Population genetics and disease ecology of European wild boar

Daniel Johannes Goedbloed

Thesis committee

Promotors
Prof. Dr H.H.T. Prins
Professor of Resource Ecology
Wageningen University

Prof. Dr R.C. Ydenberg
Professor of Fauna Management and Conservation
Wageningen University
Professor of Behavioural Ecology
Simon Fraser University, Burnaby, Canada

Co-promotor
Dr W.F. van Hooft
Assistant professor, Resource Ecology Group
Wageningen University

Other members
Prof. Dr E.W. Brascamp, Wageningen University
Prof. Dr H. Leirs, University of Antwerp, Belgium
Dr G.A. de Groot, Alterra, Wageningen UR
Prof. Dr B. Zwaan, Wageningen University

This research was conducted under the auspices of the C.T. de Wit Graduate School for Production Ecology and Resource Conservation

Population genetics and disease ecology of European wild boar

Daniel Johannes Goedbloed

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 13 November 2013
at 4 p.m. in the Aula.

Daniel Johannes Goedbloed
Population genetics and disease ecology of European wild boar
123 pages.

PhD thesis, Wageningen University, Wageningen, NL (2013)
With references, with summaries in English and Dutch

ISBN 978-94-6173-723-6

Table of contents

Chapter 1

General introduction

Infectious diseases are the number two cause of human death, responsible for 25% of worldwide mortality [1]. Essentially all human infectious diseases ultimately originate from wildlife populations, and adapted to humans directly or via livestock populations [2]. A number of human infectious diseases have emerged from wildlife populations in recent decades [3], including Avian influenza, Lyme disease and West Nile virus. The link between human, livestock and wildlife diseases has led to formulation of the 'One world, one health' concept, which integrates wildlife conservation, public and animal health [4]. However, infectious disease dynamics usually differ strongly between humans, livestock and wildlife due to differences in, e.g., host density, contact networks, environmental stress levels and application of medicine [3]. The epidemiology of human and livestock infectious diseases is relatively well-studied [5, 6], but the driving forces behind wildlife disease dynamics are largely unknown due to challenges in sampling, laboratory diagnostics and a high diversity of ecological interactions [7, 8]. In order to understand the effects of human disturbance on wildlife epidemiology, and in order to evaluate the risks of wildlife infectious diseases for biodiversity conservation, livestock industry and public health, a good understanding of the driving forces of wildlife disease dynamics is required. In this thesis I contribute to wildlife epidemiology and disease ecology by exploring the effects of various factors on the prevalence of two respiratory diseases in free-living European wild boar.

Some definitions

Allele: one of a number of alternative forms of a gene or genetic locus.

Disease dynamics: the change over time of disease prevalence.

Driving force (driver): a factor that propels and/or controls a process.

Epidemiology: the sum of factors determining disease prevalence (or the scientific study thereof).
Genetic load: the decrease in fitness of the average individual in a population due to presence of deleterious alleles in the gene pool.
Pathogen: a parasitic (micro)organism that causes disease (i.e., damages its host).
Prevalence: the frequency in a population (often expressed as a percentage).
Zoonosis: an infectious disease that can be transmitted from animals to humans.

Wildlife disease ecology

Wildlife species host a wide range of pathogens and are considered to be an important factor in the maintenance, emergence and spread of infectious diseases [3]. Some of these diseases are shared with domestic livestock or humans (zoonosis) and lead to economic, biodiversity and public health concerns [9, 10]. A wide range of possible factors may drive wildlife disease prevalence [11], which are central to wildlife disease ecology: the study of host-pathogen interactions in the context of their environment and evolution.

Ecologists previously assumed that pathogens have little impact on wildlife populations [12]. However, the last decades it has become increasingly apparent that pathogens are not only common and integral to ecosystems, but that pathogens can influence the abundance and extinction risk of populations and act as an important driving force for evolution [12, 13]. The relatively rapid mutation and adaptation rate of pathogenic microorganisms allows pathogens to 'emerge' in (or adapt to) previously unsuitable host species and places [14].

Wildlife infectious diseases show a high diversity of life history traits: from generalist to specialist, from vector-borne to air-borne to sexually transmitted, from respiratory to gastro-intestinal infection routes, from slow to fast reproduction rates and from slow to fast mutation rates. In this thesis I will focus on two host-specific directly transmitted (temporally air-borne) respiratory swine pathogens: porcine circovirus type 2 (PCV2) and the bacterial *Mycoplasma hyopneumoniae* (*Mhyo*).

Predictions from mathematical host-pathogen models

Most of our conceptual understanding of disease dynamics stems from mathematical host-pathogen models. Simplified host-pathogen models of disease dynamics occur in many forms. The classic compartmental host-pathogen models (see Figure 1.1) predict that the occurrence of disease is driven by the abundance (sometimes given as density) of susceptible hosts [15-17]. This is known as the Kermack-McKendrick threshold theorem [18]. These compartmental models assume uniform and universal contact between individuals at each time step (mass action). This is often not realistic, especially for group living host species.

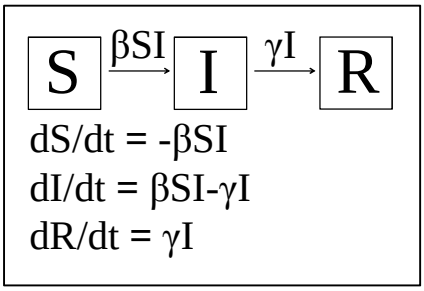


Figure 1.1 Example of a simple compartmental model. Clarification of symbols: S represents the proportion of susceptible individuals in a population, I is the proportion of infected individuals and R the proportion of recovered individuals. β is the transmission coefficient, γ is the recovery rate and t is time.

Network models of disease dynamics (particularly small world property networks, but also scale-free and random networks) have been employed to introduce heterogeneity in the contact of individuals [19, 20]. Abundance thresholds have in some cases been demonstrated using network models for disease dynamics, but outcomes depend heavily on the properties of the network [21].

Observations of high levels of pathogen aggregation, where a small fraction of host individuals harbours the majority of pathogens, are a central issue in epidemiology [15, 22]. An essential component of disease modelling in this regard is the 'transmission coefficient' (see Figure 1.1), which represents the relative capacity of a pathogen to overcome the host's innate immune defences [22]. Both pathogens and hosts are genetically heterogenic and individuals will

differ in their virulence and immune capacity respectively. The heterogeneity of hosts and pathogens in terms of their influence on the transmission coefficient is usually not properly accounted for in host-pathogen modelling [23], but it has been shown to greatly affect model outcomes [24]. The heterogeneity of hosts and pathogens and their influence on the transmission coefficient may explain observations of pathogen distribution and aggregation.

Wild boar as a model species for wildlife disease research

Wild boar (*Sus scrofa*) are the ancestors of domestic pigs [25, 26]. The two are closely related and readily interbreed as well as share their diseases. Wild boar are known to host many pathogens, including diseases with a potential for negative economic consequences such as Classical Swine Fever and Aujeszky's Disease as well as zoonotic diseases that can infect humans, e.g., Influenza viruses [27-30]. Commercial interests from the pig breeding industry have led to development of a variety of diagnostic tools for serology and disease testing, as well as to development of advanced molecular tools for genomic research and breeding purposes. Little is known about the relative immune capacity of wild boar versus domestic pigs. The few case studies that report on this topic suggest little difference in clinical symptoms [31]. Disease prevalence is usually higher in domestic herds than in free-living populations, but this is usually attributed to higher animal densities and the related increase in disease transmission efficiency [32, 33]. Immune capacity is generally a difficult subject due to physiological interactions and trade-offs, and because the role of genetic factors in wildlife disease dynamics is largely unknown [34].

The wild boar is a moderate to large sized pig (adult weight 80-150kg) with black fur. This polygynous species has a large distribution spanning most of Eurasia, where it occurs in variable but sometimes high densities (up to 60 wild boar km^{-2}) [35]. Adult males are solitary and maintain a large home range to maximise access to females. Females and sub-adults form sounders consisting of 6-20 animals or more, and tend to have a flexible home-range size and location according to season, food abundance, predation or hunting pressure [36]. Wild boar are good dispersers, with recorded life-time dispersal distances up to 250 km, and show male-biased dispersal [35, 37]. Wild boar are opportunistic omnivores [38], and have benefitted the last decades from changes in

agricultural crops, supplementary feeding and frequent good mast years due to a more temperate Western Palearctic climate [39]. Conditions in Europe have been favourable for reproduction in the last decades, allowing wild boar to breed all year round in large parts of the continent. Wild boar have a relatively short generation time with litter sizes averaging 4-8 piglets and female sexual maturity at 8-10 months of age, provided that a body mass threshold of 30kg is reached [40]. European wild boar have recently increased in number and range, reaching previously unrecorded levels of abundance [41].

European wild boar population genetics and introgression from domestic pigs

Wild boar is an intriguing study organism for many reasons, one of which is the availability of advanced genomic tools such as a full porcine genome sequence [42] and a high density Single Nucleotide Polymorphism (SNP) assay [43]. SNPs are a highly valued genetic marker, because of their high frequency in the genome and their compliance with mutation models allowing powerful statistical analysis [44, 45]. This allows relatively detailed investigation of population genetic processes, including gene flow and evolutionary adaptation. Phylogenetic relationships, gene flow and genetic adaptation are assumed to be highly relevant in the context of disease ecology research [46], because these processes reflect coevolution, host movement patterns and differences in immune capacity with regard to pathogen infection.

Large scale European wild boar gene flow and genetic population structure are mainly determined by postglacial recolonisation patterns from Mediterranean refugia after the Pleistocene ice age [47]. Regional wild boar population structure has been studied only occasionally, but may be determined by landscape barriers and human translocations [48, 49]. Genetic exchange between domestic pigs and wild boar has occurred throughout the history of domestication and pig breeding in both directions. The introgression of genetic elements from wild boar into the domestic pig genome is well studied [50]. In contrast, the extent of introgression from domestic pigs into wild boar was largely unknown at the start of this study [51]. Genetic signs of introgression had been reported in up to 2% of wild boar in Eurasia based on mitochondrial DNA [25, 26] and in 5–10% of wild boar in Europe based on a combination of

mitochondrial DNA and microsatellites [47]. Domestic pigs are subject to artificial selection and receive veterinary care as well as housing and regular feed. This is in stark contrast to wild boar, which are subject to natural selection in the wild without external support. These differences may have important consequences for disease dynamics and differences in host immunogenetic adaptation.

Thesis outline

The main aim of this thesis was to identify factors that significantly influence infectious disease prevalence in European wild boar populations. Possible driving forces of wildlife disease prevalence are not just limited to host abundance or density, but include demographic factors (e.g., host age, sex and population substructure), environmental conditions (e.g., food availability, predation pressure, ambient temperature and humidity) and individual genetic composition (e.g., inbreeding depression, outbreeding depression and inheritance of specific deleterious or beneficial alleles).

In order to test the significance of the influence of a number of these factors on disease (PCV2 and *Mhyo*) prevalence, I collected surplus wild boar blood samples from disease monitoring institutes and routine population management. Sex, age class and the location of these samples were recorded in the field. The blood samples were genotyped using Single Nucleotide Polymorphism (SNP) panels, and antibody titres against PCV2 and *Mhyo* were determined using ELISA assays. The samples were collected from the Netherlands, Luxembourg and parts of Western Germany (North Rhine-Westphalia and Rhineland-Palatinate).

In chapter 2 I address the issue of genetic introgression from domestic pigs into the wild boar population. This genetic introgression can be crucial because it may affect population substructure as well as inbreeding and outbreeding levels and because it may introduce specific domestic immune-related alleles.

In chapter 3 I describe wild boar genetic population structure in the study area. This work is required to determine if disease dynamics differ between biologically meaningful populations, or if gene flow correlates with disease prevalence patterns.

Chapter 4 deals with the disease ecology of PCV2 in wild boar. In this chapter I evaluate the influence of a number of demographic, environmental and genetic factors on PCV2 prevalence in the study area.
In chapter 5 I extend this work with an assessment of the disease ecology of *Mhyo* in wild boar. Some similarities and differences between these two diseases are discussed.
Finally, in chapter 6 I bring the previous chapters together in a discussion of the bearings of these results on existing theory and concepts in wildlife disease ecology and epidemiology.

Chapter 2

Genome-wide SNP analysis reveals recent genetic introgression from domestic pigs into Northwest European wild boar populations

Daniel J Goedbloed, Hendrik-Jan Megens, Pim van Hooft, Juanma M Herrero-Medrano, Walburga Lutz, Panoraia Alexandri, Richard PMA Crooijmans, Martien AM Groenen, Sip E van Wieren, Ron C Ydenberg, Herbert HT Prins

Abstract

Present-day genetic introgression from domestic pigs into European wild boar has been suggested in various studies. However, no hybrids have been identified beyond doubt mainly because available methods were unable to quantify the extent of introgression and rule out natural processes. Genetic introgression from domestic pigs may have far-reaching ecological consequences by altering traits like the reproduction rate or immunology of wild boar. In this study we demonstrate a novel approach to investigate genetic introgression in a Northwest European wild boar dataset using a genome-wide Single Nucleotide Polymorphism (SNP) assay developed for domestic pigs. We quantified the extent of introgression using allele frequency spectrum analysis, in silico hybridization simulations and genome distribution patterns of introgressed SNPs. Levels of recent introgression in the study area were expected to be low, as pig farming practices are prevailingly intensive and indoors. However, evidence was found for geographically widespread presence of domestic pig SNPs in 10% of analysed wild boar. This was supported by the identification of two different pig mitochondrial DNA haplotypes in three of the identified hybrid wild boar, suggesting that introgression had occurred from multiple sources (pig breeds). In silico hybridization simulations showed that the level of introgression in the identified hybrid wild boar is equivalent to first generation hybrids until fifth generation backcrosses with wild boar. The distribution pattern of

introgressed SNPs supported these assignments in four out of nine hybrids. The other five hybrids are considered advanced generation hybrids, resulting from interbreeding among hybrid individuals. Three out of nine hybrids were genetically associated with a different wild boar population than the one in which they were sampled. This discrepancy suggests that genetic introgression has occurred through the escape or release of an already hybridized farmed wild boar stock. We conclude that genetic introgression from domestic pigs into Northwest European wild boar populations is more recent and more common than expected, and that genome-wide SNP analysis is a promising tool to quantify recent hybridization in free-living populations.

Molecular Ecology 22(3): 856-866

Introduction

European and Asian pigs were independently domesticated from wild boar (Sus scrofa) [25, 26]. Even though the first domestication of European pigs is estimated to have occurred 9000 years ago [25, 26], European wild boar are still fully capable of hybridizing with domestic pigs. The process of domestication and later introgression of genetic elements from wild boar into the domestic pig genome is well studied [25, 26, 50]. In contrast, the extent of introgression from domestic pigs into wild boar is largely unknown [51]. Frequent genetic introgression from domestic pigs may lead to either hybrid vigour or to maladaptation to the natural environment [52]. In addition, regular intimate contact between pigs and wild boar may increase the risk of disease transfer and outbreaks. The extent of genetic introgression is thus a relevant parameter for wild boar conservation management and disease risk management. Genetic signs of introgression have been reported in up to 2% of wild boar in Eurasia based on mitochondrial DNA [25, 26], and in 5-10% of wild boar in Europe based on a combination of mitochondrial DNA and microsatellites [47]. The latter authors consider their estimate to be slightly inflated and report introgression in general to be lower than 5% [51]. Another study using mtDNA D-loop sequences reports only 1.6% Asian haplotypes in wild boar versus 29% in the European domestic population [53].

European wild boars have survived Pleistocene ice-ages in Mediterranean refugia [47]. Wild boars in Western Europe are considered to originate from the Iberian refugium and have a chromosome number of 2n=36. They differ in their karyotype from domestic pigs and from Balkan refugium wild boar in Eastern Europe, both with chromosome number 2n=38 [54]. Hybridization can occur, resulting in individuals with chromosome number 2n=37 [51]. Admixture between different wild boar populations may locally introduce new alleles. Single Nucleotide Polymorphism (SNP) genetic markers are found throughout any genome and represent the largest source of genetic variation [44]. Models for the mutation rate of SNPs are well established and high throughput genotyping methods are becoming increasingly efficient. These characteristics make SNPs a popular choice of marker for population genetic research [45]. Few studies have used genome-wide SNP sets in non-model organisms (e.g., [55]), as this technology is still relatively new. However, in

some cases a SNP set developed for a model species can be used effectively to study closely related non-model species [56-58].

In this study we aimed to identify the occurrence, time-frame and possible sources of genetic introgression from domestic pig into Northwest (NW) European wild boar. We used a high-density genome-wide SNP assay developed for domestic pig, the Illumina porcine SNP60 genotyping beadchip [43], for the genetic analysis of 88 wild boar from the Netherlands, Luxembourg and Western parts of Germany. This assay provided 26505 SNPs that segregated in the wild boar dataset and which were distributed across all autosomes. This amounted to a substantially higher genome coverage than commonly seen in molecular ecology studies [59]. We identified genetic introgression based on an increased abundance of rare alleles. Results from a mitochondrial (mt) DNA haplotype study were used to independently verify cases of introgression. The level of introgression from domestic pig was identified using a hybridization simulation study and the genomic distribution patterns of introgressed SNPs.

Methods

In 2008 we collected 88 wild boar blood samples from the Netherlands, Luxembourg and Western parts of Germany. Sample collection was opportunistic and without bias towards age, sex or sampling location (supplementary information Table S2.1). DNA isolation was performed following the Gentra PureGene Blood kit protocol. Samples were genotyped using the Illumina porcine SNP60 genotyping beadchip Infinium SNP assay [43] and initially analysed for all 45720 autosomal SNPs. The total genotyping rate was 0.98. During exploration using PLINK v1.06 [60], we found that SNPs with a low minor allele frequency ($0.005<MAF<0.030$) were highly abundant in the wild boar dataset (Figure 2.1a). This allele frequency spectrum was compared to that of a domestic pig dataset consisting of 20 individuals per breed for six breeds; British Saddleback, Duroc, Landrace, Large White, Pietrain and Tamworth (Figure 2.1b). These breeds were selected on the basis of occurrence in NW Europe and the availability of sufficient SNP data. MAF was in all cases calculated separately for the wild boar and domestic pig datasets. After allele frequency spectrum assessment, we excluded non-polymorphic sites and potential genotyping errors by applying a rigorous MAF threshold of 0.05 using

PLINK, as a standard procedure. This procedure therefore excluded the highly abundant rare alleles for further analysis, making sure that population genetic inferences were not influenced by potential artefacts. The procedure left 26505 segregating autosomal SNPs for population genetic analysis in the wild boar dataset. The 7083 highly abundant rare SNPs in the wild boar dataset (0.005<MAF<0.030) were analysed separately, and revealed 5038 putative introgressed SNPs, which were private to just nine out of 88 wild boar. These putative introgressed SNPs were also analysed for their allelic state in the domestic pig dataset and a sample of wild boar from the Balkans (northern Greece and Bulgaria, *n*=20) to assess the origin of the putative introgressed SNPs. To identify genetic clustering in the wild boar dataset, we performed Principal Component Analysis (PCA) using the eigenvector method as implemented in Eigensoft 3.0 [61, 62]. In addition, we performed a population assignment analysis using STRUCTURE 2.3.1 [63] based on 10 runs per number of clusters (K) for K=1-10 at 1,000,000 iterations and a burn in of 800,000. Putative hybrids were excluded from these analyses to achieve convergence between runs. The most supported partitioning (K) was identified using the method of Evanno [64]. Putative hybrids were removed to achieve convergence between runs. Observed and expected heterozygosity were calculated in R 2.13.0 using the package Adegenet [65]. Individual observed heterozygosity (Table 2.1, Ho) was calculated as the number of heterozygous SNPs divided by the total number of SNPs. In addition, part of the D-loop region of the mitochondrial DNA (mtDNA) was amplified by polymerase chain reaction (PCR) using the primers described by Luetkemeier et al. [66] (L-strand 5´CTCCGCCATCAGCACCCAAAG3´ and H-strand 5´GCACCTTGTTTGG ATTRTCG3´) yielding a 772 bp fragment. The PCR amplicons were purified and sequenced for both strands on an ABI 3130® DNA sequencer (Applied Biosystems, USA). Genome Assembly Program (GAP4, [67]) was used to view and obtain the consensus sequence of D-loop region for each individual relative to pig mtDNA sequence GenBank ID AJ00218 as a reference. Sequences were subsequently aligned by Clustal X V.2 [68] and grouped into haplotypes using the program ALTER [69]. As not all samples yielded the complete fragment (722 bp), a 624 bp fragment common to most samples was finally used for the analysis. Phylogenetic relationships among the haplotypes were determined with

Mega 5.03 [70] using the Neighbour Joining (NJ) method based on Tamura-Nei model. We included three additional NW European pig breeds: Berkshire, Bunte Bentheimer and Gloucester Old Spot in the mtDNA haplotype analysis (supporting information Table S2.2), as well as three sequences (accession numbers: DQ379224, DQ379100 and DQ379099) from Fang & Andersson [71]. Novel sequences were submitted to Genbank (supporting information Table S2.3). Hybridization simulations between domestic pigs and wild boar were performed in Excel 2010 using only monomorphic and rare SNPs with MAF<0.030 in the wild boar dataset. We used genetic data from the Veluwe population in the central Netherlands (Figure 2.3, indicated by circles, *n*=23) as the wild boar parent population. Analysis of shared polymorphisms (Table 2.3) and mtDNA haplotypes (Table 2.2) led us to specifically use the Large White (LW) and the British Saddleback (BS) pig breed (*n*=20 per breed) as parent pig populations for the hybridization simulations. LW shared most putative introgressed SNPs (80%) with the identified hybrid wild boar (Table 2.3) and harboured the observed pig haplotype HP8 (Table 2.2). BS shared 72% of putative introgressed SNPs with the identified hybrid wild boar and harboured the observed pig haplotype HP110. LW displayed 13879 SNPs with a non-wild boar allele and BS displayed 11989. The first generation hybridization (F1) was followed by seven generations of backcrossing with the parent wild boar population. We assumed Mendelian inheritance, meaning that the probability of inheritance for a typical pig allele (absent in non-hybrid wild boar) is 0.5 and 1 respectively for a heterozygous and homozygous SNP in the pig parent. Inheritance of a pig allele leads by definition to a heterozygous SNP in the next generation of hybrids. Each introgressed pig allele theoretically has a 50% probability to be inherited at each subsequent generation of backcrossing with the parent wild boar population, resulting in a halving of the total number of rare SNPs each generation. The standard deviation of the number of rare SNPs per individual for each generation was estimated on basis of 200 simulated genotypes per generation. Genomic positions of putative introgressed SNPs were analysed based on build 9 of the pig genome published by the International Swine Genome Sequencing Consortium in release 66 of the Ensembl database as Sscrofa9 (http://www.ensembl.org/Sus_scrofa/Info).

