2004
Tori	NA	NA	NA	NA	57.5±4.1	Kavak *et al.* 2004
Swedish Warmblood	NA	225.09±57.55	67.00±7.34	NA	66.60±3.97	Morrell *et al.* 2008
Swedish Warmblood	29.8±1.8	405.9±31.2	78.9±0.9	NA	72.7±1.5	Einarsson *et al.* 2009
Mixed breeds	66.2±4.9	276.1±25.1	73.3±1.5	NA	NA	Wrench *et al.* 2010
Thai native crossbreed	44.0±2.1	NA	77.8±1.3	NA	49.7±1.3	Phetudomsinsuk *et al.* 2008
Spanish thoroughbred	54	185	88.8	NA	71.6	Hidalgo *et al.* 2008
Hanoverian	44.25±21.92	206.85±88.68	64.4±13.2	NA		Labitzke *et al.* 2014
Arabian	29.90±2.99	266.83±31.97	73.07±3.04	NA	60.48± 2.98	Waheed *et al.* 2015
Andalusian donkey	76.3±6.6	240.3±21.5	89.1±1.8	NA	87.1±1.4	Dorado *et al.* 2013
Austrian Noriker	90.8±55.1	243±114	50±23	NA	38±18	Aurich *et al.* 2003
Shetland pony	23.50±12.15	234.55±185.35	66.05±10.80	NA	57.45±15.95	van Eldik *et al.* 2006

References

Aurich C., Achmann R. & Aurich J.E. 2003 Semen parameters and level of microsatellite heterozygosity in Noriker draught horse stallions. *Theriogenology* **60**, 371-8.

Brinsko S.P., Van Wagner G.S., Graham J.K. & Squires E.L. 2000 Motility, morphology and triple stain analysis of fresh, cooled and frozen-thawed stallion spermatozoa. *J Reprod Fertil Suppl*, 111-20.

Dorado J., Acha D., Ortiz I., Galvez M.J., Carrasco J.J., Diaz B., Gomez-Arrones V., Calero-Carretero R. & Hidalgo M. 2013 Relationship between conventional semen characteristics, sperm motility patterns and fertility of Andalusian donkeys Equus asinus. *Anim Reprod Sci* **143**, 64-71.

Dowsett K.F. & Knott L.M. 1996 The influence of age and breed on stallion semen. *Theriogenology* **46**, 397-412.

Dowsett K.F. & Pattie W.A. 1982 Characteristics and fertility of stallion semen. *J Reprod Fertil Suppl* **32**, 1-8.

Einarsson S., Dalin A.M. & Lundeheim N. 2009 Sperm production and sperm morphology of Swedish Warmblood stallions. *Reprod Domest Anim* **44**, 33-6.

Gamboa S., Machado-Faria M. & Ramalho-Santos J. 2009 Seminal traits, suitability for semen preservation and fertility in the native Portuguese horse breeds Puro Sangue Lusitano and Sorraia: Implications for stallion classification and assisted reproduction. *Anim Reprod Sci* **113**, 102-13.

Hidalgo M., Rodriguez I., Dorado J. & Soler C. 2008 Morphometric classification of Spanish thoroughbred stallion sperm heads. *Anim Reprod Sci* **103**, 374-8.

Jasko D.J., Lein D.H. & Foote R.H. 1990 Determination of the relationship between sperm morphologic classifications and fertility in stallions: 66 cases 1987-1988. *J Am Vet Med Assoc* **197**, 389-94.

Kavak A., Lundeheim N., Aidnik M. & Einarsson S. 2004 Sperm morphology in Estonian and Tori breed stallions. *Acta Vet Scand* **45**, 11-8.

Labitzke D., Sieme H., Martinsson G. & Distl O. 2014 Genetic parameters and breeding values for semen characteristics in Hanoverian stallions. *Reprod Domest Anim* **49**, 584-7.

Morrell J.M., Johannisson A., Dalin A.M., Hammar L., Sandebert T. & Rodriguez-Martinez H. 2008 Sperm morphology and chromatin integrity in Swedish warmblood stallions and their relationship to pregnancy rates. *Acta Vet Scand* **50**, 2.

Naden J., Amann R.P. & Squires E.L. 1990 Testicular growth, hormone concentrations, seminal characteristics and sexual behaviour in stallions. *J Reprod Fertil* **88**, 167-76.

Paccamonti D.L., Buiten A.V., Parlevliet J.M. & Colenbrander B. 1999 Reproductive parameters of miniature stallions. *Theriogenology* **51**, 1343-9.

Parlevliet J.M., Kemp B. & Colenbrander B. 1994 Reproductive characteristics and semen quality in maiden Dutch Warmblood stallions. *J Reprod Fertil* **101**, 183-7.

Phetudomsinsuk K., Sirinarumitr K., Laikul A. & Pinyopummin A. 2008 Morphology and head morphometric characters of sperm in Thai native crossbred stallions. *Acta Veterinaria Scandinavica* **50**, 41-.

Pickett B.W., Voss J.L., Bowen R.A., Squires E.L. & McKinnon A.O. 1988 Seminal characteristics and total scrotal width T.S.W. of normal and abnormal stallions. *Proceedings of the annual convention of the American Association of Equine Practitioners USA*, 487.

Quintero-Moreno A., Miro J., Teresa Rigau A. & Rodriguez-Gil J.E. 2003 Identification of sperm subpopulations with specific motility characteristics in stallion ejaculates. *Theriogenology* **59**, 1973-90.

van Eldik P., van der Waaij E.H., Ducro B., Kooper A.W., Stout T.A. & Colenbrander B. 2006 Possible negative effects of inbreeding on semen quality in Shetland pony stallions. *Theriogenology* **65**, 1159-70.

Waheed M.M., Ghoneim I.M. & Abdou M.S.S. 2015 Morphometric Characteristics of Spermatozoa in the Arabian Horse With Regard to Season, Age, Sperm Concentration, and Fertility. *Journal of Equine Veterinary Science* **35**, 244-9.

Wrench N., Pinto C.R., Klinefelter G.R., Dix D.J., Flowers W.L. & Farin C.E. 2010 Effect of season on fresh and cryopreserved stallion semen. *Anim Reprod Sci* **119**, 219-27.

CHAPTER 4 - DISCUSSION

This discussion is mainly focused on presenting the findings of the present study. For more detailed results and breed comparisons, the respective papers presented in the former chapters should be consulted.

4.1 Breed Characterization

Genetic diversity conservation is crucial for sustainable management of animal genetic resources and can be achieved by selection programmes that restore genetic diversity (Ajmone-Marsan 2010). Protecting a domestic breed and guaranteeing its survival depends on a minimum number of individuals that guarantee the essential genetic diversity and make possible adaptation to environmental changes (Frankham *et al.* 2004). Structure, diversity, gene flow and demography assessment are important to determine how endangered a population is, and to device an effective selection and conservation program (Goyache *et al.* 2003; Gutierrez *et al.* 2003; Fernández *et al.* 2004; Cervantes *et al.* 2008). Pedigrees are useful to describe genetic variability and evolution within populations, across generations (Boichard *et al.* 1997). They can be used independently or in complement of molecular analysis. In our study, pedigree results were consistent with molecular analyses, revealing low diversity in the Sorraia horse.

The small number of founders, reduced effective population size, complete genetic isolation and unbalanced use of animals for reproduction (particularly males) have led to genetic variability decline over time. The use of the same stallion in consecutive years in the same breeding farm has led to founder loss and uneven genetic contribution regarding all available males (Kjöllerström 2005; Pinheiro *et al.* 2013).

In agreement with molecular data, pedigree analysis has shown that only two of the original seven matrilineal lines exist today (Luís *et al.* 2002). Ideally founder contribution should be equivalent (Lacy 1989) but in the Sorraia three founders represent more than half of the population's genetic variability. The effective number of founders (f_e) is 7.46, with two founders no longer represented and the genetic contribution of underrepresented founders at great risk of loss (Luis *et al.* 2007a; Pinheiro *et al.* 2013). The effective number of ancestors (f_a) was four, substantially lower than f_e. Inadequate breeding strategies since the breed's foundation such as

unbalanced use of stallions for reproduction and lack of veterinarian support on reproductive performance assessment and improvement validate the low fa and fe values, with a small number of animals explaining the overall reduced genetic variability, while also leading to uneven progeny between sexes.

Generation length influences the rate of genetic progress and is important in breeding programs. In Sorraias, this value (7.94 yr) was lower than reported in other horse breeds. Mare values were slightly higher than sire values, mostly because mares are kept for reproduction as long as possible while sires are replaced more frequently. The number of international Sorraia horse breeders has increased in recent years although most living horses are still bred in Portuguese studs. Schäfer's Stud is the most important international breeder and the first established in Germany, functioning as a source for several new German subpopulations. Due to the establishment of new breeders and the public's increasing interest in this breed, the number of births per year has increased over time (the total population reaching almost 300 animals in 2011), particularly from 1987 to 2006 with about 51% of registered animals. This also allows a more efficient genetic management, by maximizing the use of all available sires and mares, currently underway, leading in turn to an improvement in fertility and progeny survival.