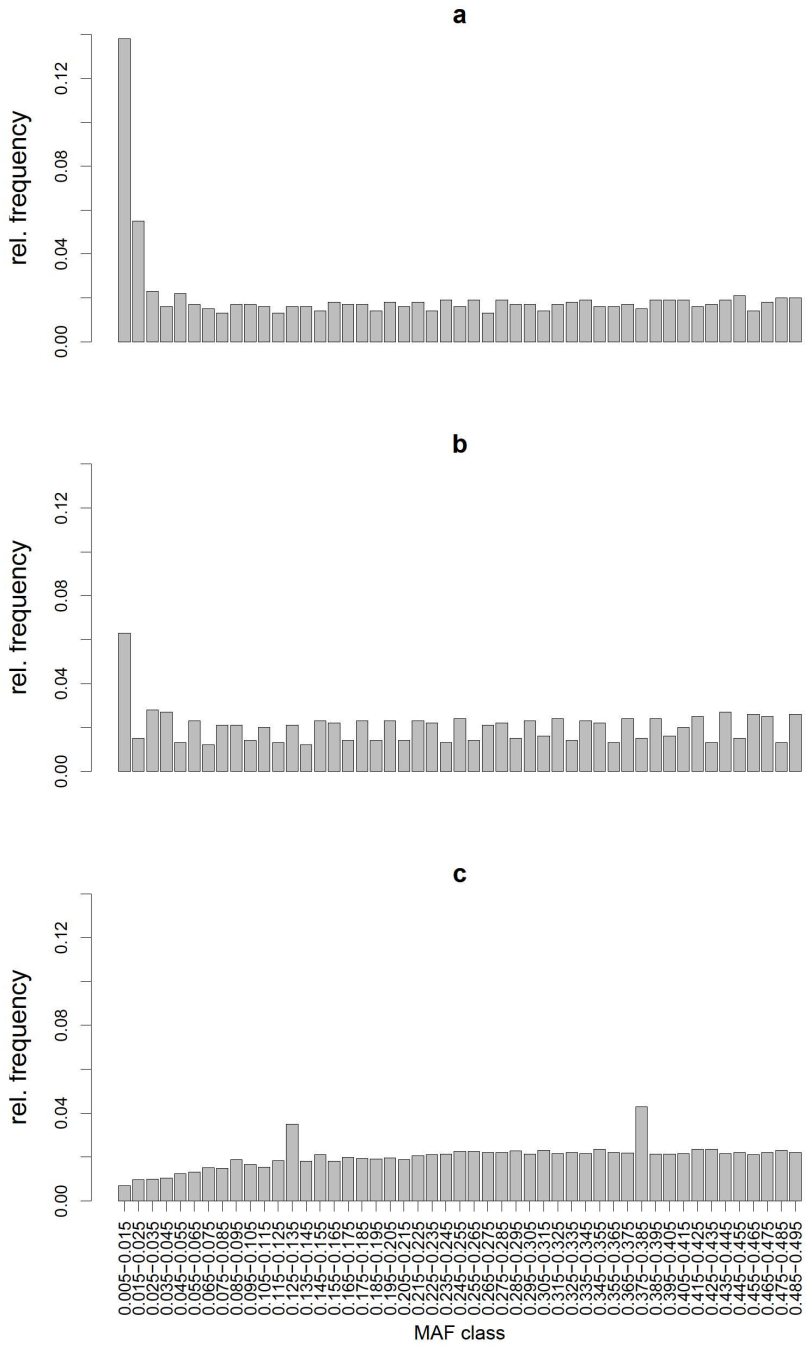


Figure 2.1 Minor Allele Frequency (MAF) distribution in a) the wild boar dataset, b) the wild boar dataset without 9 putative hybrids and c) the domestic

pig dataset. The x-axis indicates the MAF class. The y-axis indicates the frequency of each MAF class relative to the total number of SNPs in the dataset.

Results

The wild boar and domestic pig allele frequency spectra (Figure 2.1a and 2.1c respectively) differ dramatically at the lower end of the spectrum. In both cases we expected a more or less uniform distribution of SNPs across the allele frequency range based on random genetic drift and random mating. However, in the wild boar data we observed a clear excess of rare SNPs (0.005<MAF<0.030, Figure 2.1a). A large proportion (69%, 5038 SNPs) of these rare SNPs were private to just nine wild boar (Figure 2.1a and 2.1b). These putative introgressed SNPs (all heterozygous in those wild boar) almost correspond to the surplus in this MAF range, which in a uniform distribution would be expected to hold approximately 2250 SNPs rather than the observed 7083 SNPs. The nine wild boar with putative introgressed SNPs displayed higher overall levels of observed heterozygosity (Ho, Table 2.1) compared to other wild boar (Table 2.2).

Table 2.1 The number of putative introgression SNPs, observed heterozygosity (H_o) based on 26505 SNPs with MAF>0.05 and mtDNA haplotype per individual hybrid wild boar. The numbering of individuals corresponds to Figure 2.2 and 2.3.

Individual	Rare SNPs	H_o	MtDNA haplotype
1	256	0.226	HP165
2	1192	0.328	HP110
3	1086	0.325	HP110
4	129	0.202	HP8
5	580	0.207	HP19
6	1137	0.241	HP164
7	2435	0.354	HP164
8	1207	0.305	HP19
9	648	0.260	HP164

PCA separated the wild boar dataset into four genetic clusters (Figure 2.2a), with the nine putative hybrid individuals scattered across three of these clusters

(inverted triangles). The inclusion of a sample of domestic pigs in the PCA provided extra resolution, and clearly positioned these nine putative hybrid wild boar separately from the wild boar clusters, trailing off in the direction of the domestic pig (Figure 2.2b). The geographic origin of six of them (Figure 2.3) corresponded to their association with a particular genetic cluster. However, three putative hybrid wild boar (2, 3 and 5) clustered genetically with the Veluwe population (Figure 2.2, circles), but were sampled geographically in the Meinweg population in the South of the Netherlands (Figure 2.3, diamonds).

Table 2.2 Observed heterozygosity (H_o), expected heterozygosity (H_e) and mtDNA haplotype counts of the wild boar clusters, the group of hybrid wild boar and the six domestic pig breeds.

Group	*n*	H_o*	H_e*	HP16	HP16	HP1	HP11	HP	HPothe
Veluwe	2	0.18	0.19	19	0	4	0	0	0
Meinweg	2	0.16	0.16	1	0	23	0	0	0
Kirchhelle	2	0.17	0.17	0	24	0	0	0	0
Germany	1	0.20	0.20	7	0	4	0	0	0
Hybrids	9	0.26	**	2	1	3	2	1	0
L. White	2	0.33	0.35	2	0	1	0	1	16
Landrace	2	0.32	0.35	2	0	2	0	1	15
Pietrain	2	0.35	0.35	6	0	0	0	0	14
Br.	2	0.33	0.33	1	0	0	11	0	8
Duroc	2	0.33	0.34	6	0	1	0	0	13
Tamworth	2	0.33	0.32	0	0	0	8	0	12

* Standard errors are 0.001 or smaller

** not calculated as the hybrids do not constitute a population

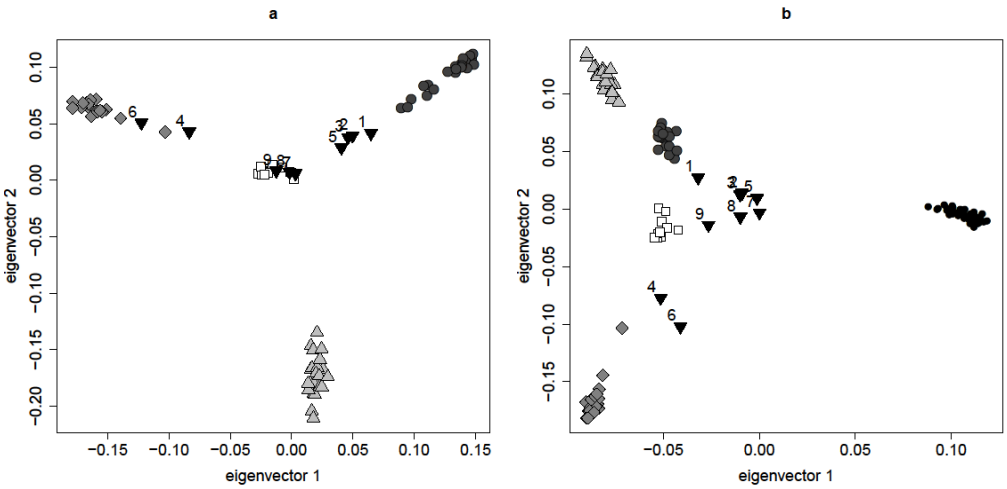


Figure 2.2 a) PCA plot based on 26505 SNPs with MAF>0.05. Four wild boar populations as inferred by STRUCTURE are indicated by different symbols. The nine individuals with putative introgressed SNPs are labelled and numbered explicitly (black inverted triangles). The first two eigenvectors explain 18% of variance in the dataset. b) PCA plot including a sample of all six domestic pig breeds considered in this study (small black dots) in the PCA analysis.

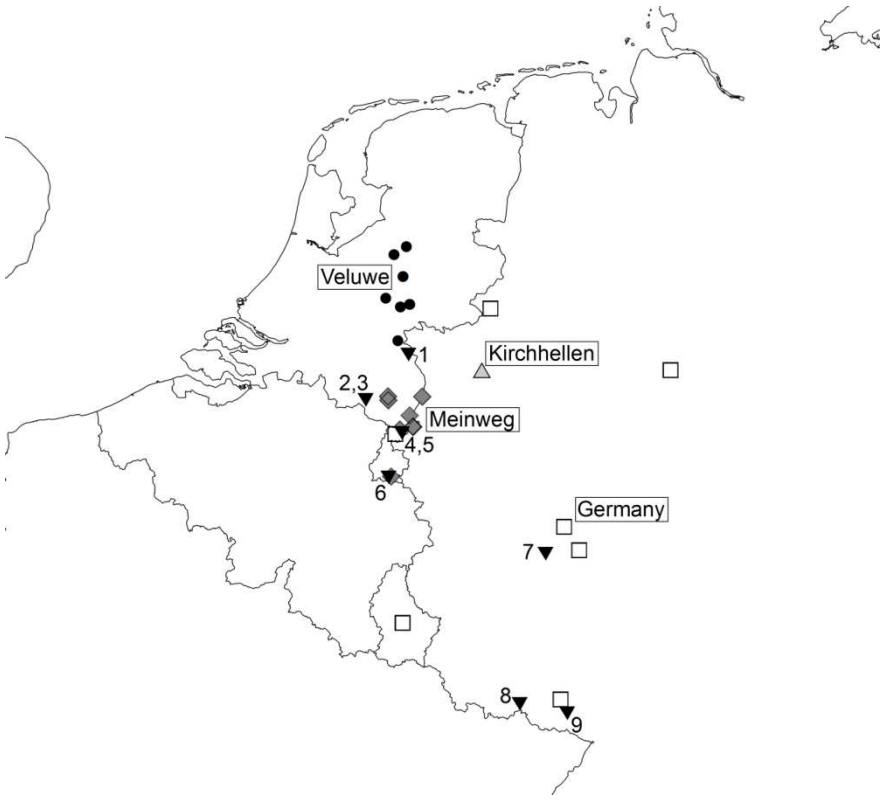


Figure 2.3 Geographic sample locations. Symbols and numbering correspond to the PCA analysis (Figure 2.2). Multiple samples may originate from one sampling location.

The most supported STRUCTURE partitioning of the data following the method of Evanno et al. (2005) was K=3 followed by K=4 (supporting information Figure S2.4). However, this method is known to favour only the first level of structure in a given dataset. In addition, the assignment of clusters for K=3 was not geographically coherent. German individuals were divided over the Meinweg and the Veluwe clusters with dubious assignment probabilities (supporting information Table S2.1). We suspect that this may be caused by a relatively low sample size of the German cluster (*n*=11 versus *n*=21, 23 and 24) as well as its wide geographic spread, resulting in high internal variation and lack of Hardy-Weinberg Equilibrium. The STRUCTURE partitioning K=4 matches fully to geographic and PCA distributions, and we therefore consider K=4 to be the most biologically meaningful structure of this dataset.

We investigated some possible sources of SNP introgression by quantifying the presence of the 5038 putative introgressed SNPs of the wild boar dataset in six domestic pig breeds (*n*=20 per breed) as well as a sample of wild boar from the Balkans (*n*=20, Table 2.3). The Large White domestic pig breed scored best, sharing approximately 80% of the putative introgressed SNPs. However, differences with other pig breeds were relatively small. Commercial pig farmers commonly use breed hybrids. Therefore we included some combinations of two breeds (*n*=40 per combination) in Table 2.3, which increased the percentage of putative introgressed SNPs explained to 86%. The percentage of shared putative introgressed SNPs between hybrid wild boar from NW Europe and wild boar from eastern Europe was only 20%.

Table 2.3 Shared SNPs between pig breeds (*n*=20 per breed) and the nine wild boar carrying putative introgressed SNPs. Six two-breed combinations (*n*=40) with a high amount of shared SNPs are also included, as well as a sample of wild boar from the Balkans (*n*=20). Percentages are calculated relative to the total amount of excessive rare SNPs in our wild boar dataset (5038).

Breed/combination	Shared SNPs	Percentage
Large White	4028	80
Landrace	3994	79
Pietrain	3868	77
British Saddleback	3647	72
Duroc	2876	57
Tamworth	1946	39
Large White * Landrace	4310	86
Large White * British Saddleback	4306	86
Large White * Pietrain	4267	85
Landrace * Pietrain	4267	85
Landrace * British Saddleback	4252	84
Pietrain * British Saddleback	4247	84
North Greece wild boar	1002	20

The wild boar in our dataset mostly displayed one of three common wild boar mtDNA haplotypes (HP164, HP165 and HP19), with three notable exceptions.

These exceptions are individuals with putative introgressed SNPs, which had a mtDNA haplotype not normally observed in wild boar (HP110 and HP8, Table 2.1). Haplotype HP110 is a rare haplotype among European pigs, because it has an Asian origin (supporting information Figure S2.5). The British heritage pig breeds and Pietrain are the only breeds in NW Europe that display this haplotype; Berkshire at a frequency of 5%, British Saddleback at 54%, Gloucester Old Spot at 40%, Tamworth at 43%, and Pietrain at 1.9% (*n*=593, supplementary information Table S2.2). Haplotype HP8 is typical for a number of mainland Europe pig breeds, including Landrace and Large White. Haplotypes HP110 and HP8 were not found in any of the 79 wild boar without putative introgressed SNPs.

The number of putative introgressed SNPs in each of the nine wild boar is indicated in Table 2.1. These numbers are decreasing (or increasing) more or less stepwise by a factor of two at each putatively assigned generation of backcrossing. This suggested a scenario of introgression followed by backcrossing with a wild boar gene pool theoretically halving the number of introgressed alleles at every generation of backcrossing.

To investigate the individual levels of introgression, we simulated hybrid genotypes using genotypes from the Veluwe wild boar population (Figure 2.3) and either of two domestic pig breeds: Large White (LW) and British Saddleback (BS). The number of putative introgressed alleles per individual wild boar observed in this study corresponded to expectations according to the hybridization simulations (Figure 2.4). Wild boar individual 7 was identified as equivalent to a first generation (F1) hybrid, wild boar individuals 2, 3, 6 and 8 were identified as equivalent to a second generation (F2) backcross to wild boar, individuals 9 and 5 were equivalent to a third generation (F3) backcross, individual 1 was equivalent to a fourth generation (F4) backcross and individual 4 was equivalent to a fifth generation backcross (Figure 2.4).

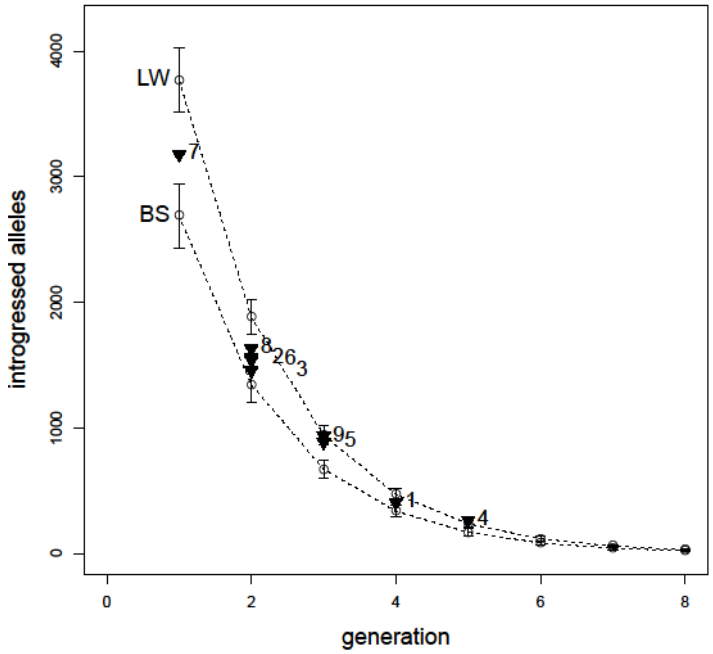


Figure 2.4 The open circles connected by dotted lines indicate the simulated mean number of introgressed pig alleles per individual (± s.d.) per generation of hybridization with Large White (LW) or British Saddleback (BS) pigs and subsequent backcrossing with wild boar. The number of putative introgressed alleles for each of the nine hybrids in our empirical dataset is indicated by inverted triangles (numbering corresponds to Figure 2.2 and 2.3).

The chromosomal positions of the introgressed SNPs are indicated for some of the identified hybrids in Figure 2.5. Individual 7 displays a wide array of introgressed alleles, resulting in a high prevalence of heterozygous SNPs across the entire genome. This pattern of genome wide heterozygosity corresponds to expectations for an F1 hybrid. Individuals 2, 5 and 1 represent subsequent generations of backcrossing with wild boar according to our hybridization simulation. The number of introgressed alleles is clearly diluted over the generations and the chromosomal positions show a clear clustering pattern that is distinct for each individual.

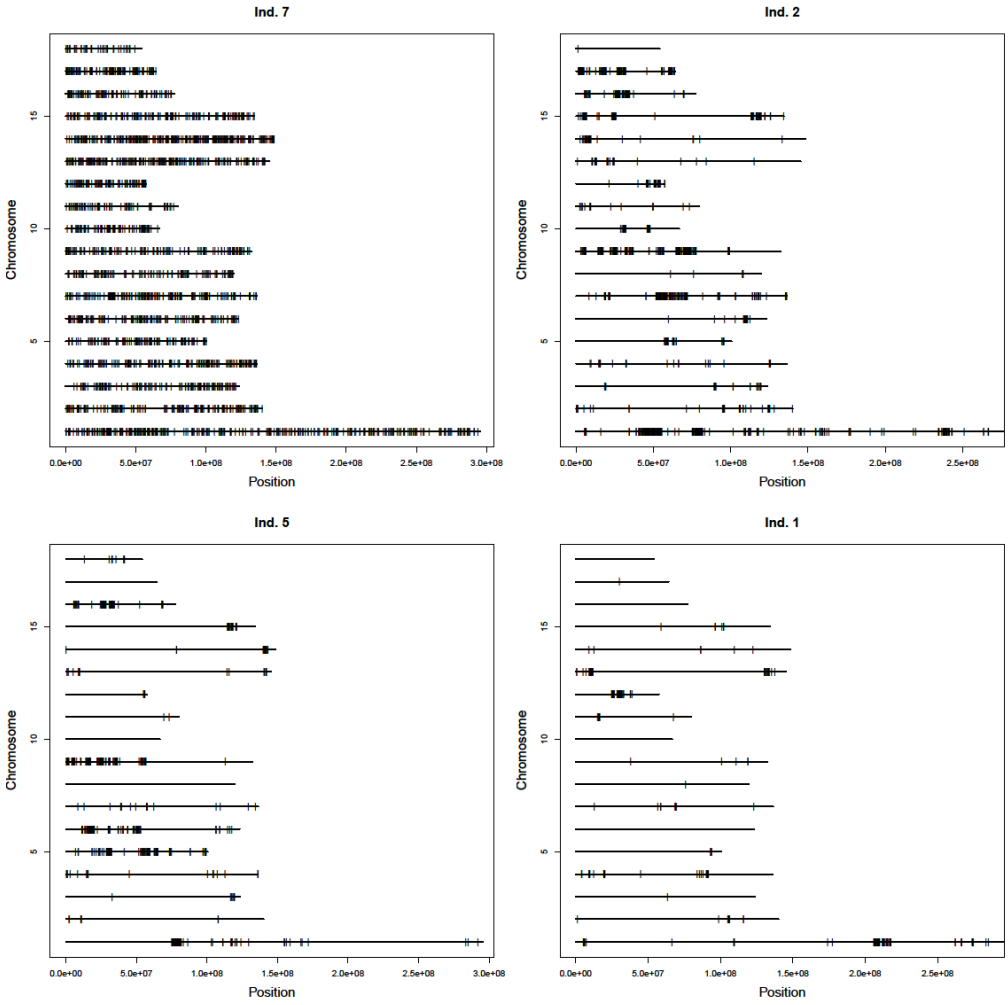


Figure 2.5 Chromosomal positions of introgressed SNPs. Individual 7 was assigned as an F1 hybrid, individual 2 as an F2 backcross with wild boar, individual 5 as an F3 backcross and individual 1 as an F4 backcross. A complete overview for all identified hybrids is given in supporting information Figure S2.6.

Discussion

Rare SNPs indicate genetic introgression from domestic pig in wild boar populations

The data presented here reveal recent hybridization and widespread genetic introgression from domestic pigs into European wild boar populations. We identified introgression by analysing the wild boar allele frequency spectrum, which showed an excess of rare polymorphisms (Figure 2.1a). These putative introgressed SNPs were exclusive to just nine individuals out of 88 sampled wild boar, from dispersed geographical origins (Figure 2.3). The nine putative hybrid wild boar also displayed elevated levels of observed heterozygosity (Table 2.1) compared to other wild boar (Table 2.2). When we included a sample of domestic pigs in a PCA, these nine individuals were positioned between the wild boar clusters and the domestic pig cluster (Figure 2.2b). The two observed typical domestic pig mtDNA haplotypes in three of these nine individuals further support a scenario of introgression from domestic pigs.

The proportion of hybrid wild boar in this dataset is 10% (Wilson Score 95% Confidence Interval: 5-19%). This is at least as high as previously reported figures (5-10%) for introgression in European wild boar [47]. High levels of recent introgression in the study area were not expected a priori since intensive indoor pig farming is prevailing in the last decades and opportunities for direct contact between pigs and wild boar are considered to be minimal. Opportunities for contact between pigs and wild boar were expected to be more prominent in parts of Eastern and Mediterranean Europe, where free-ranging pig production in semi-wild conditions is still common practice, which is the focus of the abovementioned reports [47].

Hybridization simulations and genomic distributions of introgressed alleles indicate the level of introgression

The results from the hybridization simulation study indicate that the detected cases of introgression are equivalent to F1 hybrids until F5 backcrosses with wild boar (Figure 2.4). The LW hybridization simulation resulted in slightly higher numbers of introgressed alleles, while the BS simulation resulted in slightly lower numbers of introgressed alleles (Figure 2.4). This difference is

most likely caused by different levels of outbreeding and polymorphism in these breeds, leading to different amounts of non-wild boar alleles that can potentially introgress. Contributions of multiple breeds to the genetic introgression in NW European wild boar populations may have contributed to the observed numbers of introgressed alleles per hybrid wild boar.

Mendelian inheritance and recombination (crossing over) result in the inheritance of chromosomal segments from each parent. In a scenario of hybridization followed by backcrossing with wild boar one would expect pig alleles to be found only in the chromosomal segments that originate from the parent with domestic pig ancestry. The clustered patterns of introgressed SNPs in individuals 1, 2, 3 and 5 fit this expectation (Figure 2.5), and support their assignments as recent hybrids by the hybridization simulation study. Considering a generation time of one year for wild boar, we can put these hybridization events in the last few years before sampling in 2008. Clustered patterns of introgressed genetic markers resulting from recent hybridization have to the authors' knowledge not been previously described from natural populations.

Hybrid individuals 4, 6, 8 and 9 display a more widespread distribution of introgressed SNPs across the genome (supporting information Figure S2.6). This suggests a more complex scenario of reproduction among hybrids (hybrid x hybrid). These individuals are therefore only equivalent to the assigned generations in the hybridization simulation. The actual wild x domestic hybridization may have taken place a number of generations further back in time followed by interbreeding among hybrids, which kept the number of introgressed SNPs per individual relatively high over an extended time frame. For example, a 3rd generation hybrid x 3rd generation hybrid cross would result in offspring with on average the same number of introgressed alleles as their parents, but it would be the 4th generation since the hybridization event. Sexual reproduction and recombination between different hybrid genomes with distinct individual patterns of introgressed SNP clustering will result in more widespread distribution of introgressed SNPs at every generation of reproduction among hybrids. We consider the time frame of introgression for these advanced-generation hybrids to be uncertain.

Wild boar number 7 is assigned as a first generation hybrid. Intuitively one would expect to find a first generation hybrid at the equidistance between wild and domestic in a PCA. However, one has to keep in mind that in PCA a mean centring procedure is applied. This leads to a gravitation of intermediate individuals (i.e., hybrids) to the origin (0, 0) of the PCA plot, which explains the position of wild boar number 7 at the centre of Figure 2.2 rather than of the equidistance between wild and domestic.

We show that genome-wide SNP analysis can reveal the level of introgression (F1-F5 hybrids or equivalent) by identifying putative introgressed SNPs based on allele frequency spectrum analysis, followed by a comparative analysis of the simulated number of introgressed SNPs per individual and the observed number of introgressed SNPs per individual (Figure 2.4). Assignments of generations (F1-F5 or advanced-generation hybrids) can be further validated by the identification of introgressed chromosomal segments. These methodologies can be applied to all study systems where large numbers of genome-wide genetic markers are shared between the study taxon and the source of introgression. The growing use of high density SNP sets has a promising potential to lead to important insights in the processes of hybridization and genetic introgression.

Mechanisms and sources of introgression

The putative introgressed SNPs found in wild boar are by definition polymorphic in domestic pig, because the Illumina porcine SNP60 genotyping beadchip was ascertained on four domestic pig breeds (Duroc, Pietrain, Large White and Landrace) and a small sample of wild boar [43]. A relatively small dataset of six domestic pig breeds (n=20 per breed) already accounted for 89% of the additional SNPs found.

The domestic pig breeds included in our analysis shared relatively similar proportions of putative introgressed SNPs (Table 2.3). Only Duroc and Tamworth displayed lower amounts of shared SNPs and are deemed unlikely to have been involved in the identified cases of introgression. These findings suggests that introgression was not a singular event, but that it occurred on multiple occasions originating from multiple sources or pig breeds. The presence of two distinct pig mtDNA haplotypes that are not found together in any

domestic pig breed (supporting information Table S2.2) confirms that multiple sources of introgression were involved.

The commercial Large White and Landrace breeds seemed most likely to have contributed to the introgression, as they shared the highest number of SNPs with the nine hybrid wild boar (Table 2.3). However, these breeds were well represented in the ascertainment pool of the Illumina porcine SNP60 genotyping beadchip. Overestimation of the contributions of these breeds versus breeds not included in the ascertainment pool is therefore possible. Still, these breeds share far more putative introgressed SNPs with the nine hybrid wild boar than some other breeds included in the ascertainment pool (Duroc and Pietrain). The observed mtDNA haplotype HP8 most likely entered the NW Europe wild boar gene pool through the Large White or Landrace breeds, which are the most common commercial breeds in the study area. The observed Asian mtDNA haplotype HP110 most likely originated from one of the traditional British pig breeds, as these are the only breeds in this part of the world that display significant levels of this mtDNA haplotype (supporting information Table S2.2).

Possible mechanisms for introgression are (1) crossbreeding with escaped or field-reared domestic pigs, or (2) escape/release of already hybridized (farmed) wild boar stock. Farmed wild boar are often crossbred to a certain extent with a number of domestic pig breeds to increase litter size and piglet growth rates [72]. In certain areas of Europe the documented occurrence of escaped farmed wild boar is substantial [51].

Three wild boar (individuals 2, 3 and 5) were hybrids between domestic pigs and wild boar from the Veluwe (Figure 2.2), but their geographic sampling locations fell within the range of the Meinweg population (Figure 2.3). This finding suggests that the second mechanism, escape/release of hybrid farmed wild boar, has occurred at different places. The observed mtDNA haplotypes of individuals 2, 3 and 5 (HP110 and HP19) suggests that a hybridized farmed wild boar stock with ancestry in the Veluwe wild boar population and British traditional pig breeds is present in NW Europe, and that this hybrid farmed wild boar stock has introgressed into some free-living wild boar populations.

The route by which mtDNA haplotype HP8 has entered the wild boar gene pool, which represents a separate hybridization event, remains uncertain. However, the genomic distribution pattern of introgressed SNPs in the hybrid

with this haplotype (individual 4) suggests an advanced-generation hybrid similar to individuals 6, 8 and 9. The most likely scenario seems to be escape or release of a hybrid wild boar stock influenced by Large White or Landrace pigs, which resulted from an older hybridization event followed by interbreeding among hybrids.

The relatively low number of shared introgressed SNPs between the nine identified hybrids and wild boar from the Balkans (Table 2.3) indicates that natural introgression of alleles from eastern European wild boar cannot explain our observations. We consider the low number of shared introgressed SNPs in Balkan wild boar to reflect a history of free-ranging pig farming practices with associated exchange of genetic material between domestic pigs and wild boar in Mediterranean Europe [47]. Recent genetic contributions from Eastern European wild boar into the study area are considered to be negligible.

Possible effects of introgression

The domestic pig breeds that are possibly involved in the identified introgression (Large White, Landrace, British Saddleback, etc.) carry dominant white spotting alleles. This could lead to deviating coat colour in hybrids, particularly in the first generation. Although no phenotypic details were recorded in this study, all wild boar samples were taken from animals identified in the field as true wild boar and therefore strong deviations in coat colour are unlikely. If the identified hybrids originate from a hybrid farmed wild boar stock as suggested in some cases by discrepancies in genetic association and geographic distribution, these animals may have been subject to artificial selection against the domestic phenotype during their farm history. Anecdotal reports of wild boar with deviating coat colour in Northwest Europe are very rare.

Farmed wild boar are often crossbred to a certain extent with domestic pigs to increase piglet growth rate and litter size [72]. Geographic differences in wild boar litter size have been previously reported in Western Germany [73]. These may be a result of local differences in the level of genetic introgression from domestic pig through the escape or release of hybrid farmed wild boar.

Wild boar numbers have increased markedly in Europe since the 1960s [35, 41, 74]. This population growth and accompanying range expansion has been associated with mild winters and increased food availability through augmented

mast frequency and changes in agriculture [74, 75]. In some areas genetic introgression from domestic pigs may have added to the rapid population growth in the last decades.

Acknowledgements

We thank Jörg Brün, the Animal Health Service Deventer in the Netherlands as well as the Landesuntersuchungsamt of Rhineland-Palatinate in Germany for sample contributions. Thanks also go to Robert Kraus for technical support. Finally, we thank the Royal Dutch Hunters Association (KNJV) for financial support.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S2.1. Information on individual wild boar samples including sex, age in months, sampling location (national park or municipality), sample source (collecting person or institute), PCA clustering with hybrid identification, STRUCTURE population assignment (K=4) and population assignment probabilities of STRUCTURE K=3.

Table S2.2. MtDNA haplotypes of pig breeds and wild boar populations. This table includes additional samples from the Porcine HapMap Consortium (total *n*=699).

Table S2.3. Sample details and Genbank accession numbers of the mtDNA D-loop sequences that formed the basis of the mtDNA haplotypes used in this study (see also supplementary information Table S2.2).

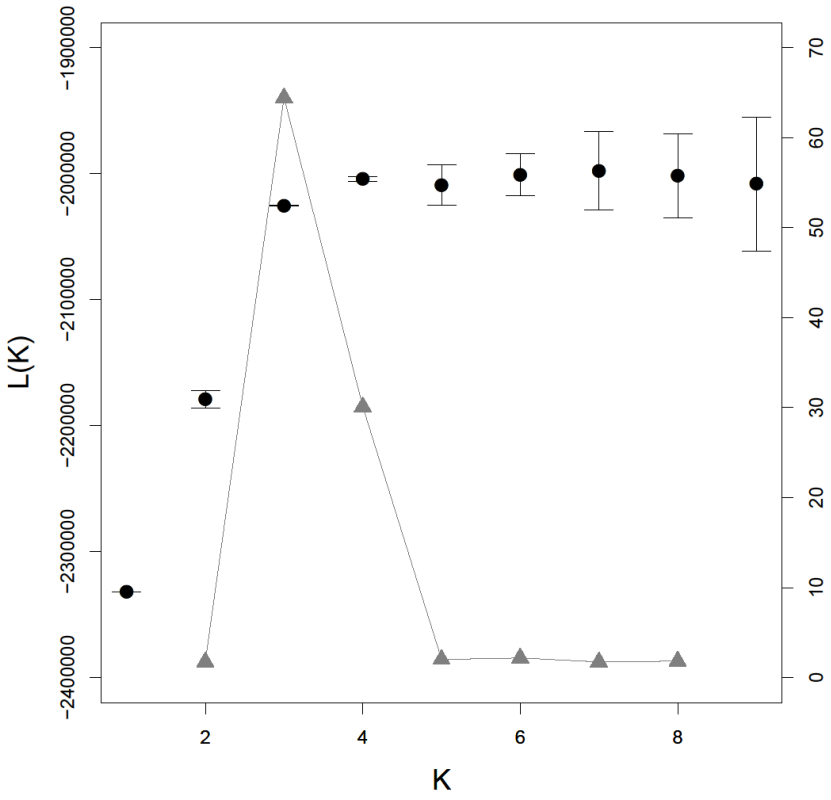


Figure S2.4 L(K) indicated by points with standard deviation bars and Delta(K) indicated by triangles connected by a solid line, per K of the performed STRUCTURE runs following the method of Evanno (2005).

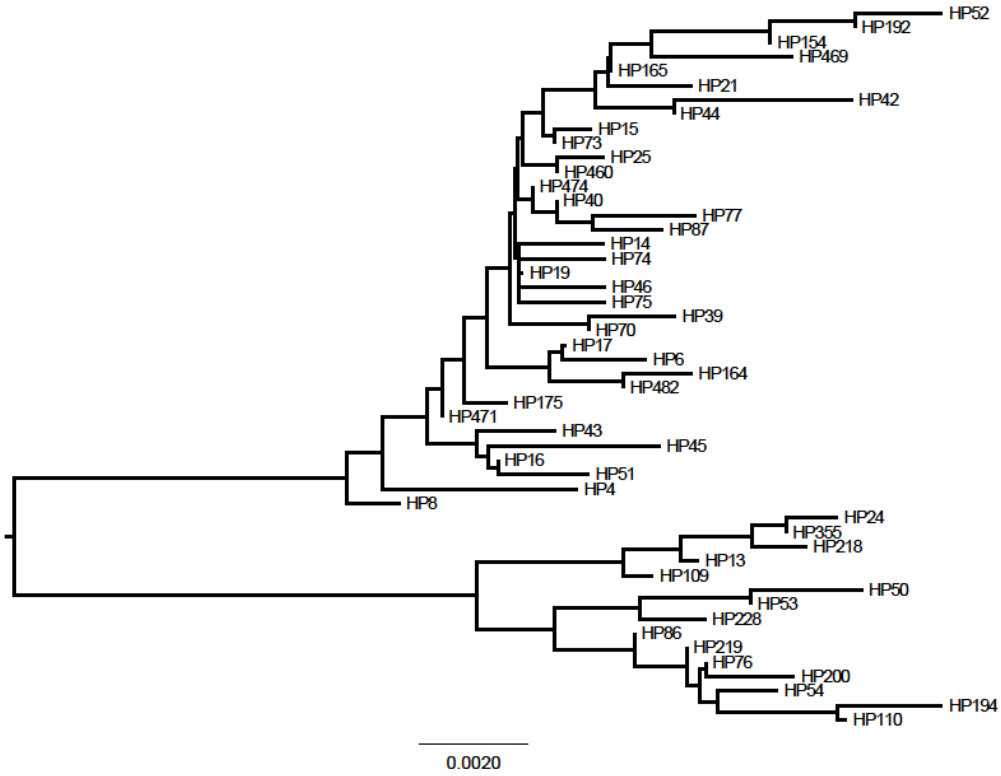


Figure S2.5 Neighbour Joining tree of swine mtDNA D-loop haplotypes. The basal split is between the European haplotypes (EU) and the Asian clades (AS). We included three sequences from Fang & Andersson [71] and follow their interpretation. Common wild boar mtDNA haplotypes in this study are HP19, HP164 and HP165. Additional haplotypes found in putative hybrids are HP8 and HP110.

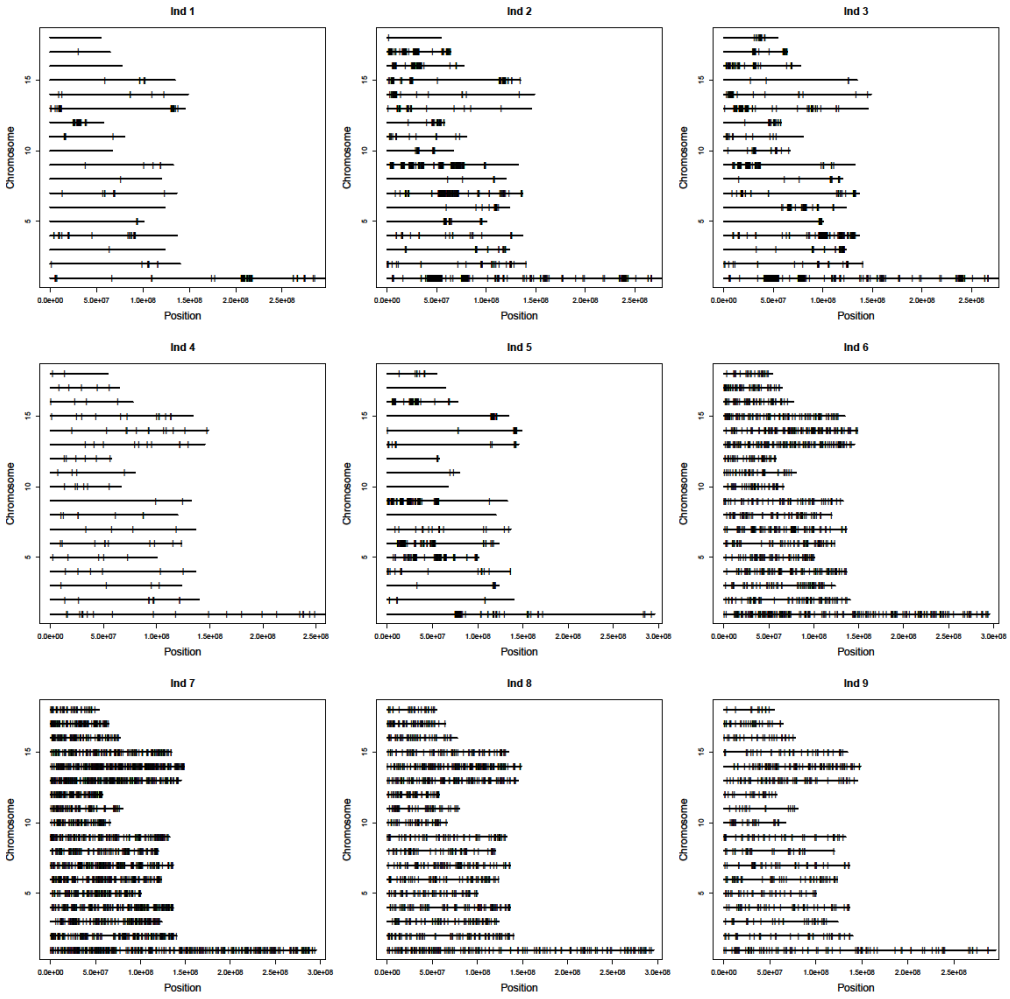


Figure S2.6 Chromosomal positions of introgressed SNPs for all nine identified hybrids.

Data accessibility

The 688 mtDNA D-loop sequences used in this study were submitted to GenBank, accession numbers ranging JQ238239-JQ273541. For more detailed information on these mtDNA D-loop GenBank accession numbers, see supporting information Table S2.3. The 45720 autosomal SNP genotypes for 88 wild boars and 120 domestic pigs (PLINK and STRUCTURE file format) were deposited in the Dryad data repository: doi:10.5061/dryad.v6f1g

Chapter 3

Reintroductions and genetic introgression from domestic pigs have shaped the genetic population structure of Northwest European wild boar

Daniel J Goedbloed, Pim van Hooft, Hendrik-Jan Megens, Katharina Langenbeck, Walburga Lutz, Richard PMA Crooijmans, Sip E van Wieren, Ron C Ydenberg, Herbert HT Prins

Abstract

Population genetic studies focus on natural dispersal and isolation by landscape barriers as the main drivers of genetic population structure. However, anthropogenic factors such as reintroductions, translocations and wild x domestic hybridization may also have strong effects on genetic population structure. In this study we genotyped 351 Single Nucleotide Polymorphism markers evenly spread across the genome in 645 wild boar (Sus scrofa) from Northwest Europe. We show that wild boar genetic population structure is influenced by historical reintroductions and by genetic introgression from domestic pigs. Six genetically distinct and geographically coherent wild boar clusters were identified in the Netherlands and Western Germany. The Dutch Veluwe cluster is known to be reintroduced, and three adjacent Dutch and German clusters are suspected to be a result of reintroduction, based on clustering results, low levels of heterozygosity and relatively high genetic distances to nearby populations. Recent wild x domestic hybrids were found geographically widespread across clusters and at low frequencies (average 3.9%). The relationship between pairwise kinship coefficients and geographic distance showed male-biased dispersal at the population genetic level. The current trend of wild boar population growth and range expansion has recently

led to a number of contact zones between clusters, and further admixture between these wild boar clusters is to be expected. In conclusion, our results demonstrate how wildlife and landscape management by humans are shaping the genetic diversity of an iconic wildlife species.

BMC Genetics 2013 14:43

Introduction

Most population genetic studies consider dispersal and isolation by landscape barriers to be the main drivers of genetic population structure [76]. However, human activities such as reintroductions, translocations and genetic introgression from domestic sources, may play an important role in certain study systems, in addition to natural dispersal and landscape patterns [49, 77, 78]. Such human activities, legal or not, are often poorly documented and their population genetic effects are mostly unknown. Molecular techniques provide increasingly powerful and affordable tools to evaluate anthropogenic influences on wildlife genetic population structure [79, 80]. The use of Single Nucleotide Polymorphisms (SNPs) in particular is promising for the fields of population and conservation genetics [45, 58].

Wild boar became extinct in large parts of Western Europe in the 19th century [35]. The species was marginalized mainly by overhunting and deforestation associated with increased agricultural land use. Extinction in Britain had already occurred in the 13th century [81]. This massive decline in Western Europe was followed by an unknown number of mostly undocumented reintroductions in the late 19th and early 20th century. One such event is the commonly known but undocumented reintroduction of wild boar to the Veluwe, the forested centre of The Netherlands, which occurred in 1904 at the orders of Hendrik, Prince-Consort of Queen Wilhelmina of The Netherlands, for the purpose of hunting [82]. These animals are thought to stem from Northeast Germany and Czech Republic.

Conditions for wild boar steadily improved during the 20th century due to hunting restrictions, reforestation, changes in agriculture and possibly climate change [75, 83]. Starting from 1960, wild boar populations throughout Europe saw rapid growth and range expansion [41, 74]. Wild boar (*Sus scrofa*) are adaptive and opportunistic omnivores as well as good dispersers, being able to travel distances up to 250 km [84] and fast breeders, with litter sizes of 4-7 once a year [35]. Dispersal is male-biased in this species [35, 37]. European wild boar population structure at the continental scale is mainly shaped by post-glacial colonization patterns [51]. It is, however, unknown how the history of marginalization, reintroductions and recent population expansion has affected

the genetic population structure at local or regional scales. In an area such as The Netherlands and Western Germany, one could expect high rates of gene flow.

Wild boar farming became popular in Europe in the second half of the 20th century to provide for a demand in luxury meat. Hybridization between wild boar and domestic breeds is common practise on these farms to achieve increased reproduction and growth rates [85]. Such hybrids have been shown to be the source of the escaped wild boar population in England [86]. Introduction of wild boar originating from hybrid farmed stocks has also been shown in mainland Europe (chapter 2). This has effectively led to genetic introgression from domestic pigs into local wild boar populations. Recent hybrids (until 5th generation backcrosses with wild boar) as well as advanced generation hybrids (resulting from reproduction among hybrids across multiple generations) were identified. However, the spatial extent of domestic introgression and its effects on the population genetic structure of European wild boar has not been studied in detail.

From an evolutionary point of view, possible adverse effects of genetic introgression from a domestic or hybrid source include genetic adaptation to captivity and possibly outbreeding depression [87], while possible advantageous effects include hybrid vigour, increased growth rates and larger litter size. These evolutionary advantageous effects may be undesirable from a management perspective, as more rapidly reproducing wild boar can be difficult to control using normal population management practices and can then cause significant damage to agricultural crops [88]. Strikingly high litter sizes and strong differences in litter size between regions have indeed been observed in wild boar in Germany [73]. In addition to evolutionary effects, also population composition and structure can be affected by hybrid introductions and restocking practices [89].

In this study we used 351 SNP markers, genotyped for 645 wild boar, including 88 samples from chapter 2, to assess the effects of historical marginalization, reintroductions and genetic introgression from domestic pigs on the population genetic structure of wild boar in The Netherlands and Western Germany.

Methods

Blood or tissue samples were taken from a total of 645 wild boar in parts of The Netherlands, Western Germany and Luxembourg at the opportunity of routine wildlife management and disease monitoring programs. This included 88 samples from chapter 2, which were genotyped using the Illumina porcine SNP60 genotyping beadchip [43]. All samples were collected in the years 2008-2010 from animals identified in the field as wild boar.

DNA was extracted using the Qiagen PureGene (Blood) kit protocol. Samples were genotyped for 384 SNPs selected from the Illumina porcine SNP60 genotyping beadchip [43] from loci known to be polymorphic in wild boar in the study area, with proportional coverage of each chromosome and random selection within each chromosome. Of these 60k SNPs, 76% proved to be polymorphic in our wild boar dataset. Random selection within the autosomal and X chromosomes was performed to minimalize ascertainment bias. The only possible remaining ascertainment bias in our SNP set is derived from the ascertainment panel of the Illumina porcine SNP60 genotyping beadchip itself, and is considered to have no effect on the inference of wild boar population structure in the study area. Less than 0.015% of the pairwise distances between the 351 randomly chosen SNPs were closer than 100,000 bp, which is considered to be the maximum range of physical linkage in wild mammals [90, 91]. Selected SNPs were genotyped on an Illumina GoldenGate bead array platform (BeadXpress, Illumina Inc.) in a 96 well, 384 SNP format [92]. Genotyping quality was assessed using GenomeStudio software (Illumina Inc.). Low genotyping quality or lack of differentiation between homozygote and heterozygote clusters lead to the removal of 33 SNPs. This left 351 non-coding SNPs for data analysis, which is roughly equivalent in statistical power to 140 microsatellites [93, 94].

Linkage Disequilibrium (LD) was analysed in PLINK v1.06 [60] by calculating all genome-wide pairwise SNP-SNP correlation coefficients (r^2) and assuming a 0.2 threshold. Principal Components Analysis (PCA) was performed to visualise genetic variation and possible clustering patterns using the eigenvector method implemented in EIGENSOFT 3.0 [61, 62]. For comparison, a sample of 120 domestic pigs from six breeds was used (Large White, Landrace, Duroc, Pietrain, British Saddleback and Tamworth, n=20 per breed). We used

STRUCTURE [63] for population assignment analysis with 10 runs per number of clusters (K) for K=1-10 with 500,000 iterations and a burnin of 800,000. Optimal partitioning was evaluated using the method proposed by Evanno *et al*. [64]. Phylogenetic network analysis was performed using SplitsTree4 [95]. A number of R packages were used: Adegenet [65] for heterozygosity calculations, Hierfstat [96] for calculation of F_{ST} values, SNPRelate [97] for the Maximum Likelihood Estimation calculation of kinship coefficients [98] based on the method of Thompson [99], and finally Vegan [100] for mantel tests in the Isolation By Distance (IBD) analysis, where genetic distance was calculated as F_{ST} /(1- F_{ST}) between all sampled locations.