The quality of genealogical information in the Sorraia Studbook is very high, with 98% known ancestors in the first generation. For the second and third generations this value is over 80%, as older animals might have founders early in the genealogy with no information on former ancestors. With records starting in 1937, mean complete number of generations was 8.17 and 13 maximum number of generations in the extant population.

Average inbreeding (F), mean kinship (mK) and average relatedness (AR) reported herein are extremely high and far from those reported in other livestock breeds. Over half of registered animals have 25% inbreeding or higher (Pinheiro *et al.* 2013). Inbred births appeared as soon as 1942 and increased 5.2% per generation throughout the years, far from the 1% limit established by FAO (1998), however, average inbreeding coefficient per year has stabilized (or even decreased) especially over the last decade. The most feasible management practice of the breed has been to avoid inbreeding on each stud farm by choosing a different stallion to be crossed with mares each season. Unfortunately, the most genetically important animals, as ranked by the individual value of mK, are reaching post-reproductive age and are limited by geographical constraints.

In order to select the most genetically influential stallions, breeders regularly refer to the Sorraia Breeders Association for genetic analysis simulations. Breeder's awareness and involvement in these conservation efforts is increasing, as information forums are held by the Sorraia Breeders Association. All efforts have given positive results with recently registered animals having below average inbreeding coefficients.

Sorraia Breeders Association has also implemented a reproduction strategy based on selecting breeders by minimizing mean kinship (mK), to the best of our knowledge for the first time in a horse breed, and also promotes stallion exchange between stud farms. Ovulation and pregnancy monitoring by ultrasonography and semen evaluation has also been carried out, in an increasingly widespread way, in order to improve pregnancy and survival rates of, for example, conceptuses. The definition of breeding strategies to minimize inbreeding and maximize genetic diversity is extremely important, as is the continued analysis of their results that underlie future decisions and allow the achievement of defined objectives.

In conservation management programs of AnGR, a breed's phenotypic characterization is of utmost importance as basic information, complementing genetic and historical information (Melo *et al.* 2011). This is fundamental for the establishment of national AnGR inventories and to effectively monitor populations, establishing the basis for which to start management breeding programs (FAO 2012). Morphometric measurements also give valuable data for instituting realistic breed standards, as well as to define breed's deviation from them in different generations.

Regarding Sorraia stallion morphology, comparing 26 body measurements in two groups 20 years apart, OLD and NEW, revealed a wider distribution of all analysed variables in the latter, mainly due to higher number of individuals. Morphometric values were higher, on average, in the OLD group. Inbreeding and age were higher in the NEW. Overall average height was 145.96cm, and 16% of the NEW animals were taller than the described 148cm standard (Oom 1992; Oom *et al.* 2004). Stallions have become shorter in 20 years, although this difference in height measured at the withers was not statistically significant. We have validated the results obtained by Oom (1992), proving the Sorraia to be shorter than the Lusitano (Oom & Ferreira 1987; Oom 1992; Solé *et al.* 2013; Vicente *et al.* 2014a), but taller than the Garrano (Oom 1992; Santos & Ferreira 2012) and Terceira ponies (Lopes *et al.* 2015). A breed's body conformation ratio (withers height/body length) is important to define the purpose for which horses

are bred. The Sorraia is a well proportionate horse with a ratio ≈1 and square body shape, eumetric and mediline (Oom 1992) and should be considered a horse and not a pony. Withers height is slightly higher than croup height, which is also in accordance with the results found in the Lusitano (Oom & Ferreira 1987; Oom 1992), the most relevant Portuguese saddle horse. Chest circumference has increased and cannon bone perimeter decreased over time in Sorraia stallions. Sorraias are now used daily in riding centres where higher aerobic capacity is preferred and are being bred mostly for leisure and sport, whereas before they were mainly used in reproduction or cattle driving and selected for stronger bone structure. Feeding differences and genetics could also explain these results.

Age had a significant negative impact on a few measurements that could be explained by overall loss of body condition with advancing age. To the best of our knowledge, this is the first report that correlates coat colour and body measurements, demonstrating that animals with a specific coat colour (dark mouse dun) are shorter than others (mouse and yellow mouse duns).

Studies in horses on the effects of inbreeding on morphometric traits are limited. In the Lusitano, Oom (1992) and Vicente *et al.* (2014b) reported a significant ($p<0.01$) detrimental effect of inbreeding on withers height, as did Gandini *et al.* (1992) in the Italian Haflinger horse. It is peculiar that inbreeding, that has increased over time in the Sorraia breed, had so little effect on morphological traits in stallions, however, it is known that inbreeding depression is more severe on fitness and life-history traits than on morphological ones (Coltman & Slate 2003; Carolino & Gama 2008). This could also be due to improved management and handling (nutrition and training) compared to 20 years ago, potentially masking the detrimental effects of significant increase in inbreeding coefficients over time. Another possibility is the purging of deleterious recessive alleles (Lacy *et al.* 1996; Leberg & Firmin 2008) from the Sorraia horse population over generations and through natural selection. Additionally, since only one stallion is used per herd, per year, using a shorter stallion would lead to a decrease in height in the progeny that wasn't due to inbreeding but to trait heritability.

Principal component analysis (PCA) is a very useful tool to visualize relationships between individuals using different input data, positioning individuals in a 2D/3D graph and grouping them by similarity (Almeida 2007). It has been used in horse morphometric studies to determine breed relationships (Gómez *et al.* 2012; Solé *et al.* 2013), breeding goals (Cervantes *et al.* 2009), or stud farms (Zechner *et al.* 2001;

Sobczuk & Komosa 2012). In our study, OLD and NEW samples did not separate into different clusters, showing the breed's great homogeneity in morphometric measurements, together with their high correlation, and also the small number of animals in group OLD.

While inbreeding depression on the majority of traits analysed on the Sorraia horse haven't been alarming thus far, studying these effects is very important as morphological conformation can determine an animal's athletic aptitude and vigour. This is pertinent for reproductive success as interactions between mares and stallions mainly occur extensively in natural settings. Conformational traits can also allude to potential inbreeding depression in other characteristics and should be considered in the management-breeding plan.

As suggested by Lacy *et al.* (1996), sample sizes should be big enough as to allow for the detection of statistically significant results of inbreeding depression. In our case, increasing sample size would improve the statistical power of the analysis, although it was not an easy task to find handled stallions to be measured.

Genome-wide variability in the Sorraia horse has decrease over time, as expected in a breed with only 12 founders that has been managed as closed population since 1937.

Our STRs analysis results represent an improvement in the analysed parameters relatively to previous ones that included Sorraia samples (Aberle *et al.* 2004; Luis *et al.* 2007a; Luis *et al.* 2007b), but it is still lower than in other horse breeds (Aurich *et al.* 2003; Curik 2003; Giacomoni *et al.* 2008; Costa *et al.* 2010; Mittmann *et al.* 2010; Gupta *et al.* 2012). These results were expected as markers were chosen based on their high variability in the Sorraia breed. A few of the newly tested STRs showed a great potential for replacing some of those currently used in parentage testing, increasing the efficacy of the test.

There was no polymorphism in Y-linked *loci*, in agreement with the results of Wallner (2004) and Kakoi *et al.* (2005). This is a result of the stronger selection in males throughout horse evolution and domestication, where only a reduced number of stallions would contribute to subsequent generations (Lindgren *et al.* 2004). In the Sorraia, this is particularly truthful since most breeding is done extensively and only one stallion is used per year, per herd (Oom 2006).

Despite the increase in inbreeding, that lead to decreased genetic variation and to inbreeding depression, mean d^2 and individual heterozygosity have improved since

2007 (Luis *et al.* 2007a). These results show the positive results of the implemented Sorraia Breeders Association management-breeding plan.

When comparing the Portuguese (PT) and German (GER) populations using STR or SNP data, we found that GER had better values in general for the genetic parameters analysed than PT. This is likely a direct result of the different breeding systems, validating that using one stallion per mare increases the number of stallions used yearly in GER, consequently increasing genetic variation and decreasing inbreeding in this population.