Results

The 351 genotyped SNPs had an overall call rate of 0.98 and 0.013% of all pairwise SNP combinations interfered with linkage equilibrium. Of these pairwise SNP combinations in LD, a quarter (0.003% of total pairwise SNP combinations) were most likely based on physical linkage up to distances of 100,000 basepairs [90, 91], while three quarters (0.010% of total pairwise SNP combinations) were found beyond this distance (up to 750kb), but still within the same chromosome.

We screened for wild boar-domestic pig hybrids by applying a STRUCTURE likelihood assignment minimum threshold of 0.25 (25%) to a sample of domestic pigs (*n*=120, see Methods). Individual assignment proportions for K=1-7 are indicated in Figure 3.1. The assignment threshold of 0.25 was chosen basd on the absence of false positive hybrids among the 88 previously studied samples (Table 3.1, see also chapter 2). At this threshold, all five recent hybrids (up to fifth generation backcrosses with wild boar) identified previously by Allele Frequency Spectrum Assessment (based on introgressed alleles, chapter 2) were correctly identified by STRUCTURE, in contrast to the four advanced generation hybrids (Table 3.2). The STRUCTURE algorithm identified a total of 25 recent hybrids in 645 wild boar samples (3.9%, 95% Wilson Score CI: 2.6-5.7%). This percentage is similar to previous reports [51], but here it represents recent hybrids identified by allele frequency signatures that rapidly degrade over generations, whereas previous studies may have reported

hybrids based on long-term genetic signatures (e.g., mitochondrial DNA haplotypes).

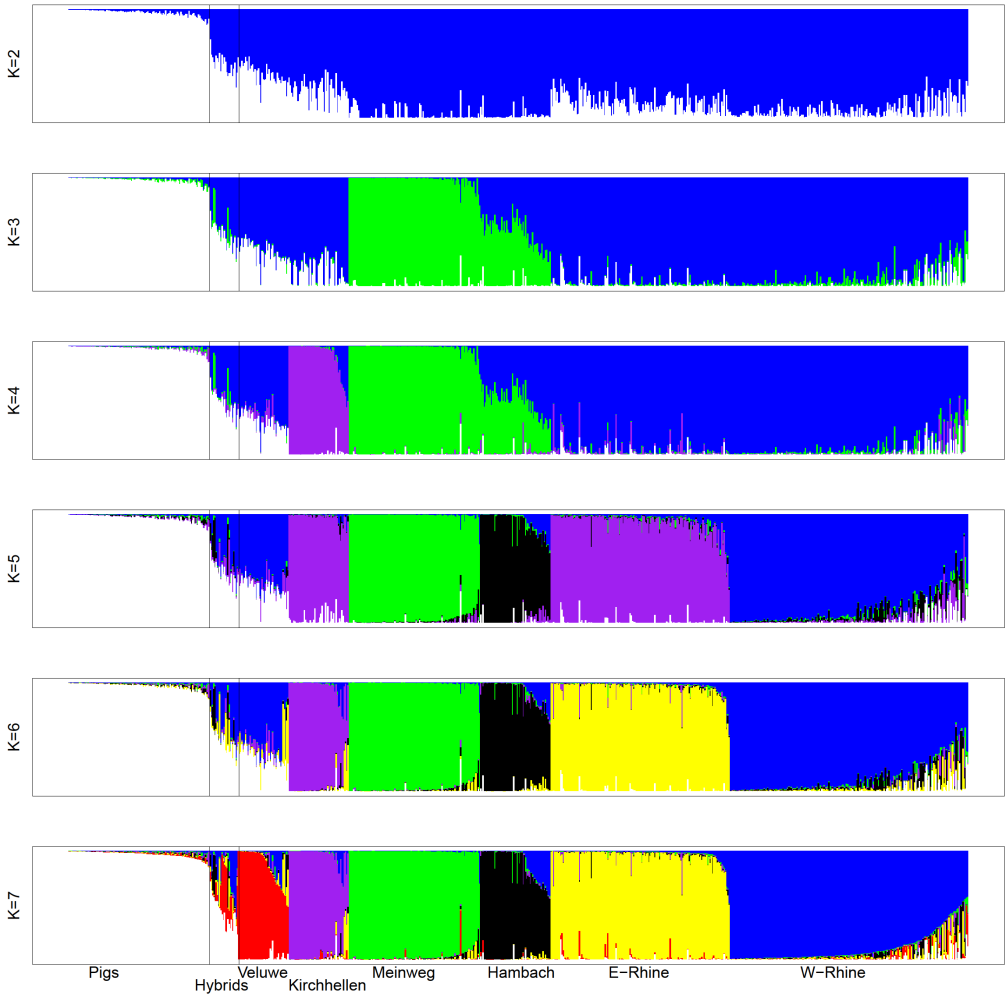


Figure 3.1 Population assignment proportions per individual based on results from STRUCTURE for K=2-7. Recent wild x domestic hybrids, sampled in the field as wild boar, are delimited by vertical lines. Results for K=5 were not ambiguous across runs. Majority rule results (*n*=10) are presented here, but the inclusion of E-Rhine in Kirchhellen at K=5 is not fully supported, as various alternative clustering patterns were also inferred. Evanno's method favoured optimal partitioning at K=7 (Figure S3.1).

Table 3.1 Results of hybrid detection using STRUCTURE at different assignment thresholds. Comparisons were made to results from chapter 2, which identified nine hybrids from a total of 88 samples using analysis of introgressed allelic states with the SNP60 genotyping beadchip.

Assign threshold	>0.30	>0.25	>0.20	>0.15	>0.10	
Total hybrids [1]	18	25	30	36	45	
Shared hybrids [2]	3	5	6	6	7	
SNP60 only [3]	6	4	3	3	2	Type II error
STRUCTURE only [4]	0	0	1	4	4	Type I error

[1] the total number of hybrids detected in this study by STRUCTURE

[2] the number of hybrids from the SNP60 study that was correctly detected also by STRUCTURE

[3] the number of hybrids from the SNP60 study that were not identified by STRUCTURE (type II error)

[4] the number of individuals that were incorrectly labelled as hybrids by STRUCTURE (type I error)

Table 3.2 The nine previously studied SNP60 hybrid individuals, listed by their being detected or not by STRUCTURE at an assignment threshold of 0.25 (see Table 3.1). Individual numbering corresponds to chapter 2. The level of introgression is based on the number of introgressed domestic alleles per individual and expressed as being equivalent to the number of generations since hybridization according to simulations (chapter 2). The type of hybrid (recent versus advanced generation) is distinguished based on the genomic distribution of introgressed alleles (clustered or spread out respectively, see chapter 2).

	Individual	level	type
Detected	7	1st	Recent
	2	2nd	Recent
	5	3rd	Recent
	1	4th	Recent
	3	2nd	Recent
Not detected (type II error)	9	3rd	Advanced
	6	2nd	Advanced
	8	2nd	Advanced
	4	5th	Advanced

Both STRUCTURE clustering and PCA show a clear wild - domestic separation (Figure 3.1 and 3.2). The recent hybrids that are detected by STRUCTURE are associated with intermediate positions between wild boar and domestic pigs as well as the origin of the plot (0,0) in the PCA (Figure 3.2). The four individuals identified as advanced generation hybrids using SNP60 genotyping (chapter 2) are scattered across the wild boar clusters, without visible association to the domestic pig cluster.

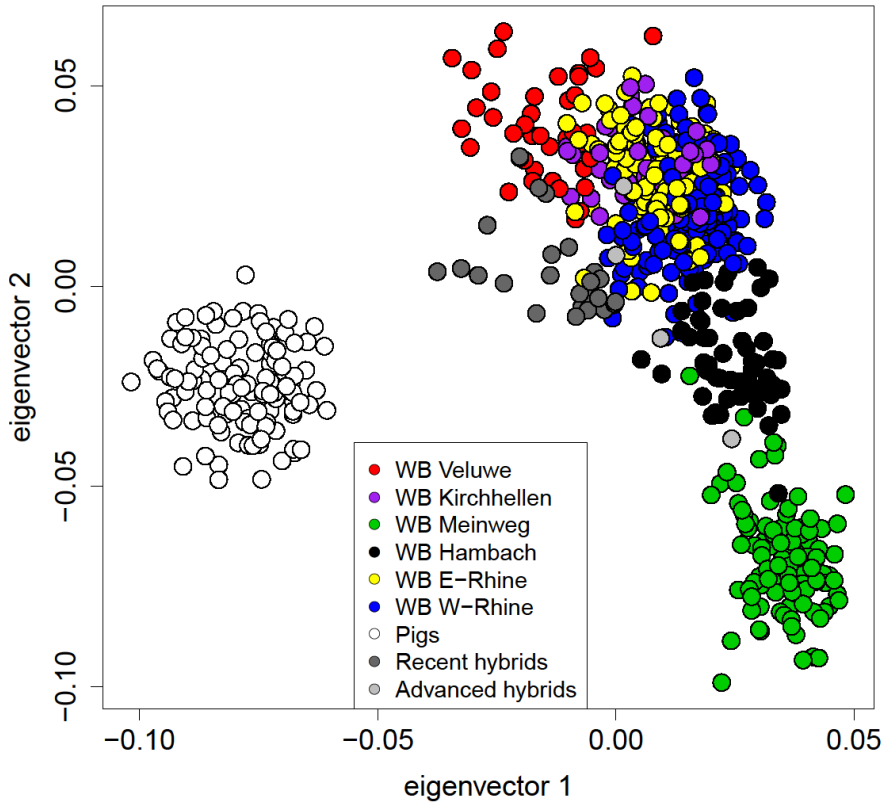


Figure 3.2 PCA plot of the wild boar and a sample of domestic pigs (colours correspond to Figure 3.1), indicating genetic variation along the first two eigenvectors. The 25 recent wild boar x domestic pig hybrids identified by STRUCTURE (threshold assignment proportion 0.25) are indicated in dark grey and four additional advanced generation hybrids with introgressed pig alleles identified in chapter 2 are indicated in light grey.

Table 3.3 Genetic wild boar clusters with the corresponding sample size (*n*), observed heterozygosity (H_o) and number of hybrids (based on geographic association, excluding 5 hybrids with uncertain geographic assignment).

cluster	*n*	H_o*	hybrids
Pigs	120	0.36	
Veluwe	43	0.36	0
Meinweg	112	0.35	2 (1.8%)
West Rhine	207	0.41	12 (5.8%)
Hambach	60	0.40	2 (3.3%)
East Rhine	153	0.40	3 (2.0%)
Kirchhellen	50	0.34	1 (2.0%)

* standard errors were 0.01 or smaller

Following the method of Evanno *et al*. [64], six genetic wild boar clusters were identified (Table 3.3, and Supporting information, Figure S3.1). These genetic clusters were supported by separation along the first four eigenvectors in a PCA (Figure 3.3), which explained 43% of the total variation. F_{ST} values indicated moderate ($0.05<F_{ST}<0.15$) to high ($0.15<F_{ST}<0.25$) genetic differentiation between the inferred clusters (Table 3.4). In addition, the identified genetic clusters were geographically non-overlapping (Figure 3.4), with one possible exception (Hambach, in black). This geographic separation supports the inferred clustering and its interpretation as a biologically meaningful population structure. The River Rhine seems to act as a boundary between genetic clusters, although some gene flow occurs across the Rhine in Germany. Isolation by Distance (IBD) across clusters was near significant (p=0.061), even though it was not significant within some of the clusters (Table 3.5). A Fisher's combined probability test indicated that overall, the within cluster IBD is significant (p=0.008) in the study area.

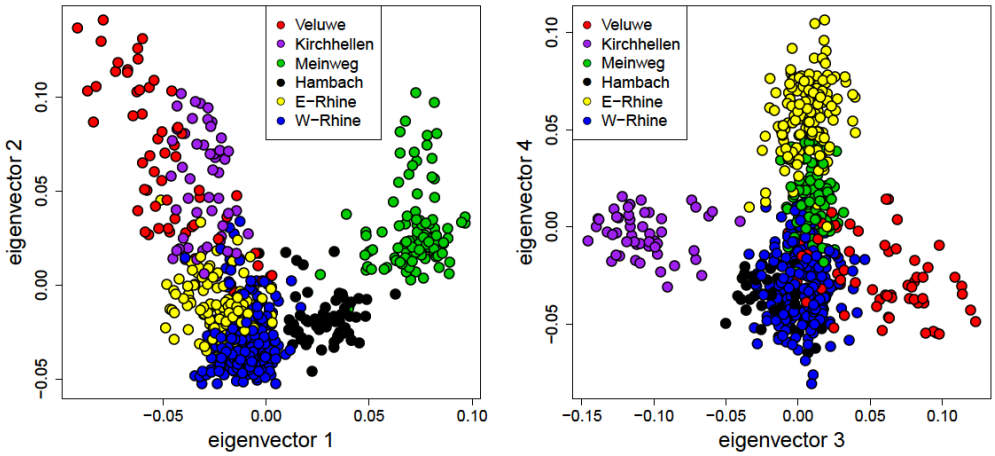


Figure 3.3 PCA plots indicating the first four eigenvectors of the wild boar data only. Colours indicate the six clusters identified by STRUCTURE. Putative hybrids are not indicated in this figure. Eigenvectors 1-4 explain 43% of variance in the dataset.

Table 3.4 Autosomal F_{ST} values between wild boar clusters (and domestic pigs). Above the diagonal: F_{ST} values without hybrids. Below the diagonal: F_{ST} values with hybrids.

	Kirchhelle	Meinwe	Veluw	East-	West-	Hambac
Pigs	0.193	0.234	0.150	0.158	0.162	0.192
Kirchhelle		0.215	0.170	0.125	0.124	0.171
Meinweg	0.212		0.214	0.139	0.121	0.108
Veluwe	0.149	0.189		0.111	0.108	0.165
East-Rhine	0.123	0.137	0.093		0.050	0.098
West-	0.119	0.117	0.086	0.047		0.069
Hambach	0.168	0.106	0.140	0.096	0.066	

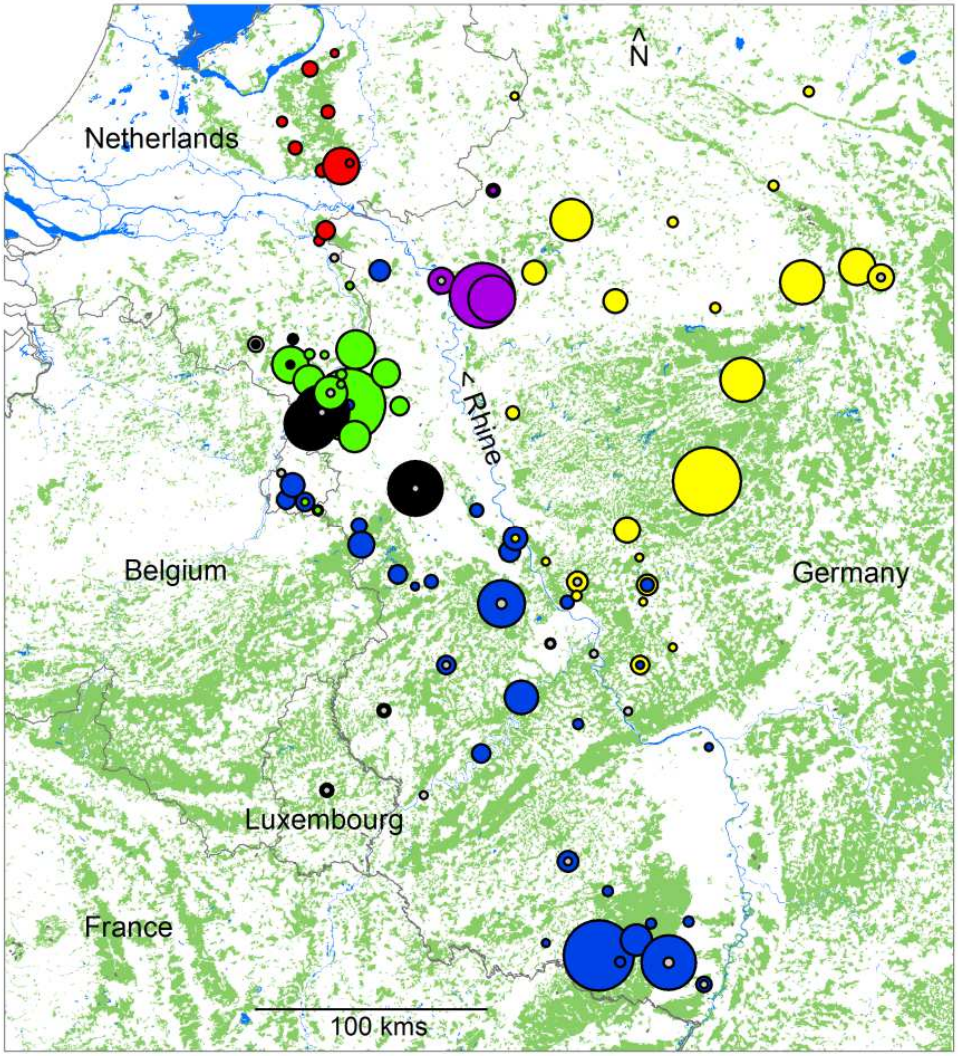


Figure 3.4 Map of the study area. Country borders are indicated by black lines, forests are indicated in soft green and inland water features in light blue. Dots indicate wild boar sampling sites. The size of the dot is relative to the sample size. The colours indicate genetic clustering by STRUCTURE and correspond to other Figures. Hybrids identified by STRUCTURE (domestic cluster assignment proportion >0.25) are indicated in grey.

Table 3.5 Isolation by distance (IBD) analysis results for the full dataset and the different clusters separately, using mantel tests (10,000 permutations, 10 repeat average). P-values indicate the significance of IBD across sampling locations in that particular dataset or cluster. The maximum pairwise geographic distance within the cluster or dataset is also given.

	Nr	max. dist.	*p*-value
Full dataset	101	402	0.061
Veluwe	10	76	0.326
Meinweg	15	50	0.166
Kirchhellen	4	44	0.334
Hambach	5	86	0.084
E-Rhine	30	240	0.085
W-Rhine	44	343	0.020

Phylogenetic network analysis displayed monophyly for the domestic pigs and the six wild boar clusters (Figure 3.5). The hybrids identified in this study are divided into three separate lineages. We recalculated the F_{ST} values after excluding all identified recent hybrids to avoid possible biases due to both increased genetic variation within clusters and decreased variation across clusters caused by the scattered presence of hybrids. This exclusion of hybrids resulted in on average 0.0093 (8%) higher pairwise F_{ST} values (Table 3.4), and represents a confounding effect of scattered hybrids on population differentiation.

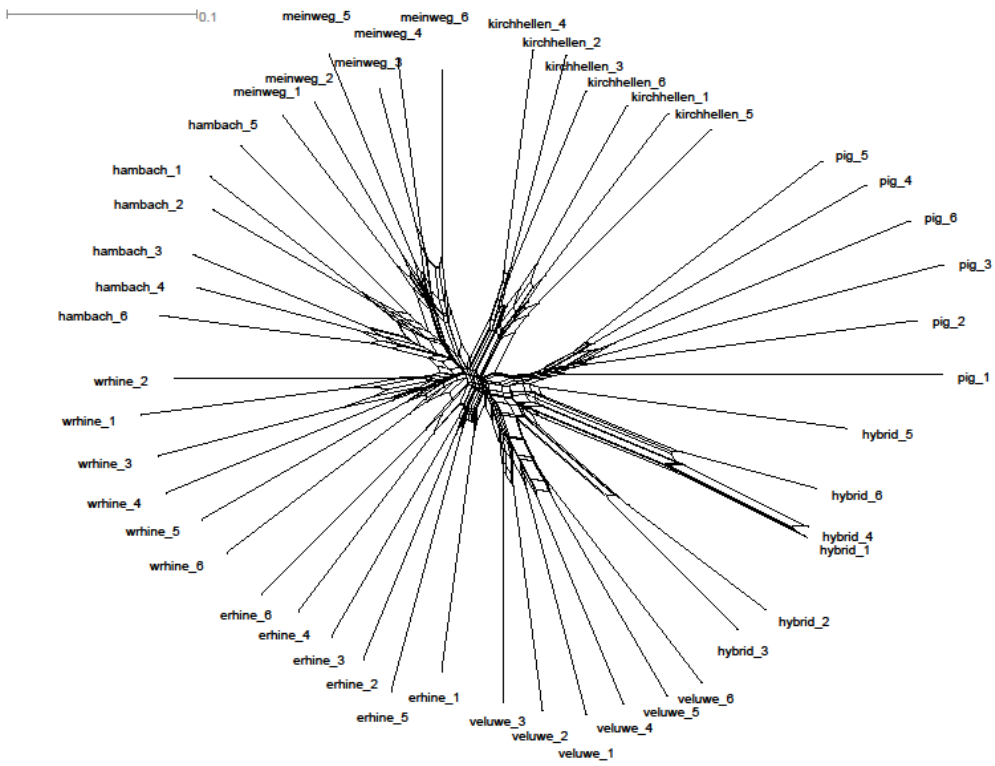


Figure 3.5 NeighborNetwork of six representative samples per wild boar cluster and one sample per domestic pig breed (see methods). The number of samples was chosen for optimal balance in information content and clarity of the figure. Distances are based on the uncorrected P (or Hamming) method.

The pairwise kinship coefficient is a measure of relatedness (consanguinity) between two individuals. Analysis of pairwise kinship coefficients in the wild boar dataset showed a decrease of pairwise kinship over geographic distance (Figure 3.6 and Figure S3.2). Females displayed relative site-fidelity (higher levels of kinship at distances less than 25 km) and males showed relatively high dispersal rates (indicated by higher kinship coefficients at distances between 25 and 150 km), demonstrating effects of male-biased dispersal in this species at the population genetic level. These kinship effects of dispersal up to distances of 150 km attest to the high dispersal capacity of wild boar and correspond to occasional high dispersal distances observed in mark-recapture studies (e.g., [84]).

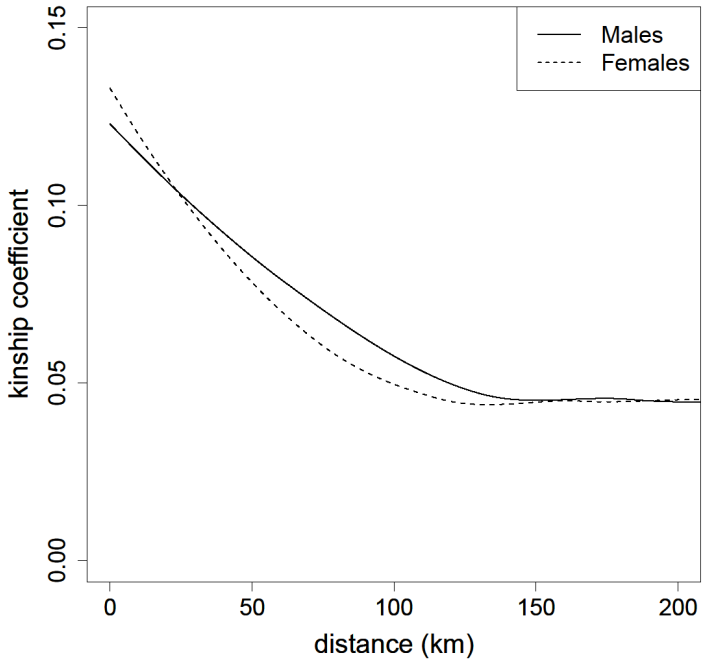


Figure 3.6 Pairwise kinship coefficients of both sexes versus geographic distance. Results are based on local polynomial regression analysis. Females show relative site fidelity at pairwise distances less than 25 kilometres, and males show higher kinship coefficients at distances between 25 and 150 kilometres, indicating higher dispersal rates.