STRUCTURE analysis of STR data organized our samples into three clusters (K=3): cluster 1 with animals related to stud farm # 4; cluster 2 contains animals from the original German (#9 and 10) and a Portuguese stud farms (#13), whose animals were bought (or are direct descendants) from #3; cluster 3 contains animals from the founder stud farm (#2, 3, 5 and 8), as well as animals from studs 7, 11 and the most recent German stud (#1). This analysis validates this breed's historic evolution as well as their breeders. It is interesting to note how this analysis shows the admixture between stud farms implemented by the Sorraia Breeders Association in recent years, visible in a few stud farms (e.g. #4 and 7) where animals have almost a 50/50 contribution from different sources.

Average inbreeding calculated using SNPs was much lower than the one obtained from pedigree data, with both values positively and significantly correlated. High-density SNP-chips are presently used to calculate inbreeding coefficients (de Cara *et al.* 2011; Lenstra *et al.* 2012; Curik *et al.* 2014; Knief *et al.* 2015), however, pedigree-based inbreeding calculations should not be neglected, especially if dealing with highly inbred individuals, as the Sorraia, where SNPs might have restricted power (Wang 2016). In Metzger *et al.* (2015) runs of homozygosity (ROH) based inbreeding calculations (total length of ROHs divided by genome length) in two Sorraia horses were in agreement with the current population average (Kjöllerström *et al.* 2015). These results show that using ROHs from sequencing data is suitable for inbreeding calculations in this extremely inbred breed.

We found higher heterozygosity in STRs than in SNPs, as did Khanshour (2013) in Arabian horses, possibly due to SNPs' bi-allelic nature and lower mutation rate. Additionally, our STRs were chosen based on their higher variability in the Sorraia breed, with subsequent higher heterozygosities. No correlation between STRs or SNP heterozygosities was found, in agreement with the different distributions.

We found ROH regions in 11 chromosomes, some of them common to more than one horse. It was interesting that chromosome 27, despite being one of the smallest chromosomes had the third biggest ROH, and that chromosome 2, one of the largest, had the smallest. Since short and long ROHs are signatures of ancient and recent inbreeding, respectively (Zhang *et al.* 2015), ancient inbreeding due to small ROH length was most common in Sorraia horses. There were a few overlapping ROH segments between our study and the one of Metzger *et al.* (2015), referring to other six horse breeds.

Contrary to PCA analysis using morphometric measurements, and despite STRUCTURE separating our samples from 13 different breeders into three main clusters, FCA (STR data) and PCA (SNP data) analysis ordered our total population into two main groups: Portugal and "old" German breeders. In this analysis, animals from the most recent German breeder (#1) are still non-distinguishable from the Portuguese due to newly imported animals. The second principal component does not separate Portuguese breeders due to the Sorraia Breeders Association efforts to encourage stallion exchange, also explaining why there weren't three clusters as in STRUCTURE analysis. These results also agree with the low F_{ST} found between PT and GER.

None of the Sorraia specific CNVs found was ubiquitous, attesting to the variable nature of these markers. Array CGH analysis revealed 213 CNVs: 53 gains and 160 losses, with an average 26.6 CNVs per horse, close to the results found by Doan *et al.* (2012b) and Ghosh *et al.* (2014) in different horse breeds. Adjacent and overlapping CNV calls were arranged into 71 CNVRs, of which 7 were Sorraia-specific. The number of CNVs in Sorraia horses (n=213) was close to that found by Doan *et al.* (2012b) (n=282), higher than in Metzger *et al.* (2013) (n=50), but lower than in Doan *et al.* (2012a), Dupuis *et al.* (2013), Ghosh *et al.* (2014) and Wang *et al.* (2014) (2368, 2797, 950 and 700, respetively). The number of CNVRs in Sorraia was also lower than found in other horse breeds (Doan *et al.* 2012a; Dupuis *et al.* 2013; Ghosh *et al.* 2014; Wang *et al.* 2014; Ghosh *et al.* 2016). We found two novel CNVs involving genes *ATP10D* and *RFX3*. Although adequate for disease study, CNVs are not suitable to evaluate the degree of genetic diversity in this extremely inbred breed, as there were no fixed CNVs and individual variation in CNVs was high.

Parentage testing and genetic variability analysis for management purposes in the Sorraia will continue to be done using STRs rather than SNPs, since STR genotyping can be done at a fraction of the cost of SNPs, and there was higher variability and heterozygosity in the former.

Our genome-wide results represent an improved evaluation and a broad overview on the extant genetic variation in the Sorraia breed resulting from the multi-method approach. They offer new possibilities to increase heterozygosity and reduce inbreeding by choosing the best breeding pair combinations each year. Therefore this data should be included in the current management breeding program as it would allow for a quicker genetic recovery and more spot-on conservation of this extremely endangered animal genetic resource.

4.2 Reproductive fitness

Reproductive success is extremely important for conservation and management. Therefore determining the possible causes of reduced fertility noted in this endangered breed is essential. Additionally, early detection of animals that will show poor breeding performance in the future can save time, and breeders' resources.

Inbreeding increased since the breed's foundation and all births from 1960 onward are inbred (Kjöllerström *et al.* 2015). Inbreeding is thus inevitable in the Sorraia breed and can only be managed in order to minimize its increase, as has been done by the Sorraia Breeders Association. Inbreeding depression has been demonstrated in other equid populations with increasing kinship between breeding animals causing reduced fitness. An increase in juvenile mortality and decrease in life expectancy throughout successive generations of inbreeding have been reported (Ryder & Wedemeyer 1982). In the Przewalski horse it has been shown that animals with inbreeding coefficients of 0.25 or higher had fewer descendants (possibly due to inbreeding affecting ovarian function) (Collins *et al.* 2012) and had shorter life expectancies than those with lower inbreeding (Bouman & Bos 1979). In the Lusitano breed, Oom (1992) showed that animals with higher inbreeding coefficients were more likely to produce non-viable offspring.

Mean fertility rates were 46.74% and 35.79% in stallions and mares, respectively (Kjöllerström *et al.* 2015) similar to those reported for the Lusitano and Garrano (Gomes & Oom 2000). Stallion fertility might be overestimated due to the effect of compensatory mating. Mean stallion and mare fertility rates are lower than in other breeds, possibly due to extremely high inbreeding and lower semen quality (Gamboa *et al.* 2009). Somewhat unexpectedly, the proportion of non-viable animals is low (Kjöllerström *et al.* 2015). Due to practice of extensive breeding management, recurrent in this breed, it is important to point out that determining the exact number of neo-natal deaths, abortions and foetus reabsorptions is difficult to report, potentially underestimating fertility rates in this breed. Comparatively, Lusitano (Barbosa & Abreu 1986) and Przewalski (Monfort *et al.* 1994) mares have much lower abortions and stillborns percentages. We found mare fertility to be significantly and negatively

influenced by inbreeding, which is concerning if we consider the above-mentioned fertility under- and overestimation.

Even though our results were based on a larger data set over a longer time-period than previous ones (Oom *et al.* 1991; Matos 1996), we could only find a weak relationship between foal inbreeding and viability. This discrepancy might be caused by husbandry conditions disguising inbreeding depression effects (Kalinowski *et al.* 2000; Kalinowski & Hedrick 2001), or due to residual genetic variability possibly overshadowing inbreeding depression. Additionally, natural selection and genetic drift might have also played an important role in purging some of the negative effects of inbreeding over the years, as suggested by Lacy (2000).

Early detection of chromosomal abnormalities can save time and resources of breeders as affected animals will later on in life have reduced breeding performance. This fact substantiates the importance of implementing a systematic cytogenetic analysis to help provide definitive diagnosis for cases of reduced fertility in this critically endangered horse breed. Most sex-chromosome abnormalities correspond to normal phenotypes, so a higher percentage of particularly difficult breeding cases may be attributed to chromosomally abnormal animals that remain barren despite best efforts from veterinarians and breeders.

With this purpose in mind, we found the first evidence of sex chromosome mosaicism in the Sorraia horse in a mare (Kjöllerström *et al.* 2011). This abnormality was described for the first time by Chandley *et al.* (1975), by the co-occurrence of normal female (64,XX) and X monosomic (63,X0) cells. This karyotype is typically associated to sub fertility, an irregular or absent oestrus cycle, smaller body size, normal external genitalia and gonadal dysgenesis (Hughes & Trommershausen-Smith 1977; Chowdhary & Raudsepp 2000). Individuals with this karyotype might be sub fertile or infertile depending on the proportion of the two cell lines in the gonads (Chandley *et al.* 1975; Reid *et al.* 1987; Bugno *et al.* 2001; Wieczorek *et al.* 2001). Our results explain the low fertility of the studied mare. This is the first time that this specific mosaic has been linked to masculine body conformation and a stallion-like behaviour, proving that a chromosomal aberration may result in different phenotypes, even normal ones. We analysed the progeny of this mare, one stallion and one filly also highly inbred, looking for chromosomal abnormalities, with negative results.

We also looked for chromosomal abnormalities in a representative sample of the Sorraia population. All analysed Sorraia stallions were chromosomally normal (64,XX in females and 64,XY in males), excluding chromosome abnormalities as possible subfertility causes.

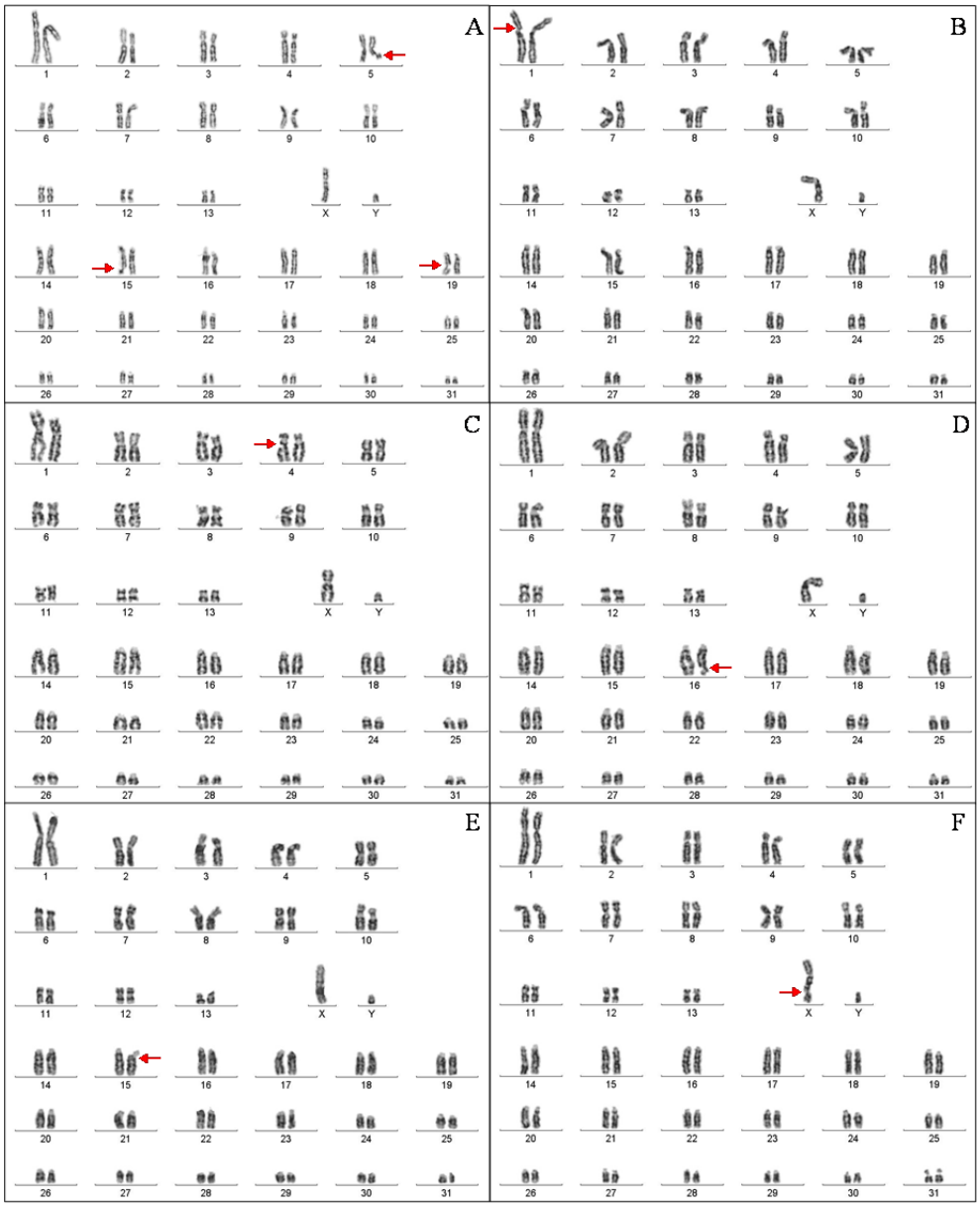


Figure 9 - Chromosome breaks in Sorraia horses. A through E are from one individual, F is from a different animal.

Two of the analysed stallions had chromosome breaks in a few metaphases (Figure 9), although these breaks were not present in all cells nor were they consistently in the same location. Therefore, we did not determine the exact location of breakpoints by FISH, although GTG-banded karyotypes of some animals were successfully obtained

and the pairs with the breaks identified (Figure 9). These breaks might, however, be located in regions where genes responsible for DNA repair are present, leading to chromosome instability. This could potentially explain why one of the karyotyped animals (Figure 9, A-E) died at a younger age than what is considered normal for this breed (20 years rather than 30).

In the future, the possibility of using X and Y whole-chromosome painting probes on metaphase and interphase nuclei (Raudsepp & Chowdhary 2008) will increase the efficiency and accuracy of karyotyping tests, reducing the time needed to detect sub fertility caused by sex chromosome aneuploidies (Breen *et al.* 1997; Wieczorek *et al.* 2001; Bugno *et al.* 2006; Bugno *et al.* 2008).

Karyotyping assesses diploid somatic cells, while sperm-FISH determines the chromosomal complement of mature sperm. Most sperm-FISH studies investigated sperm of chromosomally and reproductively normal males to assess diploidy and aneuploidy rates in normal sperm (Raudsepp & Chowdhary 2016).

Sperm FISH was done for the first time in chromosomally normal Sorraia stallions with known poor semen quality. We combined information on sperm characteristics, inbreeding and fertility rates with chromosome analysis, sex chromosome sperm-FISH and molecular testing in order to look for a susceptibility genotype for impaired acrosome reaction (IAR). This could be a valuable tool in the conservation of this breed with its useful value in fertility evaluation. Almost all Sorraia stallions semen parameters were half of those reported in other horse breeds (Dowsett & Knott 1996; Gamboa & Ramalho-Santos 2005; Gamboa *et al.* 2009), even of those with comparable closed breeding systems, low genetic diversity and high inbreeding (Ducro 2011; McCue *et al.* 2012). Inbreeding depression might be the cause behind this poor semen quality, suboptimal motility and morphological characteristics.

Aneuploidies of X and Y chromosomes in Sorraia stallions were higher than in normal males of other domestic species. The most common is XY sperm, suggesting that meiosis I (MI) errors prevail over meiosis II (MII) ones. Higher rates of MI errors have been previously described in normal stallions (Bugno *et al.* 2010; Bugno-Poniewierska *et al.* 2011; Bugno-Poniewierska *et al.* 2014). Contrary to studies in men and stallions (Egozcue *et al.* 2003; Bugno-Poniewierska *et al.* 2011; Templado *et al.* 2011), no correlation between sperm aneuploidies and age was observed in Sorraia stallions. This is probably a result of the small sample and reduced age classes of the

stallions analysed, not allowing age-related effects assessment. The rate of sex chromosome aneuploidies increased with morphologically normal sperm reduction, as in humans (Piomboni *et al.* 2014), although not significantly, again due to the small study cohort, requiring further study. We did not include an autosomal probe as internal control and thus were unable to differentiate disomy from diploidy, or nullisomy from lack of hybridization. Inbreeding negatively affects Sorraia stallions fertility (Kjöllerström *et al.* 2015), thus it was surprising to find no correlation between sperm aneuploidies and inbreeding.

Few molecular markers have been associated with stallion fertility so far (Giesecke *et al.* 2010a; Giesecke *et al.* 2010b; Raudsepp *et al.* 2012; Schrimpf *et al.* 2014; Schrimpf *et al.* 2015). *FKBP6* is a gene involved in synaptonemal complex, meiotic pairing and segregation (Crackower *et al.* 2003). It is a susceptibility locus for impaired acrosome reaction (IAR) in Thoroughbred stallions (Raudsepp *et al.* 2012), associated in fertilization with membrane fusion and acrosome reaction. Sorraia stallions have poor quality acrosomal membrane (Gamboa & Ramalho-Santos 2005), low fertility rates, poor semen quality and high rate of sex chromosome aneuploidies, justifying IAR susceptibility testing. IAR stallions are homozygous A/A-A/A in *FKBP6* exon 4 (Raudsepp *et al.* 2012). None of Sorraia stallions had the IAR susceptibility genotype and were all G/G-A/A, evidencing low genetic variability. This particular genotype was found in 22% of reproductively normal stallions from other breeds and has not been associated with fertility traits (Raudsepp *et al.* 2012). *FKBP6* non-synonymous SNP g.11040379C>A (A/A) was recently associated with higher per-cycle pregnancy rate in Hanoverian stallions (Schrimpf *et al.* 2015). As all Sorraia stallions were sub fertile, with poor sperm characteristics, this association between A/A genotype and improved fertility remains unclear. These *FKBP6* variants might be associated to underlying functional variants of other genes whose phenotypic effects may possibly depend on the genetic basis of the particular population or breed. *FKBP6* genotyping results in Sorraia stallions are intriguing and require further investigation.