Discussion

Population genetic patterns and historical reintroductions

The largest wild boar populations in this study are found in Germany (West-Rhine and East-Rhine, Figure 3.4). They are relatively closely related (Table 3.4 and Figure 3.5) and most likely represent historically continuous wild boar populations. A high density of closely connected forest patches facilitates dispersal and genetic homogenisation in this part of the study area, and is only bisected by a natural barrier: the River Rhine (Figure 3.4). This barrier is not complete, as a few individuals seem to have crossed the Rhine in Germany. The

barrier function of the River Rhine is, however, apparently sufficient to cause clear population differentiation between these clusters (F_{ST} =0.050, Table 3.4).

The wild boar found just South of the Rhine in the Netherlands, which belong to the Veluwe cluster (Figure 3.4), most likely represent an anthropogenic translocation event, as the intermediate terrain contains no forest and is intersected by two major rivers (the River Rhine and either the Waal or the IJssel). No wild boar were observed in this area bank until 1983.

The North-western section of the study area is characterised by a low level of fragmented forest cover, which is the main habitat for wild boar in Europe [35]. Historical records show that substantial forest patches appeared in this part of the study area only after the advent of artificial fertilizers and its associated reduction of landscape-wide grazing pressure at the beginning of the 20th century [83]. It is unlikely that wild boar occurred in the North-western part of the study area before 1900, due to a lack of suitable habitat (forest). One cluster (Veluwe) in this North-western section certainly originates from reintroductions in 1904, and the other three clusters (Meinweg, Hambach and Kirchhellen) most likely also arose from reintroductions in the 20th century. This is supported by clear genetic differentiation of each of these clusters (Table 3.4, Figure 3.1 and 3.5) with the other clusters, which may be explained by founder effects and subsequent reproductive isolation. The observed heterozygosity of these four populations is lower than in the Rhine populations (Table 3.3) supporting a historical population bottleneck or founder effect. The only exception is the Hambach cluster, which displays observed heterozygosity levels similar to the Rhine populations, but this may be explained by historical genetic introgression from domestic pigs, as discussed below. The absence of IBD in the (putatively) reintroduced populations: Veluwe, Meinweg, Hambach and Kirchhellen (Table 3.5), could be due to a history of introduction or translocation. On the other hand, absence of IBD may also be caused by a lack of statistical power due to small sample size (number of locations) and relatively small geographical range in these clusters. Wild boar from the Meinweg, Hambach and Kirchhellen are genetically well differentiated (Table 3.4, Figure 3.1, 3. 3and 3.5), even more so than the Veluwe cluster. The sources of the putative reintroductions in Meinweg, Hambach and Kirchhellen are unknown.

The Hambach cluster has a small geographical distribution with two localised foci (Figure 3.4). These two foci consist of small isolated forest patches, one of which is formed by a large brown coal mine in Germany (the Tagebau Hambach, opened in 1978, total surface 8500 hectare) and forested former refuse dump sites and fringes. This area was originally cleared of forest and only in 1980-1982 were the first dump sites (Sophienhohe) reforested, thereby creating opportunities for wild boar (re)colonisation. The other forest patch (Echt-Montfort, the Netherlands) was unoccupied by wild boar until 1983. Only one individual assigned by STRUCTURE to the Hambach cluster (from the Echt-Montfort patch) was included in chapter 2. This individual was then identified as an advanced generation domestic-wild hybrid. Mitochondrial DNA haplotype analysis performed in that study revealed a typical domestic pig mitochondrial haplotype in this individual. The sudden appearance of this clearly distinct wild boar cluster in Hambach and in Echt-Montfort in the 1980s, together with the evidence of genetic influences from domestic pig suggest anthropogenic introduction, most likely from a captive wild boar source. A domestic hybrid origin or influence in this cluster would also explain the relatively high levels of observed heterozygosity in such a small population (Table 3.3).

We assume the three populations (putatively) reintroduced in the early 20th century (Veluwe, Meinweg and Kirchhellen) to have existed in complete reproductive isolation initially. However, wild boar populations across Europe have increased their numbers dramatically since the 1960s [35, 41, 74]. The contact zones between wild boar clusters found in this study based on the geographical overlap of clusters (e.g., Meinweg, Hambach and West-Rhine as well as Kirchhellen and East-Rhine, see Figure 4) are considered to be a consequence of these population expansions and therefore relatively recent. STRUCTURE identified a relatively small number of admixed wild boar (Figure 3.1), all associated with contact zones. This low frequency of admixture supports a recent onset of contact between clusters.

Identification and effects of genetic introgression from domestic pigs

The mechanism for genetic introgression from domestic pigs into wild boar populations is most likely deliberate or accidental introduction of hybrid farmed

wild boar [47] (see also chapter 2). The STRUCTURE algorithm identified 25 geographically scattered recent hybrids in 645 wild boar samples (3.9%). Hybrids are not more frequent in (putative) reintroduced populations, and seem to be recently introduced to various parts of the study area, possibly for the purpose of restocking local hunting grounds.

The STRUCTURE algorithm relies solely on typical domestic pig allele frequencies for domestic-wild hybrid detection. Allele frequencies may change over time due to genetic drift and admixture with local wild boar gene pools. The figures based on hybrid detection by STRUCTURE therefore only represent recent genetic introgression from domestic pigs and are likely to underestimate or disregard historical genetic introgression. Hybrid identification using a STRUCTURE assignment threshold of 0.25 to the domestic pig cluster reliably identified all recent hybrids studied in chapter 2, but not the advanced generation hybrids. This result indicates that allele frequency signatures from both source populations (wild and domestic) were indeed only detectable in relatively recent hybrids (approximately up to five generations of backcrossing, see chapter 6).

Phylogenetic analysis indicated multiple separate lineages within the hybrid group (Figure 3.5), suggesting that different hybridisation events are responsible for the detected genetic introgression from domestic pigs. This corresponds to findings from mtDNA haplotype analysis (chapter 2), which also suggested multiple origins of wild-domestic hybrids in this area.

If low numbers of hybrids are introduced in already occupied wild boar habitat, they would be expected to mate mostly with local wildtype individuals, leading to a rapid dilution of hybrid genetic signal over a few generations (chapter 2). However, if hybrids are to be introduced in areas previously unoccupied by wild boar, reproduction will occur mostly among hybrids. Over time this could lead to local dominance of advanced generation hybrids and a persistent hybrid genetic signal. Advanced generation hybrids such as those produced by the latter scenario would not be identifiable as being of partly domestic origin by STRUCTURE, because allele frequencies are likely to have diverged over time from those of the source populations due to genetic drift and admixture with local wild boar gene pools. However, these hybrids should be detected when analysing Allele Frequency Spectrum Assessment, which is based on introgressed allelic states (see chapter 2). Such a scenario of older

hybridisation followed by introduction to the wild and reproduction among hybrids may have shaped the Hambach cluster.

Exclusion of recent hybrids from our total dataset resulted in an average population F_{ST} increase of 0.0093, corresponding to 8% of the average population F_{ST} (Table 3.4). This demonstrates that domestic introgression may affect the results of population differentiation analysis in certain study systems. Here, only recent hybrids (approximately up to fifth generation backcrosses) could be excluded. Long-term effects of domestic introgression most likely also exist (e.g., in the Hambach cluster), potentially affecting genetic population structure further. The LD among SNPs beyond 100kb distance found in this wild boar dataset may also be a consequence of recent genetic introgression, although effects of population substructure and small local population sizes could not be ruled out or corrected for. As a general recommendation for population genetic analysis, we propose that hybrid detection should be performed in all cases where genetic introgression is deemed possible, to avoid associated biases in population differentiation (F_{ST}) or LD, as well as erroneous interpretations of population structure.

Conclusions

The presence of six well-defined genetic clusters in the study area can be attributed to two factors: the presence of a natural barrier: the River Rhine, and a history of marginalization, extinction and subsequent anthropogenic reintroductions in the Northwest of the study area. Widespread genetic signatures of recent accidental or deliberate restocking of local populations with hybrid farmed wild boar have been found, which confounded population differentiation statistics, but which do not seem to affect the existing population structure.

In this study we demonstrate the effect of past landscape and population management on current population structure in an iconic wildlife species. Effects of historical deforestation and overhunting followed by reintroductions and restocking from farms are evident. Wild boar populations in the study area are currently expanding their range. Previously isolated populations are admixing in recently formed contact zones. The relative contribution of each of the current populations to future wild boar diversity may depend on a number of factors

including the effective size of populations, habitat connectivity, founder effects, restocking activity, introductions and translocations.

Acknowledgements

We thank the Animal Health Service Deventer (GDD) in the Netherlands and dr. Karl Zimmer from the Landesuntersuchungsamt of Rhineland Palatinate in Germany for sample contributions. Thanks go to Bert Dibbits and Thijs Bosch for technical assistance. We appreciate comments on the manuscript by Robert H. Kraus. Finally we thank the Royal Dutch Hunters Association (KNJV) and Stichting De Eik for financial support.

Supporting information

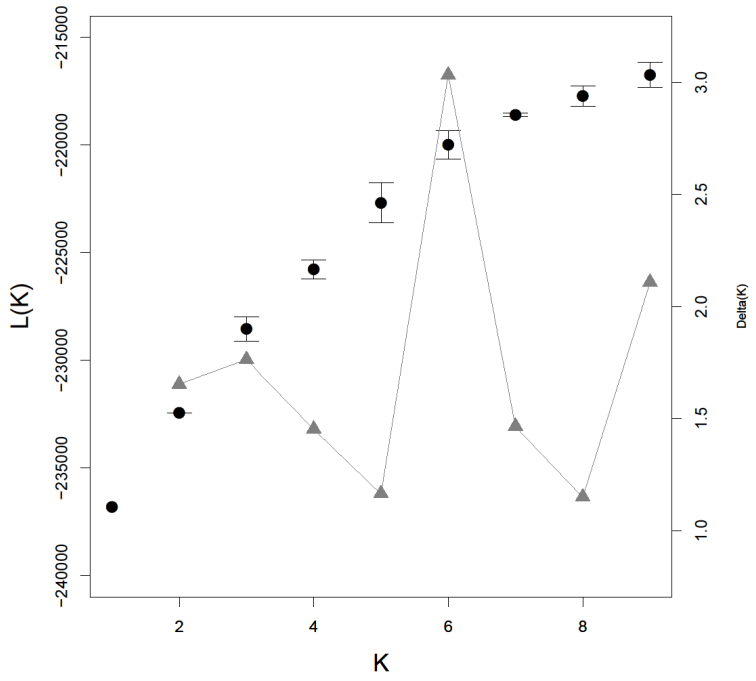


Figure S3.1 The STRUCTURE likelihood parameter L(K) ± s.d. and Evanno's ΔK (grey line) plotted per number of clusters (K) for K=1-10 for the wild boar dataset. Note that the domestic pig cluster was excluded.

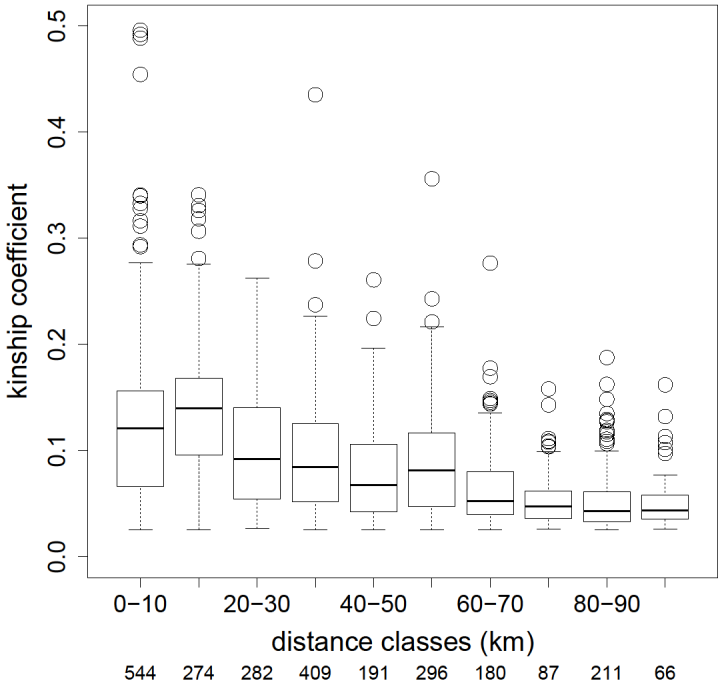


Figure S3.2 Boxplot indicating the variance of kinship coefficients over 10km geographic distance classes. Sample sizes per distance class are given below the x-axis.

Data accessibility

The 351 SNP genotypes of the 645 wild boar (plink format) are available as additional files 1 and 2 with the online version of this article. The 351 SNP genotypes of the 120 domestic pigs (plink format) are available as additional files 3 and 4.

Chapter 4

Climatic conditions, host age and host heterozygosity influence porcine circovirus type 2 disease dynamics in European wild boar

Daniel J Goedbloed, Pim van Hooft, Hendrik-Jan Megens, Thijs Bosch, Walburga Lutz, Sip E van Wieren, Ron C Ydenberg, Herbert HT Prins

Abstract

Zoonotic and emerging diseases are an important public health concern. However, little progress has been made in recent decades to increase our fundamental understanding of the source of these diseases: wildlife disease dynamics. The study of wildlife disease dynamics is complicated by challenges in sampling, diagnostics and a myriad of potentially relevant factors. Here we report the influence of a number of demographic, environmental and genetic factors on porcine circovirus type 2 (PCV2) antibody prevalence in European wild boar (*Sus scrofa*). We genotyped 462 wild boar individuals from the Netherlands and Western parts of Germany in the years 2008-2010 using a genome-wide 351 SNP assay, and performed PCV2 serology on these samples using ELISA assays. We show that individual PCV2 antibody status is dependent on age, year of sampling and genetic heterozygosity using logistic generalized linear regression analysis. The age effect is most likely caused by cumulative PCV2 exposure and intracellular hiding by this virus followed by increasing chances of activation and associated antibody responses. Year of sampling is a significant factor for PCV2 prevalence, most likely because differences in winter temperature affected external survival of the pathogen in aerosols and thereby transmission rates. A positive correlation between heterozygosity and PCV2 prevalence was found to be caused by mortality of

low-heterozygosity individuals. Pairwise heterozygosity correlations between loci suggest a global mechanism of many mildly deleterious effects across the genome, and support the interpretation of inbreeding depression in the context of PCV2 resistance. These findings indicate that PCV2 can act as a selective force in wild boar populations, and suggest that viral disease mortality is mediated by host heterozygosity. In conclusion, this work demonstrates that various types of factors (demographic, environmental and genetic) can significantly influence disease dynamics in free-living wildlife populations.

In preparation for submission

Introduction

Zoonotic and emerging diseases are an important concern for public health and biodiversity [3, 101]. Much attention is given to related safety precautions [5], but relatively little progress has been made in recent decades to increase the fundamental understanding of the natural source of zoonotic and emerging diseases: natural wildlife disease dynamics. This is not without reason, as wildlife disease dynamics are illusive due to challenges in sampling, diagnostics and a myriad of potentially relevant factors [7]. However, in order to manage disease risks and outbreaks effectively, we need to identify the factors that influence wildlife disease transmission and prevalence.

Current theory on disease dynamics mostly stems from mathematical epidemiological modelling [15, 24]. These models rely on a number of generally accepted assumptions, including equal susceptibility to infection for each individual, equal disease related mortality risks and homogeneous-mixing movement/contact patterns within populations, among others. Most assumptions have hardly been evaluated empirically, and even then only averaged parameters for transmission rates are estimated from empirical prevalence series [24, 102]. A thorough mechanistic understanding of transmission rates and the factors that influence it is lacking for most study systems. It is therefore a key issue in wildlife epidemiology to assess the relative importance of different factors on disease prevalence and to understand the mechanisms behind these driving forces. Relevant factors include host demography, host genetic background, pathogen genetic background and various environmental factors.

This study focusses on a relatively simple case of a single wildlife host - single pathogen relationship. European wild boar (*Sus scrofa*) are an interesting wildlife host study species because they are abundant and harbour many diseases [103]. The pathogen Porcine Circovirus type 2 (PCV2) was selected for its characteristics of host-specificity and direct transmission. In addition, it is a relatively common pathogen in European wild boar populations [32], which facilitates statistical analysis. Finally, because PCV2 has an economic impact on the pig breeding industry [104], diagnostic tests were available.

PCV2 is a small non-enveloped single-strand DNA virus that is associated with postweaning multisystemic wasting syndrome (PMWS) also referred to as porcine circovirus associated disease (PCVAD) in swine, causing

diarrhoea, heavy breathing and cell lesions in various organs, most commonly lungs and lymph nodes [105]. This virus depends completely on the polymerase activity of host cells for replication. The main cell types infected by PCV2 are immune cells such as macrophages, which accumulate PCV2 through phagocytosis of pathogens or infected cells. These cells have a low replication rate, and PCV2 has been found to survive phagocytosis and persist over long time-frames in macrophages without significant levels of reproduction [106]. This intracellular hiding behaviour of PCV2 has most likely evolved to evade the temporary antibody response following initial infection, and allows the virus to be silently transported throughout the body of the host to replicate at a later time. Studies have shown that co-infection with other diseases (porcine parvovirus, porcine respiratory and reproductive syndrome virus) may induce immune cell replication and thereby activate PCV2 replication to reactivate the sub-clinical PCV2 infection to full PMWS [107-109]. PCV2 is considered to be a non-lethal pathogen in the pig industry, but it has been suggested to play a role in piglet mortality in high-density wild boar populations [33]. Not all pigs develop PMWS after PCV2 exposure. Individual symptomatic differences are ascribed to genetic heterogeneity [110, 111]. Genetic characterisation of the PCV2 strains that circulate in domestic pigs and wild boar suggested that the PCV2 infection dynamics in these two populations are largely independent, with some level of transmission from domestic to wild [112].

The relative importance of host genetic factors in wildlife disease dynamics is largely unknown [34]. However, genetic heterogeneity in a wildlife host species can influence immune capacity and thereby affect individual infection risk and survival [113-115]. Acevedo-Whitehouse *et al*. [116] demonstrated that genetic heterozygosity is an important predictor of resistance to and suppression of bovine tuberculosis bacteria in free-living wild boar. High host genetic diversity, usually measured by the observed individual level of heterozygosity, is considered to increase resistance to infectious diseases [117]. Heterozygosity effects are mainly tested using heterozygote-fitness correlations (HFCs). There are two mechanistic hypotheses explaining the presence of HFCs. The local effect hypothesis states that the genotype of specific genes is affecting fitness and that this HFC extends to other nearby loci through Linkage Disequilibrium (LD). The global effect hypothesis assumes widespread

deleterious effects from inbreeding or outbreeding depression. HFCs associated with disease prevalence are often assumed to display a negative relationship, where high host heterozygosity is associated with good host immune defence and thus lower disease prevalence (due to a better capacity to prevent pathogens from entering the host system and developing an infection). Theoretical explanations include inbreeding depression (deleterious recessive effects in relatively homozygous individuals) or heterozygote advantage (the heterozygous genotype has higher relative fitness than both homozygous genotypes).

Methods

We randomly collected 462 wild boar blood samples from routine disease monitoring programs in the Netherlands and adjacent parts of Western Germany (Northrhine-Westphalia and Rhineland-Palatinate) during the years 2008-2010. These disease monitoring programs sampled culled animals from regular hunting practices. A small sampling bias may possibly have arisen during hunting, but this was considered sufficiently minor to avoid undesired effects. These samples were genotyped with a 351 SNP assay (chapter 3) derived from the Illumina SNP60 beadchip genotyping assay [43]. Serology was performed using the commercially available SERELISA PCV2 Ab Mono Blocking kit from Synbiotics®. ELISA end-product optical densities of each sample (duplo average, S) were compared to the optical density of the negative control (duplo average, N), giving ratio S/N. A ratio of ≤ 0.40 is considered positive for PCV2 antibodies according to the manufacturer's specifications, while a ratio >0.40 is considered negative. This identified whether significant levels of IgM and IgG antibodies against PCV2 antigens were present. IgM and IgG antibodies have a halflife in pigs of 3.5 and 14 days respectively [118]. Significant PCV2 antibody titres are therefore assumed to represent recent infection (approximately within 1 month before sampling). Age class (juvenile <1 year old, yearling 1-2 years old, and adult >2 years old), year of sampling and month of sampling were recorded by disease monitoring programs in the field. Age was estimated based on body size, coat pattern and dentition. However, age information was missing for 26 individuals. Sex was determined genetically based on homozygosity or heterozygosity of X-chromosomal loci. Genetic population assignment (chapter 3), domestic pig hybrid genetic ancestry (identifiable until 5^{th} generation

backcrosses with wild boar, chapter 2) and observed individual heterozygosity were included in a multiple logistic generalized linear regression analysis in R, taking into account the missing age data of 26 individuals. Fisher's exact tests and logistic linear regression were additionally employed to assess class effects. Pairwise heterozygosity correlation chi-square tests between SNP loci were performed in Excel to assess if heterozygosity effects were global or local effect. In these chi-square tests observed frequencies of pairwise heterozygosity combinations were compared to expected frequencies of pairwise heterozygosity combinations based on allele frequencies. Results from the pairwise correlations were evaluated using a false discovery rate-corrected combined probability test.

Results

In a multiple logistic generalized linear regression analysis, PCV2 antibody status was affected by age class (p=0.005), year of sampling (p=0.010), and genetic heterozygosity (p=0.002, Table 4.1). Both forward and backward selection resulted in the same model. All of the included factors were also significant (p<0.05) in a single-factor logistic regression analysis. Non-significant factors (p>0.05) included sex, population assignment, month of sampling and hybrid status.

Table 4.1 Multiple logistic generalised linear regression model of individual wild boar PCV2 antibody status. Forward and backward selection resulted in the same model. Consistently non-significant factors included sex, hybrid status, month of sampling and population assignment.

factor	coefficient	s.e.	z-value	*p*-value
intercept	1271	418.4	3.039	0.0024
age	0.511	0.181	2.827	0.0023
heterozygosity	7.176	2.368	3.031	0.0024
year	-0.635	0.208	-3.050	0.0047

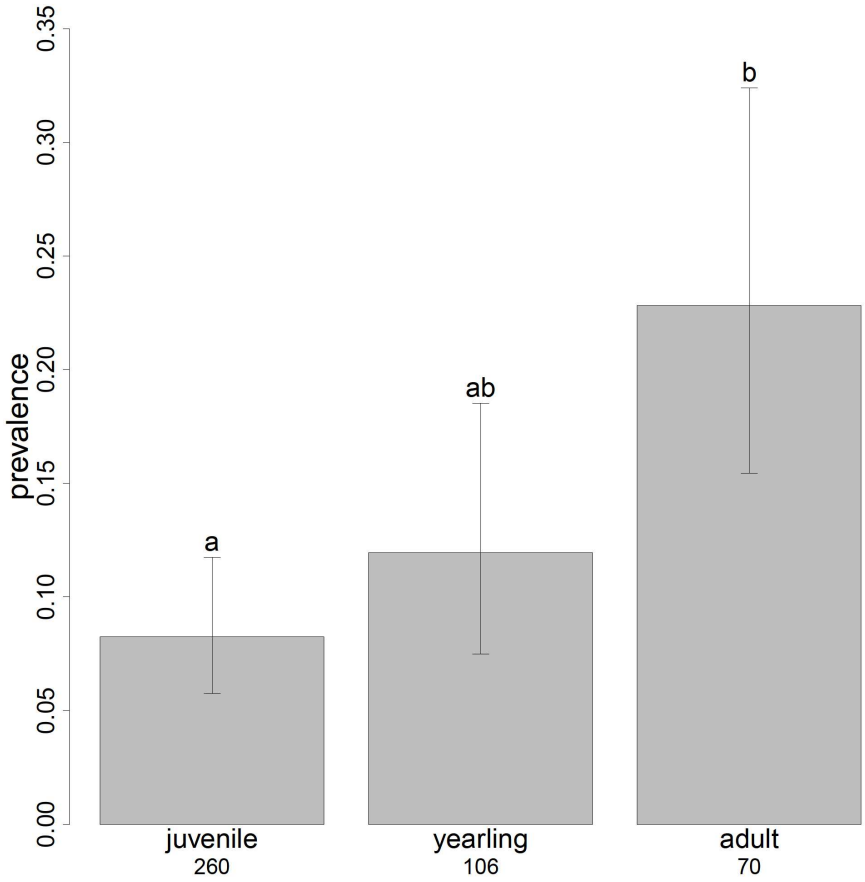


Figure 4.1 PCV2 antibody prevalence of each wild boar age class. Error bars indicate Wilson Score 95% Confidence Intervals. Sample sizes per age class are indicated along the x-axis. Fisher Exact Tests indicated significant differences

between age classes (p=0.001). Letters above the confidence intervals indicate significant group differences.