Although studying Sorraia stallions using cytogenetic and gene sequencing tools excluded large chromosomal abnormalities and IAR as likely subfertility causes, their low fertility could be partly attributed to high rates of sperm sex chromosome aneuploidies.

CHAPTER 5 - FINAL REMARKS

The techniques described herein should be included in the genetic evaluation of breeding Sorraias and further studies should be implemented, analysing different samples and traits. As a result of the detected significance of inbreeding depression, these studies should also be undertaken in Sorraia mares and compared to those of stallions to understand if it is differentially affecting the sexes.

To achieve conservation goals we must not neglect the hypothesis of taking, in some particular cases, full advantage of the benefits of assisted reproductive technologies (ART). This would maximize the use of breeding animals in the population, even those with low fertility rates, and the retention of genetic variability of those most genetically important, even after death.

Sorraia breeding and conservation strategies should include information regarding signs of inbreeding depression to establish a long-term self-sustaining population and the breeds' evolution should be tracked with the help of long term-studies.

Due to their high inbreeding coefficients all Sorraias are genetically very similar, as far as the genetic markers used could detect. With the availability and decreasing cost of next generation sequencing tools, it would be of great interest to sequence and compare whole-sequence data of a few selected animals with well-characterized fertility phenotypes. This would help look for signs of inbreeding depression, giving us an insight into fertility-related genes that we currently know little about and cost the equine industry millions of dollars every year, at least until a more informative set of fertility markers is available. Sequencing entire genomes could potentially give the results that SNP and CNV data could not as they do not contain the entire horse genome.

The next step is to continue SNP analysis looking for signs of selection and significant areas potentially linked to fertility or disease. Additionally, genotype imputation could allow the comparison of Sorraia and other horse breeds that have been analysed with the 50K-Equine SNP chip (McCue & Mickelson 2013; McCoy & McCue 2014) in order to disentangle their relationships. This will be done in collaboration with other groups.

It is obviously necessary to continue the implementation of the conservation management-breeding plan, which is essential to promote the control of the extremely high levels of inbreeding currently observed, assuming the maintenance of a long-term self-sustaining population is intended. Stallions should continue to be chosen yearly for each herd according to genetic suitability, and stallion rotation between stud farms, or yearly individual mare and stallion matings, should be encouraged. As noted in our

results comparing Portuguese and German populations, bringing a few selected German stallions for extensive breeding in Portuguese stud farms (and vice-versa) would help to maximize the retention of the genetic diversity still existing in the population, as a whole, and decrease inbreeding. These measures will continue to aid genetic and demographic improvement of the Sorraia horse, preventing future additional genetic erosion.

It is of paramount importance to preserve this animal genetic resource of worldwide interest. As it has not been subjected to directional selection to enhance particular traits, the primitive characteristics and hardiness can only be preserved if this breed is allowed to reproduce extensively. Human intervention can never be excluded as we have seen preservation benefits from management and conservation plans that give continuity to the implemented multidisciplinary approach.

Despite all efforts and considering our results showing extremely high and increasing inbreeding levels, low genetic variability and some recent evidence of inbreeding depression in the breed (Kjöllerström *et al.* 2015), the introduction of new animals from genetically close breeds might have to be considered by the Sorraia Breeders Association, in the future and as an exceptional and extreme case, in order to improve genetic health and prevent the permanent loss of this iconic and important horse breed. The stated increasing number of births per year is, nevertheless, a great hope for the long-term sustainability of this endangered national genetic resource.

References

Aberle K.S., Hamann H., Drogemuller C. & Distl O. (2004) Genetic diversity in German draught horse breeds compared with a group of primitive, riding and wild horses by means of microsatellite DNA markers. *Anim Genet* **35**, 270-7.

Ajmone-Marsan P. (2010) A global view of livestock biodiversity and conservation-GLOBALDIV. *Anim Genet* **41 Suppl 1**, 1-5.

Almeida P.A.R. (2007) Diversidade genética e diferenciação das raças portuguesas de ovinos com base em marcadores de DNA - microssatélites: uma perspectiva de conservação. p. 305. Universidade de Trás-os-Montes e Alto Douro, Vila Real.

Aurich C., Achmann R. & Aurich J.E. (2003) Semen parameters and level of microsatellite heterozygosity in Noriker draught horse stallions. *Theriogenology* **60**, 371-8.

Barbosa M. & Abreu J.V. (1986) Alguns parâmetros reprodutivos em cavalos lusitanos e árabes. *Revista Portuguesa de Ciências Veterinárias* **478**, 147-69.

Boichard D., Maignel L. & Verrier É. (1997) The value of using probabilities of gene origin to measure genetic variability in a population. *Genet Sel Evol* **29**, 5-23.

Bouman J.G. & Bos H. (1979) Two symptoms of inbreeding depression in Przewalski horses living in captivity. In: *Genetics and hereditary diseases of the Przewalski horse* (eds. by Boer LEM, Bouman J & Bouman I), pp. 111-7. Foundation for the Preservation and Protection of the Przewalski Horse, Rotterdam – The Netherlands.

Breen M., Langford C.F., Carter N.P., Fischer P.E., Marti E., Gerstenberg C., Allen W.R., Lear T.L. & Binns M.M. (1997) Detection of equine X chromosome abnormalities in equids using a horse X whole chromosome paint probe (WCPP). *Vet J* **153**, 235-8.

Bugno-Poniewierska M., Kozub D., Pawlina K., Tischner M., Jr., Tischner M., Slota E. & Wnuk M. (2011) Determination of the correlation between stallion's age and number of sex chromosome aberrations in spermatozoa. *Reprod Domest Anim* **46**, 787-92.

Bugno-Poniewierska M., Pawlina K., Tischner M. & Tischner M. (2014) Age-related effects on sex chromosome aberrations in equine spermatozoa. *Journal of Equine Veterinary Science* **34**, 34.

Bugno M., Jablonska Z., Tischner M., Klukowska-Rotzler J., Pienkowska-Schelling A., Schelling C. & Slota E. (2010) Detection of sex chromosome aneuploidy in equine spermatozoa using fluorescence in situ hybridization. *Reprod Domest Anim* **45**, 1015-9.

Bugno M., Pieńkowska-Schelling A., Schelling C., Włodarczak N. & Słota E. (2006) Utilization of the probe generated by chromosome microdissection, for detection of equine X chromosome aneuploidy. *Ann. Anim. Sci.* **6**, 205-10.

Bugno M., Slota E. & Zabek T. (2001) Two cases of subfertile mares with 64,XX/63,X mosaic karyotype. *Ann. Anim. Sci.* **1**, 7–11.

Bugno M., Zabek T., Golonka P., Pienkowska-Schelling A., Schelling C. & Slota E. (2008) A case of an intersex horse with 63,X/64,XX/65,XX,del(Y)(q?) karyotype. *Cytogenetic and genome research* **120**, 123-6.

Carolino N. & Gama L.T. (2008) Inbreeding depression on beef cattle traits: estimates, linearity of effects and heterogeneity among sire-families. *Genet Sel Evol* **40**, 511-27.

Cervantes I., Baumung R., Molina A., Druml T., Gutiérrez J.P., Sölkner J. & Valera M. (2009) Size and shape analysis of morphofunctional traits in the Spanish Arab horse. *Livest Sci* **125**, 43-9.

Cervantes I., Molina A., Goyache F., Gutiérrez J.P. & Valera M. (2008) Population history and genetic variability in the Spanish Arab Horse assessed via pedigree analysis. *Livest Sci* **113**, 24-33.

Chandley A.C., Fletcher J., Rossdale P.D., Peace C.K., Ricketts S.W., McEnery R.J., Thorne J.P., Short R.V. & Allen W.R. (1975) Chromosome abnormalities as a cause of infertility in mares. *J Reprod Fertil Suppl*, 377-83.

Chowdhary B.P. & Raudsepp T. (2000) Cytogenetics and physical gene maps. In: *The genetics of the horse* (eds. by Bowling AT & Ruvinsky A), pp. 171-241. CABI Publishing, Oxon, UK.

Collins C.W., Songsasen N.S., Vick M.M., Wolfe B.A., Weiss R.B., Keefer C.L. & Monfort S.L. (2012) Abnormal reproductive patterns in Przewalski's mares are associated with a loss in gene diversity. *Biol Reprod* **86 (2)**, 1-10.

Coltman D.W. & Slate J. (2003) Microsatellite measures of inbreeding: a meta-analysis. *Evolution* **57**, 971-83.

Costa M.A., Bressel R.M., Almeida D.B., Oliveira P.A., Bassini L.N., Moreira C.G., Manzke V.H., Siewerdt F. & Moreira H.L. (2010) Genotyping in the Brazilian Criollo Horse Stud Book: resources and perspectives. *Genet Mol Res* **9**, 1645-53.

Crackower M.A., Kolas N.K., Noguchi J., Sarao R., Kikuchi K., Kaneko H., Kobayashi E., Kawai Y., Kozieradzki I., Landers R., Mo R., Hui C.C., Nieves E., Cohen P.E., Osborne L.R., Wada T., Kunieda T., Moens P.B. & Penninger J.M. (2003) Essential role of *FKBP6* in male fertility and homologous chromosome pairing in meiosis. *Science* **300**, 1291-5.

Curik I. (2003) Inbreeding, Microsatellite Heterozygosity, and Morphological Traits in Lipizzan Horses. *J Hered* **94**, 125-32.

Curik I., Ferenčaković M. & Sölkner J. (2014) Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livest Sci* **166**, 26-34.

de Cara M.A.R., Fernández J., Toro M.A. & Villanueva B. (2011) Using genome-wide information to minimize the loss of diversity in conservation programmes. *J Anim Breed Genet* **128**, 456-64.

Doan R., Cohen N., Harrington J., Veazey K., Juras R., Cothran G., McCue M.E., Skow L. & Dindot S.V. (2012a) Identification of copy number variants in horses. *Genome Res* **22**, 899-907.

Doan R., Cohen N.D., Sawyer J., Ghaffari N., Johnson C.D. & Dindot S.V. (2012b) Whole-genome sequencing and genetic variant analysis of a Quarter Horse mare. *BMC Genomics* **13**, 78.

Dowsett K.F. & Knott L.M. (1996) The influence of age and breed on stallion semen. *Theriogenology* **46**, 397-412.

Ducro B.J. (2011) Relevance of Test Information in Horse Breeding. p. 170. Wageningen University, Wageningen, The Netherlands.

Dupuis M.C., Zhang Z., Durkin K., Charlier C., Lekeux P. & Georges M. (2013) Detection of copy number variants in the horse genome and examination of their association with recurrent laryngeal neuropathy. *Anim Genet* **44**, 206-8.

Egozcue J., Blanco J., Anton E., Egozcue S., Sarrate Z. & Vidal F. (2003) Genetic analysis of sperm and implications of severe male infertility--a review. *Placenta* **24 Suppl B**, S62-5.

FAO (1998) *Secondary Guidelines for Development of National Farm Animal Genetic Resources Management Plans: Management of small populations at risk*. Food and Agriculture Organization of the United Nations, Rome, Italy.

FAO (2012) Phenotypic characterization of animal genetic resources. FAO Animal Production and Health Guidelines, Rome.

Fernández J., Toro M.A. & Caballero A. (2004) Managing Individuals' Contributions to Maximize the Allelic Diversity Maintained in Small, Conserved Populations. *Conserv Biol* **18**, 1358-67.

Frankham R., Ballou J.D. & Briscoe D.A. (2004) *A Primer of Conservation Genetics*. Cambridge University Press.

Gamboa S., Machado-Faria M. & Ramalho-Santos J. (2009) Seminal traits, suitability for semen preservation and fertility in the native Portuguese horse breeds Puro Sangue Lusitano and Sorraia: Implications for stallion classification and assisted reproduction. *Anim Reprod Sci* **113**, 102-13.

Gamboa S. & Ramalho-Santos J. (2005) SNARE proteins and caveolin-1 in stallion spermatozoa: possible implications for fertility. *Theriogenology* **64**, 275-91.

Gandini G.C., Bagnato A., Miglior F. & Pagnacco G. (1992) Inbreeding in the Italian Haflinger horse. *J Anim Breed Genet* **109**, 433-43.

Ghosh S., Das P.J., McQueen C.M., Gerber V., Swiderski C.E., Lavoie J.P., B P.C. & Raudsepp T. (2016) Analysis of genomic copy number variation in equine recurrent airway obstruction (heaves). *Anim Genet*.

Ghosh S., Qu Z., Das P.J., Fang E., Juras R., Cothran E.G., McDonell S., Kenney D.G., Lear T.L., Adelson D.L., Chowdhary B.P. & Raudsepp T. (2014) Copy Number Variation in the Horse Genome. *PLoS Genet* **10**, e1004712.

Giacomoni E.H., Fernandez-Stolz G.P. & Freitas T.R. (2008) Genetic diversity in the Pantaneiro horse breed assessed using microsatellite DNA markers. *Genet Mol Res* **7**, 261-70.

Giesecke K., Hamann H., Sieme H. & Distl O. (2010a) INHBA-associated markers as candidates for stallion fertility. *Reprod Domest Anim* **45**, 342-7.

Giesecke K., Sieme H. & Distl O. (2010b) Infertility and candidate gene markers for fertility in stallions: a review. *Vet J* **185**, 265-71.

Gomes J.C. & Oom M.M. (2000) Caracterização demográfica do núcleo de garranos bravios no PNPG (in Portuguese). *Revista Portuguesa de Zootecnia* **1**, 67-78.

Gómez M.D., Azor P.J., Alonso M.E., Jordana J. & Valera M. (2012) Morphological and genetic characterization of Spanish heavy horse breeds: Implications for their conservation. *Livest Sci* **144**, 57-66.

Goyache F., Gutiérrez J.P., Fernández I., Gomez E., Alvarez I., Díez J. & Royo L.J. (2003) Using pedigree information to monitor genetic variability of endangered populations: the Xalda sheep breed of Asturias as an example. *J Anim Breed Genet* **120**, 95-105.

Gupta A.K., Chauhan M., Bhardwaj A. & Tandon S.N. (2012) Microsatellite markers based genetic diversity and bottleneck studies in Zanskari pony. *Gene* **499**, 357-61.

Gutierrez J.P., Altarriba J., Diaz C., Quintanilla R., Canon J. & Piedrafita J. (2003) Pedigree analysis of eight Spanish beef cattle breeds. *Genet Sel Evol* **35**, 43-63.

Hughes J.P. & Trommershausen-Smith A. (1977) Infertility in the horse associated with chromosomal abnormalities. *Aust Vet J* **53**, 253-7.

Kakoi H., Hirota K., Gawahara H., Kurosawa M. & Kuwajima M. (2005) Genetic diagnosis of sex chromosome aberrations in horses based on parentage test by

microsatellite DNA and analysis of X- and Y-linked markers. *Equine Vet J* **37**, 143-7.
Kalinowski S.T. & Hedrick P.W. (2001) Inbreeding depression in captive bighorn sheep. *Anim Conserv* **4**, 319-24.
Kalinowski S.T., Hedrick P.W. & Miller P.S. (2000) Inbreeding Depression in the Speke's Gazelle Captive Breeding Program. *Conserv Biol* **14**, 1375-84.
Khanshour A. (2013) Genetic diversity and population structure of the Arabion horse populations from Syria and other countries. p. 135. VIBS, Texas A&M University, College Station, Texas.
Kjöllerström H.J. (2005) O Studbook como instrumento na análise genética e demográfica de populações ameaçadas: o caso do Cavalo do Sorraia (in Portuguese). In: *Animal Biology Department*, p. 50. Faculty of Sciences, University of Lisbon, Portugal, University of Lisbon, Portugal.
Kjöllerström H.J., Collares-Pereira M.J. & Oom M.M. (2011) First evidence of sex chromosome mosaicism in the endangered Sorraia Horse breed. *Livest Sci* **136**, 273-6.
Kjöllerström H.J., Gama L.T. & Oom M.M. (2015) Impact of inbreeding on fitness-related traits in the highly threatened Sorraia horse breed. *Livest Sci* **180**, 84-9.
Knief U., Hemmrich-Stanisak G., Wittig M., Franke A., Griffith S.C., Kempenaers B. & Forstmeier W. (2015) Quantifying realized inbreeding in wild and captive animal populations. *Heredity* **114**, 397-403.
Lacy R.C. (1989) Analysis of founder representation in pedigrees: Founder equivalents and founder genome equivalents. *Zoo Biology* **8**, 111-23.
Lacy R.C. (2000) Considering Threats to the Viability of Small Populations Using Individual-Based Models. *Ecological Bulletins*, 39-51.
Lacy R.C., Alaks G. & Walsh A. (1996) Hierarchical Analysis of Inbreeding Depression in *Peromyscus polionotus*. *Evolution* **50**, 2187-200.
Leberg P.L. & Firmin B.D. (2008) Role of inbreeding depression and purging in captive breeding and restoration programmes. *Mol Ecol* **17**, 334-43.
Lenstra J.A., Groeneveld L.F., Eding H., Kantanen J., Williams J.L., Taberlet P., Nicolazzi E.L., Sölkner J., Simianer H., Ciani E., Garcia J.F., Bruford M.W., Ajmone-Marsan P. & Weigend S. (2012) Molecular tools and analytical approaches for the characterization of farm animal genetic diversity. *Anim Genet* **43**, 483-502.
Lindgren G., Backstrom N., Swinburne J., Hellborg L., Einarsson A., Sandberg K., Cothran G., Vila C., Binns M. & Ellegren H. (2004) Limited number of patrilines in horse domestication. *Nat Genet* **36**, 335-6.
Lopes M.S., Mendonça D., Rojer H., Cabral V., Bettencourt S.X. & da Câmara Machado A. (2015) Morphological and genetic characterization of an emerging Azorean horse breed: the Terceira Pony. *Front Genet* **6**, 62.
Luís C., Bastos-Silveira C., Cothran E.G. & Oom M.M. (2002) Variation in the mitochondrial control region sequence between the two maternal lines of the Sorraia horse breed. *Genetics and Molecular Biology* **25**, 309-11.
Luis C., Cothran E.G. & Oom M.M. (2007a) Inbreeding and genetic structure in the endangered Sorraia horse breed: implications for its conservation and management. *J Hered* **98**, 232-7.
Luis C., Juras R., Oom M.M. & Cothran E.G. (2007b) Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. *Anim Genet* **38**, 20-7.