A Fisher's exact test of age class differences in PCV2 prevalence showed a positive relationship, with the adult life stage displaying the highest prevalence (Figure 4.1). Between year differences in PCV2 prevalence indicated that the year 2008 had a significantly higher PCV2 prevalence than the years 2009 and 2010 (Figure 4.2).

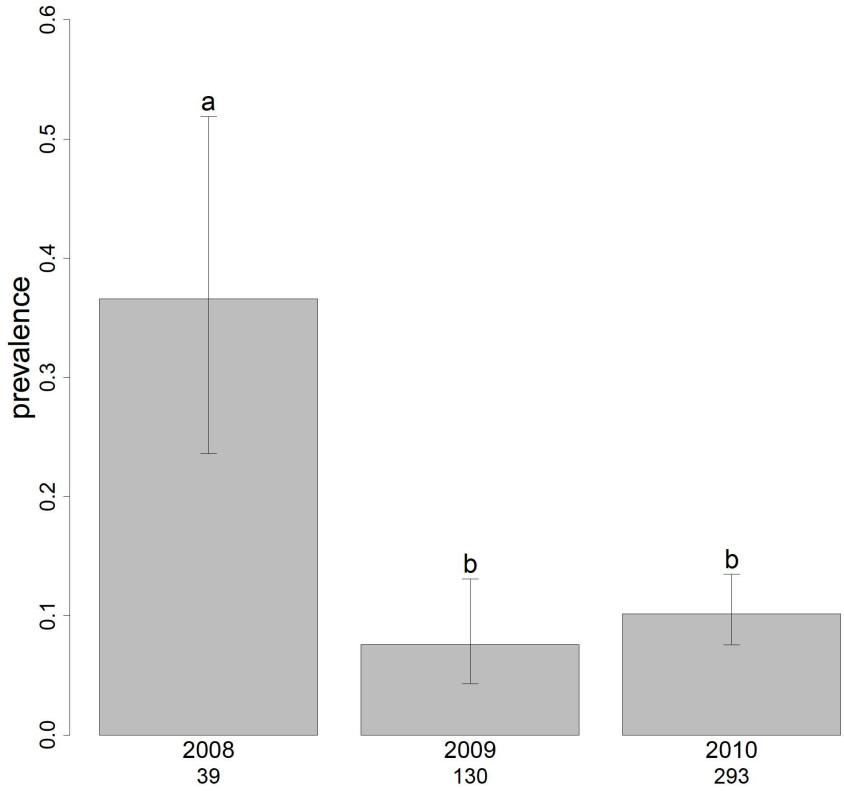


Figure 4.2 PCV2 antibody prevalence per sampling year (2008-2010). Error bars indicate Wilson Score 95% Confidence Intervals. Sample sizes per year are indicated along the x-axis. Fisher Exact Tests indicated significant differences between years (p=0.001). Letters above the confidence intervals indicate significant group differences.

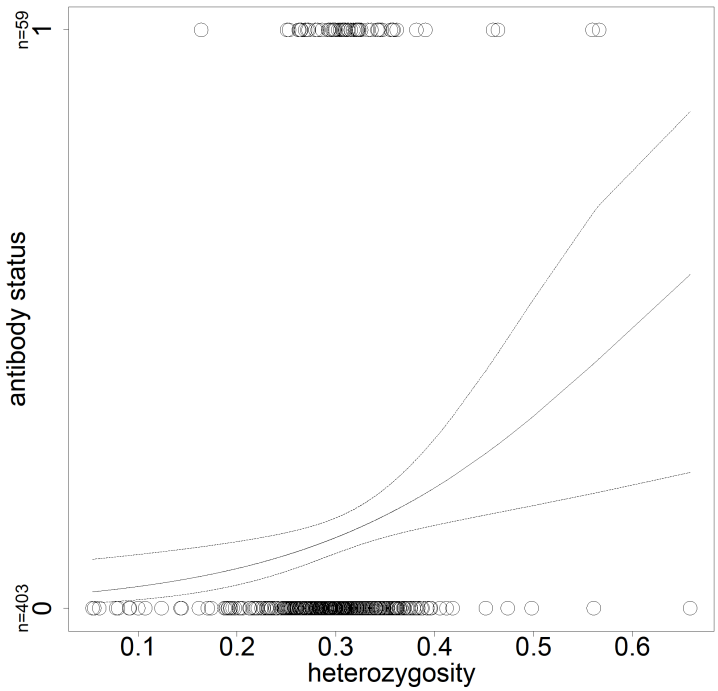


Figure 4.3 PCV2 antibody status (1=presence of antibodies, 0=absence) versus genetic heterozygosity in wild boar. Sample sizes per group are indicated along the y-axis. The solid curve shows the result of a single-factor logistic generalized linear regression between PCV2 antibody status and individual heterozygosity, y=exp(-3.883=6.368*x)/(1+exp(-3.883+6.368*x)), with a *p*-value of 0.004 and a Nagelkerke R-square of 0.035. The dotted lines indicate 95% confidence intervals.

The logistic regression curve of individual PCV2 antibody status versus heterozygosity indicates a positive relationship between PCV2 prevalence and wild boar heterozygosity (Figure 4.3). This significant positive relationship between individual genetic heterozygosity and presence of PCV2 antibodies was present in juveniles and yearlings, but not in adults according to separate logistic regression (Table 4.2 and Figure 4.4).

Table 4.2 Single-factor logistic generalized linear regression results for the effect of heterozygosity on PCV2 antibody presence in the different age classes.

age class	factor	coefficient	s.e.	z-value	*p*-value
juveniles	intercept	-4.617	0.966	-4.778	1.77e-06
	heterozygosity	7.631	2.911	2.621	0.009
yearlings	Intercept	-7.182	2.147	-3.345	0.001
	heterozygosity	17.361	6.641	2.614	0.009
adults	intercept	0.303	1.803	0.168	0.867
	heterozygosity	-5.005	5.928	-0.844	0.399

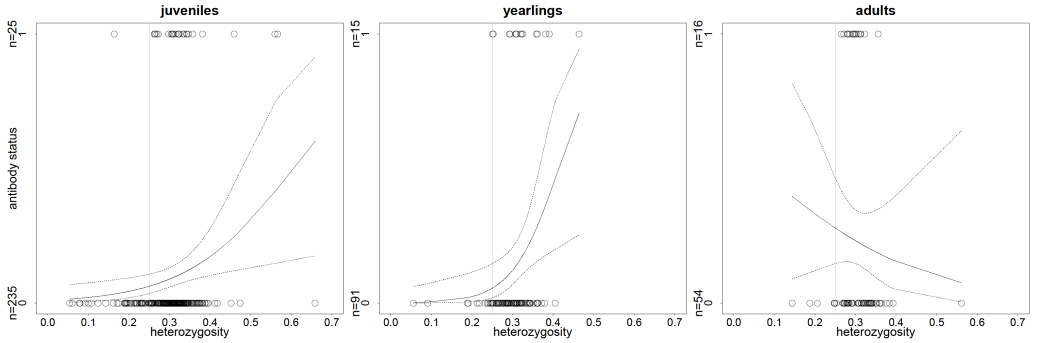


Figure 4.4 PCV2 antibody status (1=presence of antibodies, 0=absence) versus genetic heterozygosity for the different wild boar age classes. Sample sizes per group are indicated along the y-axis. The solid curves show results from single-factor logistic generalized linear regression analyses between PCV2 antibody status and individual heterozygosity (Table 4.2), which were significant for juveniles (p=0.009) and yearlings (p=0.009), but not for adults (p=0.399). Nagelkerke R-square values are 0.057, 0.133 and 0.016 respectively. The dotted lines indicate 95% confidence intervals. Grey vertical lines indicate the 0.25 threshold used for the random sampling analysis of selective disappearance of low-heterozygosity individuals.

The chance to draw only 1 or less juveniles with a level of heterozygosity below 0.25 (see Figure 4.4), based on random sampling (n=25 individuals) from the juveniles without antibodies (n=235) was 396 out of 9999 (p=0.040). The chance to draw 0 yearlings with a level of heterozygosity below 0.25, based on random sampling (n=15) from the yearlings without antibodies (n=91) was 963

out of 9999 (p=0.096). This means that there were significantly less low-heterozygosity (<0.25) juveniles with a PCV2 infection. This effect was near-significant in yearlings, and absent in adults (p=0.150) .

Pairwise correlation chi-square tests of heterozygosity between loci indicated that heterozygosity was correlated across the entire genome more often than expected by chance (fraction of significantly heterozygosity correlated locus-pairs=0.10, average pairwise chi-square=0.59, combined probability test p<0.0001, the false discovery rate corrected alpha for the combined probability test was alpha=0.027). This finding points to a global HFC for PCV2 antibody prevalence, and suggests that many mildly deleterious effects across the genome rather than a few strongly deleterious effects at specific loci are responsible for the effect of individual heterozygosity on PCV2 prevalence [119].

Discussion

The effect of age class on PCV2 prevalence in this study (Table 4.1 and Figure 4.1) is linked to the infection mechanism of the pathogen. After primary infection, PCV2 is known to persist silently in infected macrophage cells [106]. Accumulated exposure to PCV2 over a wild boar's life time leads to accumulated presence of silent intracellular PCV2 pathogens with age. These silent viruses can be activated by co-infection [107, 109] or possibly other signals, leading to an increasing frequency of active PCV2 infection with age.

The between year differences in PCV2 prevalence in this study (Table 4.1 and Figure 4.2) are most likely caused by climatic factors. The European winter of 2007-2008 was average to mild, with mean January temperatures of 4^{o}C and a fair amount of rainfall in the study area (from the E-OBS dataset, www.ecad.eu/eobs May 2013). The 2008 PCV2 prevalence found in this study (23.6-51.9%) corresponds reasonably well to independent reports of November-January 2008-2009 PCV2 prevalence (50.7%) from Southern Germany [120], but the 2009-2010 PCV2 prevalence data from this study are remarkably low in comparison. The two subsequent winters in the study area were unusually cold with 2009 mean January temperatures in the study area of -3^{o}C and 2010 mean January temperatures of -5^{o}C, and heavy snowfall. Sub-zero temperatures can be expected to affect the transmission of PCV2 by aerosols negatively due to desiccation of the virus, which has been shown for other aerosol-transmitted

pathogens [121, 122]. In addition, survivability of actively infected (infective) hosts may be lower in harsh winters, further decreasing transmission rates.

The positive association between heterozygosity and PCV2 antibody prevalence (Table 4.1 and Figure 4.3) is most likely caused by selective disappearance of low-heterozygosity individuals, as shown by the resampling analysis. This selection effect is most evident in juveniles (and a proportion of yearlings, see also Figure 4.4). The mortality of low-heterozygosity individuals may be caused by a direct selection effect of PCV2 infection, or it may be caused by lethal effects of inbreeding unrelated to PCV2 infection. In the latter case PCV2 infection is not the cause of death, and the observed patterns in PCV2 prevalence represent an independent effect of inbreeding. This latter explanation suggests a very strong inbreeding depression in wild boar leading to mortality through an unknown mechanism. No other indications for such a strong inbreeding depression were found, and we consider a direct selection effect of PCV2 infection on mortality to be the most likely explanation of the data. In this interpretation, infection by the PCV2 pathogen is the direct causative agent of mortality. This form of natural selection is however mediated by heterozygosity, because only the most inbred individuals suffer actual mortality upon PCV2 infection, while others survive (but probably pay some energetic cost to deal with the infection). The mortality effect is most pronounced upon first exposure to the PCV2 (i.e., in the case of a clear pre- to post-selection contrast). Adult individuals (>2 years old) and some yearlings may already have been exposed to PCV2 earlier in life, and therefore represent post-selection individuals, even if they do not currently display significant PCV2 antibody levels. These older post-selection animals do not show a significant relationship between heterozygosity and PCV2 prevalence (Table 4.2 and Figure 4.4), which suggests that disease mortality is less frequent at later life stages.

If selection pressures are stable over time, one would expect a significant increase of heterozygosity with age. However, a general linear model analysis showed no significant relationship between age and heterozygosity (p-value=0.890). The lack of such a relationship may be due to between year differences in selection pressure associated with winter conditions or other obscuring factors. Such age effects can then only be reliably demonstrated using cohort studies. The current study was however not set up for this approach.

The pairwise correlation of heterozygosity between loci suggests that in this case a global HFC for PCV2 antibody prevalence is more likely than a local HFC. If this is interpretation correct, a combination of a global mechanism and a HFC would further suggest that many mildly deleterious effects across the genome rather than a few strongly deleterious effects at specific loci are responsible for the effect of individual heterozygosity on PCV2 prevalence [122]. This supports the interpretation of inbreeding depression, as the most homozygous individuals are selectively disappearing (Figure 4.4). The global effect positive HFC with PCV2 prevalence found in this study does not necessarily indicate that the wild boar are suffering from inbreeding depression for other traits as well. Disease resistance is a trait under relatively strong selection pressure, and may show selection effects before other traits.

In conclusion, our results show that different types of factors influence PCV2 prevalence: demographic (age class), environmental (climatic conditions) and genetic (heterozygosity). A global HFC caused by selective disappearance of low-heterozygosity individuals suggests inbreeding depression in the context of PCV2 disease resistance. Epidemiological models of wildlife disease dynamics that are based solely on demographic and spatial parameters may underestimate the influence of genetic heterogeneity and population genetic dynamics.

Acknowledgements

We thank Mike Nieuwland and Ger de Vries-Reilingh for technical assistance. Thanks go to the Animal Health Service Centre Deventer (GDD) and Karl Zimmer from the Landesuntersuchungsamt Rheinland-pfalz for sample contributions. Finally, we thank the Royal Dutch Hunters Association (KNJV) for financial support.

Data accessibility

The 351 SNP genotypes (plink format) of all 462 wild boar and the 120 domestic pig sample, as well as a tab delimited text file containing information on individual PCV2 antibody status, hybrid status, year of sampling, month of sampling, age, sex, observed heterozygosity and population assignment of all 462 wild boar are available upon request from danielgoedbloed@hotmail.com.

Chapter 5

Elevated *Mycoplasma hyopneumoniae* infection rates in European wild boar with partial domestic ancestry

Daniel J Goedbloed, Pim van Hooft, Walburga Lutz, Thijs Bosch, Sip E van Wieren, Ron C Ydenberg, Herbert HT Prins

Abstract

Wildlife form a natural reservoir for a range of potentially zoonotic diseases. Genetic components are known to determine individual immune capacity and thereby affect the risk of disease outbreaks in wildlife populations. Such immune-related genetic components are subject to natural selection exerted by pathogens in an assumed evolutionary arms race known as the 'Red Queen Hypothesis'. However, domestic animals are often protected by veterinary care, which relaxes the pathogen-driven selection pressure maintaining genetic immune functions. Since the 1940s veterinary practice has included large scale application of antibiotics in livestock. Such a degree and timescale of veterinary protection against pathogens, particularly bacteria, can be expected to lead to a reduced immune capacity in domestic animals compared to wildlife, but evidence from literature is lacking. Effects of genetic introgression from domestic animals on wildlife disease risk have never been demonstrated. Here we show significantly increased disease antibody prevalence in a domestic pig hybrid subgroup of free-living European wild boar. A fraction (4.3%) of 463 opportunistically sampled wild boar from the Netherlands and Western parts of Germany was identified as genetic wild-domestic hybrids, based on population assignment analysis using a 351 Single Nucleotide Polymorphism (SNP) assay. Antibody prevalence against the bacterial pathogen *Mycoplasma hyopneumoniae* was significantly higher in hybrids than in non-hybrid wild boar, based on a multiple logistic generalized linear regression. This finding demonstrates

increased infection rates in a wildlife subgroup with partial domestic ancestry and indicates that genetic introgression from domestic to wild can increase the risk of disease outbreaks in wildlife populations. These results provide a novel argument in support of the 'Red Queen Hypothesis' of host-pathogen coevolution, as it suggests a significant effect of relaxed pathogen selection pressure in domestic pigs. Finally, the detrimental effects of veterinary protection on immune capacity in captive-reared stocks raise concern for disease risks of wildlife restocking activities.

In preparation for submission

Introduction

Wildlife diseases have received much attention in recent decades [3, 101]. They are considered to be a reservoir for potentially zoonotic diseases and an important risk factor for biodiversity conservation, livestock health and public health [9, 10]. Most epizootic outbreaks and emerging diseases are due to spill-over from domestic stocks to wildlife, rather than vice-versa [3, 112]. There is however much public concern for the risk of spill-back from wildlife reservoirs, because of economic interests [11, 123].

High animal densities in captive conditions are a fertile ground for many pathogens. This is demonstrated for example by findings of higher wild boar disease prevalence in fenced, high-density, farm-like hunting estates compared to more natural conditions [32, 33]. The increased disease risks of high-density captive breeding are often partially countered by providing veterinary care, including use of antibiotics. Antibiotics have been used since the 1940s curatively, preventively and (until the 1970s) as a growth-promoter on a large scale in livestock [124]. This time period equates to approximately 36 generations for wild boar, which have an average generation time of 2 years [125]. This relatively fast generation time is due to an early average age of first reproduction (8-9 months) and a large contribution of juveniles to the total reproduction of a population (adults older than 2 years usually make up only 20-25% of a population [35]). Domestic farrowing sows have a generation time of 1.8 years [126], leading to a period of 40 generations in domestic pigs. Such a period is limited and may be too restrictive for deleterious alleles to increase in the population through genetic drift. On the other hand, the small population size of purebred pig lines may enhance the rate of genetic drift, facilitating loss of immune functions. While the use of antibiotics has received much attention with regard to antibiotic-resistance in pathogens [127], little is known about the evolutionary effects on the immune capacity of the host.

Genetic factors determine individual immune capacity and thereby risk of disease infection in wildlife populations [13, 128]. Such genetic factors are subject to natural selection exerted by pathogens. Host-pathogen interactions in an evolutionary context are viewed as a continuous arms race, known as the 'Red Queen Hypothesis' [129-131]. However, long-term exposure to domestic conditions including use of antibiotics and further veterinary support may relax

or even relieve the selection pressure driving the evolution of a host species' immune functions. In that case, the immune capacity of the domestic host population to independently deal with these pathogens may become compromised.

This study concerns the relatively simple case of a single wildlife host and a single host-specific pathogen. We investigated European wild boar (*Sus scrofa*), which is known to host many diseases that are shared with the closely related domestic pig [103, 123]. We focussed on the respiratory disease Porcine Enzootic Pneumonia caused by the bacterial pathogen *Mycoplasma hyopneumoniae* (*Mhyo*) because of its characteristics of host-specificity and direct transmission. Porcine Enzootic Pneumonia prevalence can reach up to 50% in commercial slaughter pigs [132]. *Mhyo* is a contagious pathogen that primarily infects lung epithelium where it induces coughing, cell lesions and a reduced response of phagocytes and lymphocytes [133, 134]. Transmission of *Mhyo* in domestic pigs during contact experiments occurs rapidly, with symptoms and infectious stage presenting on average at 15 days (range 5-28) after exposure [135]. Airborne transmission through aerosols over long distances has also been demonstrated for *Mhyo* [136]. The immune response to *Mhyo* varies considerably between individual domestic pigs [137], and current control measures such as vaccination or medication supress symptoms but do not lift the infection entirely [138]. Outbreaks of *Mhyo* in domestic pig farms have been shown to display seasonality, because the survival of the pathogen outside the host is dependent on air humidity and temperature [121, 122, 139].

Northwest European wild boar populations have been demonstrated to contain approximately 4-10% wild-domestic hybrids, most likely due to release of wild-domestic hybrids from a captive farmed stock (chapter 2 and 3). Farmed wild-domestic hybrids have a wildtype appearance and are bred to supply a commercial demand in exclusive wild boar meat. Previously, recent hybrids were identified with levels of domestic genetic introgression up to five generations of backcrossing to wild boar, as well as advanced-generation hybrids produced by repeated reproduction among hybrids (chapter 2). This provides an opportunity to assess the effect of partial domestic ancestry (and thus a history of genetic adaptation to captivity) on disease resistance in free-living wild boar.

Methods

We randomly collected 462 wild boar blood samples from the Netherlands and adjacent parts of Western Germany (Northrhine-Westphalia and Rhineland-Palatinate) from routine disease monitoring programs in the years 2008-2010 targeting culled animals from regular hunting practices. This dataset corresponds to that of chapter 4 and includes 34 animals analysed previously for genetic introgression from domestic pigs using a 60k Single Nucleotide Polymorphism (SNP) assay (chapter 2). Age class (juvenile <1 year old, yearling 1-2 years old, and adult >2 years old), year of sampling and month of sampling were recorded in the field. Age was estimated based on body size, coat pattern and dentition. Age information was however missing for 26 individuals. Sex was determined genetically based on homozygosity or heterozygosity of X-chromosomal loci. All samples were genotyped using a 351 SNP assay and compared to the genotypes of a domestic pig sample (n=120, breeds: Large white, Landrace, Duroc, Pietrain, British saddleback and Tamworth,). We performed Principal Component Analysis (PCA) using the eigenvector method as implemented in Eigensoft 3.0 [61, 62]. In addition, we performed a population assignment analysis using STRUCTURE 2.3.1 [63] based on 10 runs using two clusters (K=2) at 500,000 iterations after a burn in of 800,000. The population assignment resulted in a clear wild-domestic separation. Recent domestic ancestry was identified using a STRUCTURE population assignment score threshold of 0.25 to the domestic pig cluster, as this value matched the results of chapter 2. We additionally screened all samples for antibodies against the bacterial pathogen Mycoplasma hyopneumoniae (Mhyo), using an in-house sandwich ELISA test based on Intervet® antigens purified by ammonium sulphate precipitation. Optical densities of ELISA end-products for each sample (duplo average, S) were compared to the optical density of the negative control (duplo average, N), giving ratio S/N. A ratio of ≤ 0.50 was considered negative for PCV2 antibodies, while a ratio >0.50 was considered positive. This threshold was based on the observed boundary between two peaks in the measured optical densities in this dataset. Multiple logistic generalised linear regression was performed in R to assess the influence of: sex, age class, year of sampling, month of sampling, genetic heterozygosity, genetic population assignment (based on the clustering

identified by the program STRUCTURE) and genetic hybrid status on Mhyo disease antibody prevalence.

Results

Genetic analysis with a 351 SNP assay showed a clear separation between wild boar and domestic pigs in a Principal Components Analysis (PCA, Figure 5.1). This separation was supported by STRUCTURE population assignment analyses (*K*=2), with some free-living wild boar displaying admixture with domestic pigs (Figure 5.2). To identify wild-domestic hybrids in our wild boar dataset, a minimum STRUCTURE population assignment proportion of 0.25 to a sample of domestic pigs (*n*=120) was applied (chapter 3). This threshold was chosen because it identified all hybrids with recent domestic ancestry (up to five generations ago) among 34 samples previously analysed with a more detailed 60k SNP assay (chapter 2) without producing any false positives. In total, 20 hybrids with recent domestic ancestry were identified this way, which corresponded to the most intermediate individuals in the PCA (Figure 5.1).

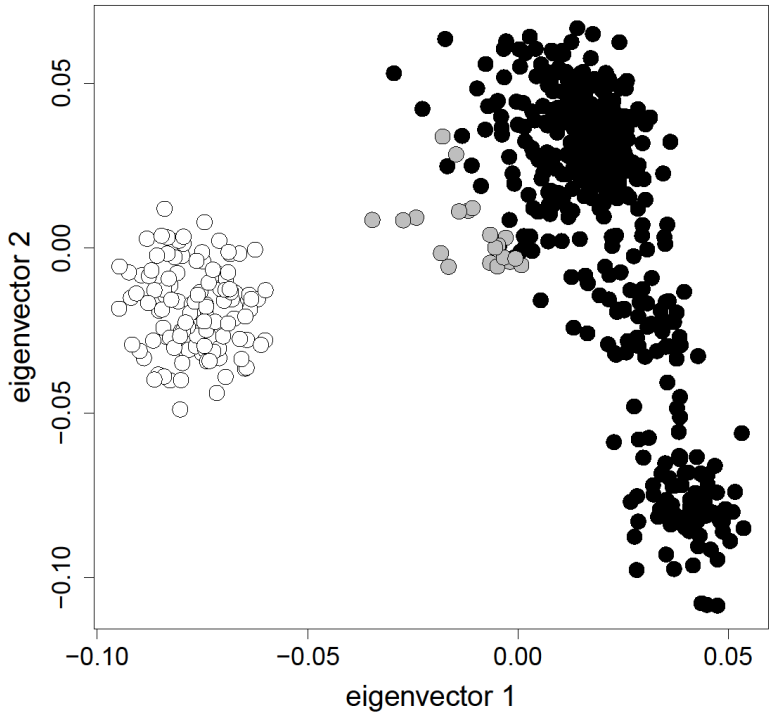


Figure 5.1 PCA plot of 351 SNP genetic diversity. Wild boar are indicated in black, domestic pigs in white, and the 20 identified hybrids in grey. Eigenvectors 1 and 2 explain 39.5% of variation in the dataset.

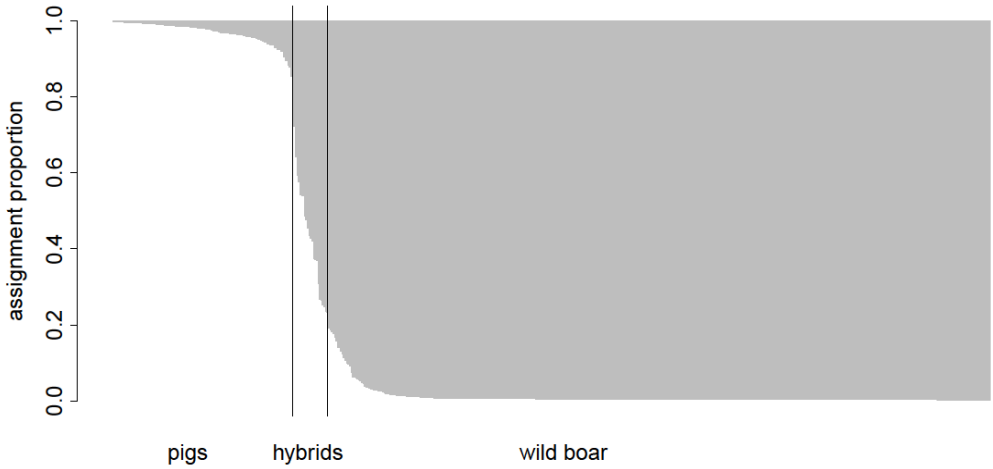


Figure 5.2 Genetic population assignment proportions. Proportions belonging to the wild boar cluster are indicated in grey and domestic pig proportions are in white. Vertical lines indicate the boundaries between the domestic pig sample, the identified hybrids and the remaining wildtype wild boar.

Strikingly, the identified hybrids with recent domestic ancestry displayed a higher *Mhyo* antibody prevalence than wildtype wild boar (Figure 5.3) indicating increased infection rates in this subgroup. Single-factor logistic regression analysis revealed a near-significant effect of hybrid status on *Mhyo* antibody prevalence in wild boar (p=0.055), while a multiple logistic generalised linear regression indicated significance (p=0.035, Table 5.1). The multiple logistic generalised linear regression model also accounted for temporal (seasonal) variation in disease prevalence by including the factors year of sampling and month of sampling. The between year differences (Figure 5.4) match those found for Porcine Circovirus type 2, reported in chapter 4 and point to a shared mechanism for these respiratory diseases. The seasonal variation of *Mhyo* prevalence in wild boar corresponds to findings of seasonality in *Mhyo* infection of domestic pigs [121, 139], which is attributed to climatic influences on pathogen transmission. Both forward and backward selection resulted in the same multiple logistic generalised linear regression model, and consistently excluded age class, sex, genetic heterozygosity and STRUCTURE population assignment (chapter 3) as significant factors.

No correlation was found between overall genetic heterozygosity and *Mhyo* antibody prevalence. This indicates that general effects of inbreeding or outbreeding did not affect immune capacity with regard to *Mhyo* in the wild-domestic hybrid wild boar.

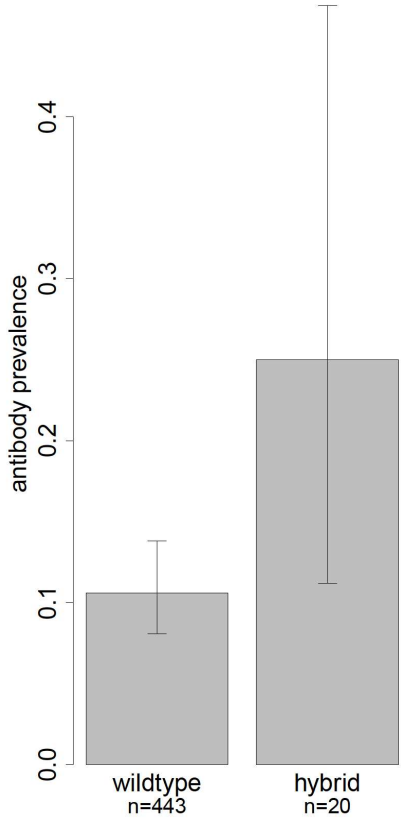


Figure 5.3 *Mycoplasma hyopneumoniae* (*Mhyo*) antibody prevalence in wildtype wild boar and hybrid wild boar with recent domestic ancestry (up to five generations ago). Error bars indicate Wilson Score 95% Confidence Intervals. Sample sizes for wildtype wild boar and domestic hybrids are given along the x-axis. Multiple logistic generalised linear regression indicated a significantly raised *Mhyo* antibody prevalence in domestic hybrids (p=0.035).

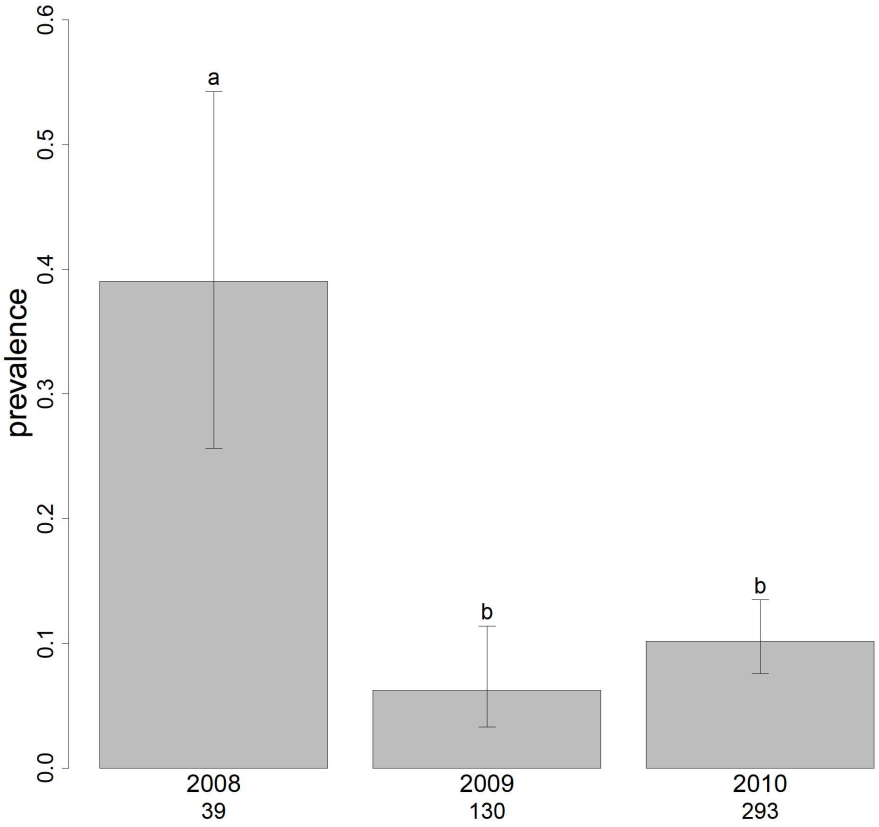


Figure 5.4 *Mycoplasma hyopneumoniae* (*Mhyo*) antibody prevalence per sampling year (2008-2010). Error bars indicate Wilson Score 95% Confidence Intervals. Sample sizes per year are indicated along the x-axis. Fisher Exact Tests indicated significant differences between years (p=0.001). Letters above the confidence intervals indicate significant group differences.

Table 5.1 Multiple logistic generalised linear regression model of individual wild boar *Mycoplasma hyopneumoniae* antibody status. Forward and backward selection resulted in the same model. Consistently non-significant factors included age class (juvenile, yearling, adult), sex, genetic heterozygosity and population assignment.

factor	coefficient	s.e.	z-value	*p*-value
intercept	1111	411.0	2.703	0.0069
year	-0.554	0.205	-2.707	0.0068
month	-0.080	0.034	-2.327	0.0200
hybrid	1.243	0.590	2.107	0.0351

Discussion

The mechanism behind the observed higher antibody prevalence against *Mhyo* in recent wild-domestic hybrids compared to wildtype wild boar could not be established with certainty. One explanation for this finding, would be a direct release effect. The identified hybrids could theoretically be released individuals. This means that they themselves would have a farm history. *Mhyo* is endemic and widespread in the wild, but captive farm-like conditions are associated with higher disease prevalence [32, 33]. This suggests that the increased *Mhyo* antibody prevalence in this subgroup could be a direct consequence of their farm history. However, *Mhyo* is a rapidly spreading pathogen with airborne transmission and development of an infectious stage in the recipient occurring within a few days [135, 136]. If a recent release of captive hybrids indeed caused *Mhyo* prevalence to increase locally in parts of the study area, this would be a temporary effect. Within days or weeks a large number of wildtype wild boar in those parts of the study area would be exposed as well and possibly infected. The observed difference in *Mhyo* antibody prevalence between the hybrid subgroup and the wildtype wild boar is therefore unlikely to be caused by direct release effects.

Another explanation for the observations presented in this study is the inheritance of degraded (or deleterious) domestic immune genes. Veterinary protection using antibiotics in domestic pigs and farmed wild boar can relax the pathogen selection pressure maintaining immune gene variation and may lead to a form of degradation of these immune genes.

Previous analysis of the same wild boar dataset for PCV2 antibodies showed no association with hybrid status (chapter 4). This supports a possible role of antibiotics in the observed immune degradation, as antibiotics protect the host against bacterial pathogens (e.g., *Mhyo*), but not viruses (e.g., PCV2). This leads to the hypothesis that only those immune genes specifically involved in resistance against bacterial diseases are degraded in domestic pigs compared to wild boar.

The wild-domestic hybrids in this study included animals resulting from up to five generations of backcrossing with wild boar (chapter 2). The fact that we detected a significant increase of *Mhyo* antibody prevalence in this hybrid group, even after a small number of generations of backcrossing to wild, indicates that the effect of domestic ancestry on immune capacity versus *Mhyo* is persistent over generations.

In conclusion, we find a significantly higher *Mhyo* antibody prevalence in a subgroup of free-living Northwest European wild boar with recent partial domestic ancestry. This finding suggests a deleterious effect of domestic ancestry on disease resistance in a wild context, arguably caused by evolutionary effects of veterinary protection and the use of antibiotics in domestic and captive stocks. If this interpretation is correct, these findings support the 'Red Queen Hypothesis' of a continuous evolutionary arms-race between hosts and pathogens, in the sense that a history of protection from pathogen selection pressure leads to degraded immune functions and therefore increased susceptibility to these pathogens compared to individuals with a continuous history of pathogen selection pressure. The finding that partial domestic ancestry increases disease risk in wild boar has implications for wildlife restocking projects and activities, where the consequences and risks for disease outbreak may be more far-reaching than previously thought if captive-reared stocks are used.

Acknowledgements

We thank Mike Nieuwland and Ger de Vries-Reilingh for technical assistance. Thanks go to Maarten Witvliet from Intervet® for contributing Mycoplasma hyopneumoniae antigen, and to the Animal Health Service Centre Deventer (GDD) and Karl Zimmer from the Landesuntersuchungsamt Rheinland-pfalz for

sample contributions. Finally, we thank the Royal Dutch Hunters Association (KNJV) for financial support.

Data accessibility

The 351 SNP genotypes (plink format) of all 462 wild boar and the 120 domestic pig sample, as well as a tab delimited text file containing information on individual *Mhyo* antibody status, hybrid status, year of sampling, month of sampling, age, sex, observed heterozygosity and population assignment of all 462 wild boar are available upon request from danielgoedbloed@hotmail.com.

Chapter 6

Synthesis

The occurrence and consequences of domestic pig hybrids in wild boar populations

Finding clear evidence of recent genetic introgression from domestic pigs in wild boar in the study area (chapter 2) was rather unexpected. Pig farming in Western Europe has heavily intensified in recent history and occurs indoors, with very limited opportunity for interaction between pigs and wild boar. The most likely route of introgression, through purposeful wild-domestic hybridisation in farmed wild boar and subsequent restocking of wild populations, indicated that undocumented events can have important consequences for population genetic processes. The results in chapters 2 and 3 show that the use of farmed wild boar in reintroductions or restocking activities is more common than previously thought and may have consequences for population differentiation and the genetic diversity of wild boar populations.

The evolutionary consequences of genetic introgression from domestic animals to their wild counterparts are largely unknown, but see [140]. Centuries of artificial selection in pigs has skewed phenotypic and genotypic variation in certain traits to the extreme (mostly in traits related to appearance, reproduction and growth rate). In addition, artificial selection and domestic conditions can be assumed to have caused degradation of traits that were no longer maintained by natural selection. Such genetic consequences of a domestic history are expected to mainly have neutral or maladaptive value in a wild context.

The fact that wild boar litter sizes show strong regional differences in Western Germany [73], may partly be a consequence of different levels of genetic introgression from domestic pigs. The effect of increased wild boar litter size on fitness and local adaptation in a natural populations has however not been evaluated. The significant effect of partial domestic ancestry on *Mycoplasma hyopneumoniae* (*Mhyo*) disease prevalence in chapter 5 implies

that genetic introgression from domestic pigs may increase disease susceptibility in free-living wild boar populations. Here, immune functions putatively serve as an example of a trait degraded by a lack of natural selection pressure in domestic pigs. Genetic introgression from domestic pigs is in this case considered to be maladaptive for *Mhyo* disease resistance.

Backcrossing with wild boar was shown in chapter 2 to rapidly reduce proportions of domestic ancestry. In addition, natural selection will remove most maladaptive traits inherited from domestic ancestry over time. Severe long-term evolutionary consequences for wild populations are therefore not likely, although genetic introgression from domestic pigs may increase genetic load in wild boar populations. Because of the long-term evolutionary risks of introgressed deleterious alleles and genetic load are difficult to quantify, I recommend to aim for minimisation of genetic introgression from domestic sources in wildlife conservation and management.

Population genetics and epidemiology are different ballgames

In theory, the distribution and movement patterns of a host species determine the occurrence of obligate host-specific pathogens. To a certain extent this has to be true: where there are no hosts, there can be no pathogens. In addition, transmission between individuals depends on host encounters and thus host movement patterns. However, genetic population structure or gene flow were not found to significantly affect disease prevalence in this thesis (chapter 4 and 5). This may be due to the vastly different temporal scales of their dynamics. The time it takes for a wild boar allele to disperse over a distance of 50 km may be 5 years (if one assumes a slightly optimistic average lifetime dispersal of 10 km per generation [35]). On the other hand, it may take only a few weeks or days for a bacterial or viral disease to travel this distance, as these pathogens will infect multiple new hosts every few days and may travel for kilometres in aerosols under favourable circumstances [135, 136].

Population genetics and disease dynamics not only differ in their temporal scales, the factors that influence their occurrence or spread differ as well. The occurrence and spread of alleles is determined by dispersal and reproduction. Dispersal of terrestrial animals is influenced by geographical distance and landscape features [76]. Reproduction (or mating) mainly depends

on spatial proximity and competition for mating rights. Alleles may further be subject to random genetic drift or natural selection. The rules of the game are quite well understood, and population genetic patterns can be modelled or analysed rather realistically (see chapter 2 and 3).

The rules for the game of disease prevalence are not so clear. The factors that determine occurrence and spread of pathogens, are often unknown and may differ strongly between different pathogens. For PCV2 and *Mhyo*, the only universal factor influencing disease prevalence seems to be climatic, causing between year differences in disease prevalence that are very similar for both pathogens (chapter 4 and 5). This similarity is most likely connected to their shared mechanism of transmission: short-distance airborne travel in aerosols produced by the coughing and sneezing of an infected host. The survival and lifespan of a pathogen in an aerosol (and thus the chances for infecting a new host) depend on the ambient temperature and humidity [139]. Dry conditions and temperature extremes can desiccate or destroy the pathogen, affecting transmission negatively. Meteorological records show extremely low winter temperatures (-20ºC in some nights) for years with relatively low disease prevalence in this study, supporting the influence of climatic conditions on respiratory disease prevalence.

Further parallels between the two diseases could not be drawn. PCV2 was influenced by host age and heterozygosity, whereas *Mhyo* was influenced only by hybrid status or partial domestic ancestry. This indicates that it is inaccurate to speak in generalisations when it comes to factors influencing wildlife disease prevalence. Wildlife disease ecology researchers will mostly need to consider each disease and perhaps each study area separately, as different factors may be driving disease prevalence in each case. More investigations of the various wildlife diseases are required to elucidate the possibly vast diversity in wildlife disease dynamics and ecology.

Disease transmission, disease mortality and the role of individual genetic composition

Theoretical knowledge of disease dynamics is mostly based on mathematical epidemiological models (chapter 1). Epidemiological models for host-specific directly transmitted diseases have in some cases successfully described

spatiotemporal patterns and dynamics of fast-spreading epidemics [141, 142] using spatially explicit modelling and simple assumptions on the transmission coefficient. However, endemic disease models that describe long-term disease persistence in host populations by incorporating birth-death dynamics, have hardly ever been evaluated using empirical data, but see [143]. The behaviour of these endemic disease models depends critically on the disease reproduction rate and thus the transmission coefficient [17, 143]. This transmission coefficient is difficult to extract from empirical data due to a lack of methods for direct quantification of exposure versus infection and the complexity of factors involved (host traits, pathogen traits and environmental conditions).

In this thesis, I show that successful transmission of two respiratory diseases is influenced by temporal climatic conditions (discussed above) and by the genetic format of the host (chapter 4 and 5). This suggests that using a single transmission coefficient and disease recovery or mortality rate (equal for all individuals) in epidemiological models is not realistic. The genetic heterogeneity of individuals, in terms overall heterozygosity and/or inherited immune capacity, can determine individual susceptibility to, or survival of, a particular disease and can thus play an important role in how diseases spread through a population, network or landscape [144]. This notion of individual genetic differences in susceptibility may prove crucial for a good understanding of the dynamics of various wildlife diseases. Theoretical work on individual differences in infectiousness indicate that such heterogeneity can significantly alter model outcomes and predictions [23, 145]. A future challenge for epidemiological modellers will be to realistically include heterogeneity in disease susceptibility at the individual level in order to more realistically and more accurately describe disease dynamics in wildlife populations.

What are the evolutionary effects of the use of human medicine?

The results of chapter 5 show that domestic hybrid wild boar (20-50% genetic domestic ancestry) have significantly higher *Mhyo* infection rates than wild-type wild boar. This finding suggests that lack of exposure to infectious agents due to nearly a century of veterinary care in pigs can lead to detectably degraded immune functions in domestic pigs and even in domestic hybrid wild boar. If this is true in wild boar, it raises questions about possibly similar evolutionary

effects of medical care in humans. A lack of pathogen selection pressure can lead to accumulation of deleterious mutations and loss of variation in immune genes. Has this also happened in humans? Has a history of human medicine causing degradation of immune functions in humans at an evolutionary time-scale?

The hygiene hypothesis states that, for humans, a lack of (childhood) exposure to infectious agents supresses natural development of the immune system and may lead to increased susceptibility to allergies, autoimmune disease, microorganisms and parasites [146-148]. This reasoning focusses on the lack of training of the immune system during early life, caused by hygienic practices, elimination of childhood diseases, reduced interactions with the environment, protection by medical care and the use of antibiotics. An evolutionary view is however missing from the discussion. What are the consequences of a lack of exposure to infectious agents for the maintenance of effective immune functions over multiple generations? Is the genetic immune capacity of a contemporary human at an equal level to humans from pre-industrial eras, or is it significantly degraded due to protection from pathogens by hygiene and medicine? The practical and ethical limitations of an experimental approach to address this issue are evident. However, these questions seem a neglected but relevant avenue for research.

Research outlook for swine epidemiology

The wild boar - domestic pig model system has great potential for epidemiological research. Swine are host to a wide array of pathogens, and the available high-quality porcine genomic information and diagnostic disease tests provide all the necessary tools to investigate host-pathogen interactions. This system has the added advantage of combining opportunities for controlled experiments in pig farms with opportunities for studying natural dynamics in wild boar populations.

Combined research on domestic pigs and wild boar is highly relevant for wildlife conservation and the pig breeding industry, and represents a wildlife-livestock interface within a species. Swine are also an important study system in the context of public health and emerging zoonoses, as swine have proven to be an important mixing vessel for pathogen adaptation to humans hosts [149, 150].

Heterozygosity fitness correlations such as found in chapter 4, and effects of (hybrid) ancestry on disease prevalence such as found in chapter 5, may open the door for genome-wide association studies in relation to disease resistance in wild boar. Experimental approaches in a domestic pig farm setup can confirm for instance effects of inbreeding of specific genes on disease resistance. Together, these approaches may increase our understanding of exactly how the co-evolutionary arms race between host and pathogen works.

In addition, it would be informative to expand the wild boar disease data with a few more diseases. Diagnostic tests such as antibody assays are available for a number of other swine diseases, *e.g.*: porcine parvovirus, porcine respiratory and reproductive syndrome virus or *Actinobacillus pleuropneumonia*. Such additional data would clarify if domestic hybrids are more susceptible than wild boar to all bacterial diseases, but not viruses. If so, this would strengthen the case for an evolutionary effect of antibiotics use in livestock industries.