Matos S.M.S. (1996) Contributo para a gestão genética do Cavalo do Sorraia (in Portuguese). In: *Animal Biology Department*, p. 44. University of Lisbon, Portugal, University of Lisbon, Portugal.

McCoy A.M. & McCue M.E. (2014) Validation of imputation between equine genotyping arrays. *Anim Genet* **45**, 153.

McCue M. & Mickelson J. (2013) Genomic Tools and Resources: Development and Applications of an Equine SNP Genotyping Array. In: *Equine Genomics* (pp. 113-24. Blackwell Publishing Ltd.

McCue M.E., Bannasch D.L., Petersen J.L., Gurr J., Bailey E., Binns M.M., Distl O., Guerin G., Hasegawa T., Hill E.W., Leeb T., Lindgren G., Penedo M.C., Roed K.H., Ryder O.A., Swinburne J.E., Tozaki T., Valberg S.J., Vaudin M., Lindblad-Toh K., Wade C.M. & Mickelson J.R. (2012) A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. *PLoS Genet* **8**, e1002451.

Melo J.B., Pires D.A.F., Ribeiro M.N., Santos D.O. & Silva H.G.O. (2011) Estudo zoométrico de remanescentes da raça equina Nordestina no município de Floresta, Pernambuco – Brasil. *AICA* **1**, 71-4.

Metzger J., Karwath M., Tonda R., Beltran S., Agueda L., Gut M., Gut I.G. & Distl O. (2015) Runs of homozygosity reveal signatures of positive selection for reproduction traits in breed and non-breed horses. *BMC Genomics* **16**, 764.

Metzger J., Philipp U., Lopes M.S., da Camara Machado A., Felicetti M., Silvestrelli M. & Distl O. (2013) Analysis of copy number variants by three detection algorithms and their association with body size in horses. *BMC Genomics* **14**, 487.

Mittmann E.H., Lampe V., Momke S., Zeitz A. & Distl O. (2010) Characterization of a minimal microsatellite set for whole genome scans informative in warmblood and coldblood horse breeds. *J Hered* **101**, 246-50.

Monfort S.L., Arthur N.P. & Wildt D.E. (1994) Reproduction in the Przewalski's horse. In: *Przewalski's horse: the history and biology of an endangered species.* (eds. by Boyd L & Houpt KA), pp. 173–94. Suny Press, New York.

Oom M. & Ferreira J. (1987) Estudo biométrico do cavalo Alter. *Rev Port Cienc Vet* **83**, 101-48.

Oom M.M. (1992) O cavalo Lusitano. Uma raça em Recuperação (in Portuguese). p. 299, University of Lisbon, Portugal.

Oom M.M. (2006) *O Cavalo do Sorraia. Um património histórico e genético a conservar (in Portuguese)*. Edições Cosmos, Portugal.

Oom M.M., Costa-Ferreira J. & Cothran E.G. (1991) Inbreeding, Reproductive Success and Genetic Variation in the Sorraia Horse of Portugal. *Anim Genet* **11**, 22-3.

Oom M.M., d'Andrade J.L. & Costa-Ferreira J. (2004) *Studbook da raça Sorraia (in Portuguese)*. Garrido Artes Gráficas, Alpiarça, Portugal.

Pinheiro M., Kjöllerström H.J. & Oom M.M. (2013) Genetic diversity and demographic structure of the endangered Sorraia horse breed assessed through pedigree analysis. *Livest Sci* **152**, 1-10.

Piomboni P., Stendardi A. & Gambera L. (2014) Chromosomal aberrations and aneuploidies of spermatozoa. *Adv Exp Med Biol* **791**, 27-52.

Raudsepp T. & Chowdhary B. (2016) Chromosome Aberrations and Fertility Disorders in Domestic Animals. In: *Annual Review of Animal Biosciences* (ed. by L A). Annual Reviews, Palo Alto, CA.

Raudsepp T. & Chowdhary B.P. (2008) FISH for mapping single copy genes. *Methods in molecular biology* **422**, 31-49.

Raudsepp T., McCue M.E., Das P.J., Dobson L., Vishnoi M., Fritz K.L., Schaefer R., Rendahl A.K., Derr J.N., Love C.C., Varner D.D. & Chowdhary B.P. (2012) Genome-wide association study implicates testis-sperm specific *FKBP6* as a susceptibility locus for impaired acrosome reaction in stallions. *PLoS Genet* **8**, e1003139.

Reid S.W., Weatherston J.F. & Robinson B.N. (1987) X chromosome mosaicism and infertility in a mare. *Can Vet J* **28**, 533-4.

Ryder O.A. & Wedemeyer E.A. (1982) A cooperative breeding programme for the Mongolian wild horse Equus przewalskii in the United States. *Biological Conservation* **22**, 259-71.

Santos A.S. & Ferreira L.M.M. (2012) The Portuguese Garrano breed: an efficient and sustainable production system. In: *Forages and grazing in horse nutrition* (eds. by Saastamoinen M, Fradinho M, Santos A & Miraglia N), pp. 481-4. Wageningen Academic Publishers.

Schrimpf R., Dierks C., Martinsson G., Sieme H. & Distl O. (2014) Genome-wide association study identifies phospholipase C zeta 1 (PLCz1) as a stallion fertility locus in Hanoverian warmblood horses. *PloS one* **9**, e109675.

Schrimpf R., Metzger J., Martinsson G., Sieme H. & Distl O. (2015) Implication of *FKBP6* for male fertility in horses. *Reprod Domest Anim* **50**, 195-9.

Sobczuk D. & Komosa M. (2012) Morphological Differentiation of Polish Arabian Horses - Multivariate Analysis. *Bull Vet Inst Pulawy* **56**, 623-29.

Solé M., Santos R., Gómez M.D., Galisteo A.M. & Valera M. (2013) Evaluation of conformation against traits associated with dressage ability in unridden Iberian horses at the trot. *Res Vet Sci* **95**, 660-6.

Templado C., Donate A., Giraldo J., Bosch M. & Estop A. (2011) Advanced age increases chromosome structural abnormalities in human spermatozoa. *Eur J Hum Genet* **19**, 145-51.

Vicente A.A., Carolino N., Ralão-Duarte J. & Gama L.T. (2014a) Selection for morphology, gaits and functional traits in Lusitano horses: I. Genetic parameter estimates. *Livest Sci* **164**, 1-12.

Vicente A.A., Carolino N., Ralão-Duarte J. & Gama L.T. (2014b) Selection for morphology, gaits and functional traits in Lusitano horses: II. Fixed effects, genetic trends and selection in retrospect. *Livest Sci* **164**, 13-25.

Wallner B. (2004) Isolation of Y Chromosome-specific Microsatellites in the Horse and Cross-species Amplification in the Genus Equus. *J Hered* **95**, 158-64.

Wang J. (2016) Pedigrees or markers: Which are better in estimating relatedness and inbreeding coefficient? *Theor Popul Biol* **107**, 4-13.

Wang W., Wang S., Hou C., Xing Y., Cao J., Wu K., Liu C., Zhang D., Zhang L., Zhang Y. & Zhou H. (2014) Genome-Wide Detection of Copy Number Variations among Diverse Horse Breeds by Array CGH. *PloS one* **9**, e86860.

Wieczorek M., Switonski M. & Yang F. (2001) A low-level X chromosome mosaicism in mares, detected by chromosome painting. *J Appl Genet* **42**, 205-9.

Zechner P., Zohman F., Sölkner J., Bodo I., Habe F., Marti E. & Brem G. (2001) Morphological description of the Lipizzan horse population. *Livest Prod Sci* **69**, 163-77.

Zhang Q., Calus M.P., Guldbrandtsen B., Lund M.S. & Sahana G. (2015) Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. *BMC Genet* **16**, 88.

Glossary

Assisted reproductive technologies (ART) - Reproductive technologies routinely used to genetically manage and study wildlife species, as well as in endangered species conservation. These methods include: artificial insemination (AI); non-invasive hormonemonitoring; in vitro oocyte maturation and culture; in vitro fertilisation (IVF); embryo transfer; germplasm banking; cloning and stem cell-based technologies.

Average relatedness coefficient (AR) - Defined as the probability that an allele randomly chosen from the whole population in the pedigree belongs to a given animal. Numerically, AR corresponds to twice the probability of two random alleles (one from the animal and other from the complete population, including itself) are identical by descent and it can be interpreted as the representation of a given animal in the whole pedigree regardless of the knowledge of its own pedigree. It is computed as the average of the coefficients integrating the row from the individual in the numerator relationship matrix and it takes into account, simultaneously, the inbreeding and the coancestry coefficients.

Copy number variations (CNVs) - Genomic structural variations that involve DNA segments ranging from 1 kb to 5 Mb, comprising insertions and deletions, as well as inversions and translocations). CNVs represent an important source of genetic variation among individuals and cover more nucleotide sequences per genome compared to SNPs markers.

Effective number of ancestors (f_a) - The minimum number of ancestors (founders or not) that are necessary to explain the population's overall genetic diversity (Boichard et al., 1997). It is computed as:

$$f_a = \frac{1}{\sum_{j=1}^{a} q_j^{\,2}}$$

q_j being the marginal contribution of an ancestor j (genetic contribution of an ancestor that isn't explained by any other previously chosen). This parameter takes into account genetic diversity losses derived from the unbalanced use of reproductive animals through generations producing bottlenecks, being considered a useful parameter and an important complement to the information supplied by the f_e. Only the animals with both parents known are considered for the calculations.

Effective number of founders (f_e) - Number of founders that would be necessary to produce the same amount of genetic diversity as in the population under study if they had identical contributions to the descendant population. It is computed as:

$$f_e = \frac{1}{\sum_{k=1}^{f} q_k^{\,2}}$$

where q_k is the probability of gene origin of the k ancestor (computed as the average relatedness coefficient - AR) and f is the total number of founders. It is equivalent to the parameter f_e obtained following the methodology proposed by James (1972) or Lacy (1989) (*founder equivalents*) if the whole pedigree is included in the calculations.

Effective population size (Ne) **-** The size of a randomly mating population of constant size with equal sex ratio and a Poisson distribution of family sizes that would (a) result in the same mean rate of inbreeding as that observed in the population, or (b) would result in the same rate of random change in gene frequencies (genetic drift) as observed in the population. These two definitions are identical only if the population is demographically stable (because the rate of inbreeding depends on the distribution of alleles in the parental generation, whereas the rate of gene frequency drift is measured in the current generation).

Expected heterozygosity (H_E) **-** Is calculated using an unbiased formula from allele frequencies assuming Hardy-Weinberg equilibrium (equation 8.4, Nei 1987).

Fitness **-** Capability of a genotype to survive and reproduce. Selection, within environments, acts against the phenotype associated with the genotype.

Founder **-** An individual at the top of the pedigree, assumed to be unrelated to all other founders. An individual is not yet a founder of the captive-born population until it has living descendants in the population.

Founder genome equivalents **-** The number of equally represented founders with no loss of alleles (retention = 1) that would produce the same gene diversity as that observed in the living, descendant population. Equivalently, the number of animals from the source population that contain the same gene diversity as does the descendant population. The gene diversity of a population is 1 - 1 / (2 * fge). In Goals, FGE is the number of founder genomes that will be incorporated into the population for each new founder added. A fge of 0.4 means that each new founder only contributed 40% of a founder genome to the population.

Generation length **-** The time elapsing from reproduction in one generation to the time the next generation reproduces. Also, the average age at which a female (or male) produces offspring. It is not the age of first reproduction. Males and females often have different generation lengths.

Hardy-Weinberg Equilibrium **-** State of a population in which the gene and genotypic frequencies remain constant from one generation to the next.

Inbreeding **-** Mating of related animals resulting in a non-zero probability that alleles at a particular locus are identical by descent. The inbreeding coefficient of an offspring is equal to the kinship between its parents.

Inbreeding Coefficient **-** Probability that the two alleles at a genetic locus are identical by descent from a common ancestor to both parents. The mean inbreeding coefficient of a population will be the proportional decrease in observed heterozygosity relative to the expected heterozygosity of the founder population.

Increase in inbreeding per generation **-** We used average F values' evolution per year of birth to evaluate this parameter, according to the formula in Gutiérrez et al. (2003) (F_n - F_{n-1} = l x b, where l = the average generation interval and b = the regression coefficient of mean inbreeding coefficient per birth year).

Individual heterozygosity **-** Number of *loci* at which an individual was heterozygous, divided by the total number of loci at which an individual was scored.

Individual increase in inbreeding ΔF_i ***(Gutiérrez et al., 2008)*** **-** Computed for each individual in the pedigree, following the modification proposed by Gutiérrez et al. (2009) to account for the exclusion of self-fertilization, as $\Delta F_i = 1 - \sqrt[t/1]{1 - F_i}$, where F_i is the inbreeding coefficient for each individual i and t the equivalent complete generations computed on the pedigree of this individual.

Kinship **-** Probability that alleles randomly selected from homologous loci in two individuals are identical by descent from a common ancestor. A measure of the genetic identity of two individuals.

Mean d^2 - Calculated as the squared distance in repeat units between the 2 alleles an individual had at a microsatellite locus, averaged over all loci at which an individual was scored, according to equation:

$$mean\ d^2 = \sum_{i=1}^{n} \frac{(i_a - i_b)^2}{n}$$

where i_a and i_b are the lengths (in base pair) of alleles a and b at locus i and n is the total number of loci at which an individual was score.

Mean kinship (mk) - The mean kinship coefficient between an animal and all animals (including itself) in the living, captive-born population. The mean kinship of a population is equal to the proportional loss of gene diversity of the descendant (captive-born) population relative to the founders and is also the mean inbreeding coefficient of progeny produced by random mating. Mean kinship is also the reciprocal of two times the founder genome equivalents.

Microarrays - Sets of miniaturized reaction areas that may also be used to test the binding of DNA fragments. Several different technologies are currently used to perform microarray experiments. DNA microarrays offer the opportunity to analyse the expression of many thousands of genes in a single experiment. They work by providing a fixed single strand of DNA to which labelled DNA fragments can bind. The DNA fragments are physically attached to an inert support (called a chip).

Microsatellies - See short tandem repeats (STRs)

Ne/N - Ratio of the Effective Population Size to the Census Size. Approximately, the proportion of animals that are currently breeders.

Next generation sequencing (NGS) - New DNA sequencing technologies that produce millions of short reads (from 25–500 bp) in a short time (1–5 days).

Observed Heterozygosity (H_O) - Proportion of individuals in a population (or proportion of genetic loci within an individual) that are heterozygous (having two different alleles at a genetic locus).

Pedigree **-** Biological relationship among members of a family.

Polymorphic information content (PIC) - Average polymorphic information content, taken across all loci. This is a measure of informativeness related to expected heterozygosity and likewise is calculated from allele frequencies (Botstein *et al.* 1980; Hearne *et al.* 1992). It is commonly used in linkage mapping.

Probability of exclusion (PE) - The average probability that the set of loci will not exclude a pair of unrelated candidate parents from parentage of an arbitrary offspring.

Probe **-** DNA fragments on an array to which test DNA will bind.

Sex reversal **-** A condition in which the karyotype appears normal but the genetic/chromosomal sex does not match with the gonadal and/or external phenotypic sex.

Short tandem repeats (STRs) **-** Highly polymorphic repeated DNA sequences, short (2–6 bp), comprised of a variable number of tandem repeats that occur in a seemingly random fashion distributed throughout the genome of all higher organisms.

Structural rearrangements **-** Structural aberrations change the constitution of one or more chromosomes, are typically caused by mistakes in meiotic recombination and DNA break repair, and can be classified as genetically balanced and unbalanced changes. There rearrangements can be: inversions, translocations, deletions, duplications.

Runs of homozygosity **-** Long stretches of homozygous genome that most likely arise when the individual is the offspring of related individuals.

Single Nucleotide Polymorphisms (SNPs) **-** Single-base-pair variations scattered within the genetic code of the individuals within a population