Conclusions

A major challenge in disease ecology is to obtain a good overview of the driving forces of infectious disease prevalence in wildlife populations. The aim of this thesis was to specifically identify the main factors that influence PCV2 and *Mhyo* disease prevalence in European wild boar populations. I show that climatic conditions represent a shared driving force for temporal dynamics in both diseases, while effects of host age and genetic factors differed between PCV2 and *Mhyo* (chapter 4 and 5). I demonstrate that PCV2 related mortality in wild boar juveniles and yearlings is mediated by heterozygosity. The most inbred individuals do not survive first exposure to this disease. In chapter 2 and 3 I show that genetic introgression from domestic pigs into wild boar populations is more recent, more common and more widespread than expected. Disease prevalence was not significantly different between populations, but *Mhyo* prevalence was higher in hybrids with partial domestic ancestry than in wild-type wild boar. I propose that this is a consequence of the inheritance of deleterious alleles that originate from a lack of pathogen selection pressure in domestic pigs due to veterinary care. The fact that different factors drive the prevalence of the two diseases examined in this thesis, indicates that speaking in generalities is inaccurate within the context of disease ecology. Research on

different model systems and pathogens is required to get a better overview of the possibly vast diversity of driving factors for wildlife disease prevalence. One of the central questions in wildlife epidemiology relates to the underlying mechanism for the observation of pathogen aggregation or overdominance [15, 22]. Genetic heterogeneity has been suggested to explain patterns of pathogen aggregation [46] and the results in this thesis support this hypothesis. I propose that an important next step for the field of wildlife disease ecology would be to incorporate genetic heterogeneity into epidemiological modelling. Population genetic theory provides sufficient information to model host heterozygosity as well as Mendelian inheritance of deleterious alleles, which will have to be embedded in an epidemiological framework.

References

1. World Health Organisation: **Removing obstacles to healthy development**. In.; 1999.
2. Wolfe ND, Dunavan CP, Diamond J: **Origins of major human infectious diseases**. *Nature* 2007, **447**(7142):279-283.
3. Daszak P: **Emerging infectious diseases of wildlife - Threats to biodiversity and human health (vol 287, pg 443, 2000)**. *Science* 2000, **287**(5459):1756-1756.
4. Karesh WB, Cook RA: **One world - one health**. *Clin Med* 2009, **9**(3):259-260.
5. Pritchard G, Dennis I, Waddilove J: **Biosecurity: reducing disease risks to pig breeding herds**. *In Practice* 2005, **27**(5):230-237.
6. Riley S: **Large-scale spatial-transmission models of infectious disease**. *Science* 2007, **316**(5829):1298-1301.
7. de Meeus T, McCoy KD, Prugnolle F, Chevillon C, Durand P, Hurtrez-Bousses S, Renaud F: **Population genetics and molecular epidemiology or how to "debusquer la bete"**. *Infect Genet Evol* 2007, **7**(2):308-332.
8. Lloyd-Smith JO, Cross PC, Briggs CJ, Daugherty M, Getz WM, Latto J, Sanchez MS, Smith AB, Swei A: **Should we expect population thresholds for wildlife disease?** *Trends Ecol Evol* 2005, **20**(9):511-519.
9. Bar-David S, Lloyd-Smith JO, Getz WM: **Dynamics and management of infectious disease in colonizing populations**. *Ecology* 2006, **87**(5):1215-1224.
10. Bengis RG, Kock RA, Fischer J: **Infectious animal diseases: the wildlife/livestock interface**. *Rev Sci Tech Off Int Epizoot* 2002, **21**(1):53-65.
11. Kruse H, Kirkemo AM, Handeland K: **Wildlife as source of zoonotic infections**. *Emerg Infect Dis* 2004, **10**(12):2067-2072.
12. McCallum H, Dobson A: **Detecting disease and parasite threats to endangered species and ecosystems**. *Trends Ecol Evol* 1995, **10**(5):190-194.
13. Altizer S, Harvell D, Friedle E: **Rapid evolutionary dynamics and disease threats to biodiversity**. *Trends Ecol Evol* 2003, **18**(11):589-596.
14. Webster RG, Wright SM, Castrucci MR, Bean WJ, Kawaoka Y: **Influenza - a Model of an Emerging Virus-Disease**. *Intervirology* 1993, **35**(1-4):16-25.
15. Anderson RM, May RM: **Regulation and Stability of Host-Parasite Population Interactions .1. Regulatory Processes**. *J Anim Ecol* 1978, **47**(1):219-247.
16. May RM, Anderson RM: **Regulation and Stability of Host-Parasite Population Interactions .2. Destabilizing Processes**. *J Anim Ecol* 1978, **47**(1):249-267.
17. Hethcote HW: **The mathematics of infectious diseases**. *Siam Rev* 2000, **42**(4):599-653.
18. Anderson RM: **Discussion: the Kermack-McKendrick epidemic threshold theorem**. *Bull Math Biol* 1991, **53**(1-2):3-32.
19. Verdasca J, da Gama MMT, Nunes A, Bernardino NR, Pacheco JM, Gomes MC: **Recurrent epidemics in small world networks**. *J Theor Biol* 2005, **233**(4):553-561.
20. Hartvigsen G, Dresch JM, Zielinski AL, Macula AJ, Leary CC: **Network structure, and vaccination strategy and effort interact to affect the dynamics of influenza epidemics**. *J Theor Biol* 2007, **246**(2):205-213.
21. Olinky R, Stone L: **Unexpected epidemic thresholds in heterogeneous networks: the role of disease transmission**. *Phys Rev E Stat Nonlin Soft Matter Phys* 2004, **70**(3 Pt 1):030902.
22. Grenfell BT, Dobson AP: **Ecology of infectious diseases in natural populations**. Cambridge: Cambridge University press; 1995.

23. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM: **Superspreading and the effect of individual variation on disease emergence**. *Nature* 2005, **438**(7066):355-359.
24. McCallum H, Barlow N, Hone J: **How should pathogen transmission be modelled?** *Trends Ecol Evol* 2001, **16**(6):295-300.
25. Giuffra E, Kijas JMH, Amarger V, Carlborg O, Jeon JT, Andersson L: **The origin of the domestic pig: Independent domestication and subsequent introgression**. *Genetics* 2000, **154**(4):1785-1791.
26. Larson G, Dobney K, Albarella U, Fang MY, Matisoo-Smith E, Robins J, Lowden S, Finlayson H, Brand T, Willerslev E *et al*: **Worldwide phylogeography of wild boar reveals multiple centers of pig domestication**. *Science* 2005, **307**(5715):1618-1621.
27. Gortazar C, Acevedo P, Ruiz-Fons F, Vicente J: **Disease risks and overabundance of game species**. *Eur J Wildl Res* 2006, **52**(2):81-87.
28. Deutz A, Kofer J: **Pig and wild boar as carriers of zoonoses**. *Berliner Munchener Tierarztl Wochenschr* 1999, **112**(8):305-310.
29. van Reeth K: **Avian and swine influenza viruses: our current understanding of the zoonotic risk**. *Vet Res* 2007, **38**(2):243-260.
30. Meng XJ, Lindsay DS, Sriranganathan N: **Wild boars as sources for infectious diseases in livestock and humans**. *Philos Trans R Soc B-Biol Sci* 2009, **364**(1530):2697-2707.
31. Depner KR, Muller A, Gruber A, Rodriguez A, Bickhardt K, Liess B: **Classical Swine Fever in Wild Boar (Sus Scrofa) - Experimental Infections and Viral Persistence**. *Deut Tierarztl Woch* 1995, **102**(10):381-384.
32. Ruiz-Fons F, Vicente J, Vidal D, Hofle U, Villanua D, Gauss C, Segales J, Almeria S, Montoro V, Gortazar C: **Seroprevalence of six reproductive pathogens in European wild boar (Sus scrofa) from Spain: The effect on wild boar female reproductive performance**. *Theriogenology* 2006, **65**(4):731-743.
33. Vicente J, Segales J, Hofle U, Balasch M, Plana-Duran J, Domingo M, Gortazar C: **Epidemiological study on porcine circovirus type 2 (PCV2) infection in the European wild boar (Sus scrofa)**. *Vet Res* 2004, **35**(2):243-253.
34. Gray GD, Gill HS: **Host Genes, Parasites and Parasitic Infections**. *Int J Parasit* 1993, **23**(4):485-494.
35. Briedermann: **Schwarzwild.** . Berlin, Germany: VEB Deutscher Landwirtschaftsverlag; 1990.
36. Dexter N: **The influence of pasture distribution, temperature and sex on home-range size of feral pigs in a semi-arid environment**. *Wildl Res* 1999, **26**(6):755-762.
37. Truve J, Lemel J: **Timing and distance of natal dispersal for wild boar Sus scrofa in Sweden**. *Wildlife Biol* 2003, **9**:51-57.
38. Schley L, Roper TJ: **Diet of wild boar Sus scrofa in Western Europe, with particular reference to consumption of agricultural crops**. *Mammal Rev* 2003, **33**(1):43-56.
39. Leaper R, Massei G, Gorman ML, Aspinall R: **The feasibility of reintroducing Wild Boar (Sus scrofa) to Scotland**. *Mammal Rev* 1999, **29**(4):239-259.
40. Fernandez-Llario P, Mateos-Quesada P: **Body size and reproductive parameters in the wild boar Sus scrofa**. *Acta Theriol* 1998, **43**(4):439-444.
41. Saez-Royuela C, Telleria JL: **The Increased Population of the Wild Boar (Sus-Scrofa L) in Europe**. *Mammal Rev* 1986, **16**(2):97-101.
42. Humphray SJ, Scott CE, Clark R, Marron B, Bender C, Camm N, Davis J, Jenks A, Noon A, Patel M *et al*: **A high utility integrated map of the pig genome**. *Genome Biol* 2007, **8**(7).
43. Ramos AM, Crooijmans RPMA, Affara NA, Amaral AJ, Archibald AL: **Design of a high-density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology**. *Plos 1* 2009, **4**(8).

44. Vignal A, Milan D, SanCristobal M, Eggen A: **A review on SNP and other types of molecular markers and their use in animal genetics**. *Genet Sel Evol* 2002, **34**(3):275-305.
45. Morin PA, Luikart G, Wayne RK: **SNPs in ecology, evolution and conservation**. *Trends Ecol Evol* 2004, **19**(4):208-216.
46. Archie EA, Luikart G, Ezenwa VO: **Infecting epidemiology with genetics: a new frontier in disease ecology**. *Trends Ecol Evol* 2009, **24**(1):21-30.
47. Scandura M, Iacolina L, Crestanello B, Pecchioli E, Di Benedetto MF, Russo V, Davoli R, Apollonio M, Bertorelle G: **Ancient vs. recent processes as factors shaping the genetic variation of the European wild boar: are the effects of the last glaciation still detectable?** *Mol Ecol* 2008, **17**(7):1745-1762.
48. Frantz AC, Bertouille S, Eloy MC, Licoppe A, Chaumont F, Flamand MC: **Comparative landscape genetic analyses show a Belgian motorway to be a gene flow barrier for red deer (Cervus elaphus), but not wild boars (Sus scrofa)**. *Mol Ecol* 2012, **21**(14):3445-3457.
49. Vernesi C, Crestanello B, Pecchioli E, Tartari D, Caramelli D, Hauffe H, Bertorelle G: **The genetic impact of demographic decline and reintroduction in the wild boar (Sus scrofa): A microsatellite analysis**. *Mol Ecol* 2003, **12**(3):585-595.
50. Larson G, Albarella U, Dobney K, Rowley-Conwy P, Schibler J, Tresset A, Vigne JD, Edwards CJ, Schlumbaum A, Dinu A *et al*: **Ancient DNA, pig domestication, and the spread of the Neolithic into Europe**. *Proc Natl Acad Sci U S A* 2007, **104**(39):15276-15281.
51. Scandura M, Iacolina L, Apollonio M: **Genetic diversity in the European wild boar Sus scrofa: phylogeography, population structure and wild x domestic hybridization**. *Mammal Rev* 2011, **41**(2):125-137.
52. Verhoeven KJ, Macel M, Wolfe LM, Biere A: **Population admixture, biological invasions and the balance between local adaptation and inbreeding depression**. *Proc Biol Sci* 2011, **278**(1702):2-8.
53. Alves PC, Pinheiro I, Godinho R, Vicente J, Gortazar C, Scandura M: **Genetic diversity of wild boar populations and domestic pig breeds (Sus scrofa) in South-western Europe**. *Biol J Linnean Soc* 2010, **101**(4):797-822.
54. Fang M, Berg F, Ducos A, Andersson L: **Mitochondrial haplotypes of European wild boars with 2n=36 are closely related to those of European domestic pigs with 2n=38**. *Anim Genet* 2006, **37**(5):459-464.
55. Kraus RHS, Kerstens HHD, Van Hooft P, Crooijmans RPMA, Van der Poel JJ, Elmberg J, Vignal A, Huang YH, Li N, Prins HHT *et al*: **Genome wide SNP discovery, analysis and evaluation in mallard (Anas platyrhynchos)**. *Bmc Genomics* 2011, **12**.
56. Miller JM, Poissant J, Kijas JW, Coltman DW: **A genome-wide set of SNPs detects population substructure and long range linkage disequilibrium in wild sheep**. *Mol Ecol Resour* 2011, **11**(2):314-322.
57. Narum SR, Banks M, Beacham TD, Bellinger MR, Campbell MR, Dekoning J, Elz A, Guthrie CM, 3rd, Kozfkay C, Miller KM *et al*: **Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms**. *Mol Ecol* 2008, **17**(15):3464-3477.
58. Willing EM, Bentzen P, van Oosterhout C, Hoffmann M, Cable J, Breden F, Weigel D, Dreyer C: **Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies**. *Mol Ecol* 2010, **19**(5):968-984.
59. Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb LW: **Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms**. *Mol Ecol Resour* 2011, **11 Suppl 1**:1-8.
60. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ *et al*: **PLINK: a tool set for whole-genome**

association and population-based linkage analyses. *Am J Hum Genet* 2007, **81**(3):559-575.
61. Patterson N, Price AL, Reich D: **Population structure and eigenanalysis**. *PLoS Genet* 2006, **2**(12):e190.
62. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: **Principal components analysis corrects for stratification in genome-wide association studies**. *Nat Genet* 2006, **38**(8):904-909.
63. Pritchard JK, Stephens M, Donnelly P: **Inference of population structure using multilocus genotype data**. *Genetics* 2000, **155**(2):945-959.
64. Evanno G, Regnaut S, Goudet J: **Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study**. *Mol Ecol* 2005, **14**(8):2611-2620.
65. Jombart T: **adegenet: a R package for the multivariate analysis of genetic markers**. *Bioinformatics* 2008, **24**(11):1403-1405.
66. Luetkemeier ES, Sodhi M, Schook LB, Malhi RS: **Multiple Asian pig origins revealed through genomic analyses**. *Mol Phylogenet Evol* 2010, **54**(3):680-686.
67. Bonfield JK, Smith KF, Staden R: **A new DNA sequence assembly program**. *Nucleic Acids Res* 1995, **23**(24):4992-4999.
68. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al*: **Clustal W and clustal X version 2.0**. *Bioinformatics* 2007, **23**(21):2947-2948.
69. Glez-Pena D, Gomez-Blanco D, Reboiro-Jato M, Fdez-Riverola F, Posada D: **ALTER: program-oriented conversion of DNA and protein alignments**. *Nucleic Acids Res* 2010, **38**:W14-W18.
70. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0**. *Mol Biol Evol* 2007, **24**(8):1596-1599.
71. Fang MY, Andersson L: **Mitochondrial diversity in European and Chinese pigs is consistent with population expansions that occurred prior to domestication**. *Proc R Soc B-Biol Sci* 2006, **273**(1595):1803-1810.
72. Goulding MJ: **Possible genetic sources of free-living Wild Boar (Sus scrofa) in southern England**. *Mammal Rev* 2001, **31**(3):245-248.
73. Gethoffer F, Sodeikat G, Pohlmeyer K: **Reproductive parameters of wild boar (Sus scrofa) in three different parts of Germany**. *Eur J Wildl Res* 2007, **53**(4):287-297.
74. Geisser H, Reyer HU: **The influence of food and temperature on population density of wild boar Sus scrofa in the Thurgau (Switzerland)**. *J Zool* 2005, **267**:89-96.
75. Bieber C, Ruf T: **Population dynamics in wild boar Sus scrofa: ecology, elasticity of growth rate and implications for the management of pulsed resource consumers**. *J Appl Ecol* 2005, **42**(6):1203-1213.
76. Manel S, Schwartz MK, Luikart G, Taberlet P: **Landscape genetics: combining landscape ecology and population genetics**. *Trends Ecol Evol* 2003, **18**(4):189-197.
77. Hartl GB, Zachos F, Nadlinger K: **Genetic diversity in European red deer (Cervus elaphus L.): anthropogenic influences on natural populations**. *Cr Biol* 2003, **326**:S37-S42.
78. Pierpaoli M, Biro ZS, Herrmann M, Hupe K, Fernandes M, Ragni B, Szemethy L, Randi E: **Genetic distinction of wildcat (Felis silvestris) populations in Europe, and hybridization with domestic cats in Hungary**. *Mol Ecol* 2003, **12**(10):2585-2598.
79. Randi E: **Detecting hybridization between wild species and their domesticated relatives**. *Mol Ecol* 2008, **17**(1):285-293.
80. Frantz AC, Pourtois JT, Heuertz M, Schley L, Flamand MC, Krier A, Bertouille S, Chaumont F, Burke T: **Genetic structure and assignment tests demonstrate illegal translocation of red deer (Cervus elaphus) into a continuous population**. *Mol Ecol* 2006, **15**(11):3191-3203.

81. Yalden DW: **The History of British Mammals**. London, UK: Poyser Natural History; 1999.
82. Pelzers E: **Beknopte historie van het wilde zwijn in Nederland**. *Zoogdier* 1990, **1**(1):25-29.
83. Prins HHT: **The origins of grassland communities in northwestern Europe**. In: *Grazing and Conservation Management.* Edited by WallisDeVries MF, Bakker JP, Van Wieren SE. Boston: Kluwer Academic Publishers; 1998: 55-105.
84. Andrzejewski R, Jezierski W: **Management of a Wild Boar Population and Its Effects on Commercial Land**. *Acta Theriol* 1978, **23**(19-3):309-339.
85. Booth WD: **Wild boar farming in the United Kingdom**. *IBEX J Mt Ecol* 1995(3):245-248.
86. Frantz AC, Massei G, Burke T: **Genetic evidence for past hybridization between domestic pigs and English wild boars**. *Conserv Genet* 2012, **13**(5):1355-1364.
87. Frankham R: **Quantitative genetics in conservation biology**. *Genet Res* 1999, **74**(3):237-244.
88. Schley L, Dufrene M, Krier A, Frantz AC: **Patterns of crop damage by wild boar (Sus scrofa) in Luxembourg over a 10-year period**. *Eur J Wildl Res* 2008, **54**(4):589-599.
89. Laikre L, Schwartz MK, Waples RS, Ryman N, Grp GW: **Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals**. *Trends Ecol Evol* 2010, **25**(9):520-529.
90. Laurie CC, Nickerson DA, Anderson AD, Weir BS, Livingston RJ, Dean MD, Smith KL, Schadt EE, Nachman MW: **Linkage disequilibrium in wild mice**. *PLoS Genet* 2007, **3**(8):1487-1495.
91. Villa-Angulo R, Matukumalli LK, Gill CA, Choi J, Van Tassell CP, Grefenstette JJ: **High-resolution haplotype block structure in the cattle genome**. *Bmc Genet* 2009, **10**.
92. Fan JB, Gunderson KL, Bibikova M, Yeakley JM, Chen J, Garcia EW, Lebruska LL, Laurent M, Shen R, Barker D: **Illumina universal bead arrays**. In: *DNA Microarrays Part a: Array Platforms and Wet-Bench Protocols.* vol. 410. San Diego: Elsevier Academic Press Inc; 2006: 57-+.
93. Herraez DL, Schafer H, Mosner J, Fries HR, Wink M: **Comparison of microsatellite and single nucleotide polymorphism markers for the genetic analysis of a Galloway cattle population**. *Z Naturforsch C* 2005, **60**(7-8):637-643.
94. Schopen GCB, Bovenhuis H, Visker MHPW, van Arendonk JAM: **Comparison of information content for microsatellites and SNPs in poultry and cattle**. *Anim Genet* 2008, **39**(4):451-453.
95. Huson DH, Bryant D: **Application of phylogenetic networks in evolutionary studies**. *Mol Biol Evol* 2006, **23**(2):254-267.
96. Goudet J: **HIERFSTAT, a package for R to compute and test hierarchical F-statistics**. *Mol Ecol Notes* 2005, **5**(1):184-186.
97. Zheng X: **SNPRelate: Parrallel Computing Toolset for Genome-Wide Association Studies**. *R package version 095* 2012.
98. Milligan BG: **Maximum-likelihood estimation of relatedness**. *Genetics* 2003, **163**(3):1153-1167.
99. Thompson EA: **Estimation of Pairwise Relationships**. *Ann Hum Genet* 1975, **39**(Oct):173-188.
100. Oksanen J, Blanchet GF, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H: **Vegan: Community Ecology Package.** *R package version 20-3* 2012.
101. Epstein PR: **Emerging diseases and ecosystem instability: new threats to public health**. *Am J Public Health* 1995, **85**(2):168-172.
102. Finkenstadt BF, Grenfell BT: **Time series modelling of childhood diseases: a dynamical systems approach**. *J Roy Stat Soc C-App* 2000, **49**:187-205.

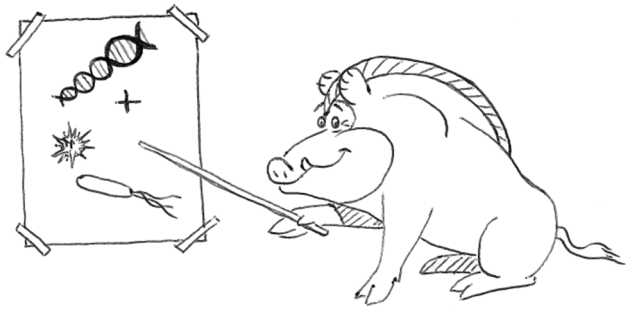
103. Ruiz-Fons F, Segales J, Gortazar C: **A review of viral diseases of the European wild boar: Effects of population dynamics and reservoir role**. *Vet J* 2008, **176**(2):158-169.
104. Segales J, Allan GM, Domingo M: **Porcine circovirus diseases**. *Animal health research reviews / Conference of Research Workers in Animal Diseases* 2005, **6**(2):119-142.
105. Allan GM, Ellis JA: **Porcine circoviruses: a review**. *J Vet Diagn Invest* 2000, **12**(1):3-14.
106. Gilpin DF, McCullough K, Meehan BM, McNeilly F, McNair I, Stevenson LS, Foster JC, Ellis JA, Krakowka S, Adair BM *et al*: **In vitro studies on the infection and replication of porcine circovirus type 2 in cells of the porcine immune system**. *Vet Immunol Immunopathol* 2003, **94**(3-4):149-161.
107. Allan GM, Kennedy S, McNeilly F, Foster JC, Ellis JA, Krakowka SJ, Meehan BM, Adair BM: **Experimental reproduction of severe wasting disease by co-infection of pigs with porcine circovirus and porcine parvovirus**. *J Comp Pathol* 1999, **121**(1):1-11.
108. Allan GM, McNeilly F, Ellis J, Krakowka S, Meehan B, McNair I, Walker I, Kennedy S: **Experimental infection of colostrum deprived piglets with porcine circovirus 2 (PCV2) and porcine reproductive and respiratory syndrome virus (PRRSV) potentiates PCV2 replication**. *Arch Virol* 2000, **145**(11):2421-2429.
109. Kennedy S, Moffett D, McNeilly F, Meehan B, Ellis J, Krakowka S, Allan GM: **Reproduction of lesions of postweaning multisystemic wasting syndrome by infection of conventional pigs with porcine circovirus type 2 alone or in combination with porcine parvovirus**. *J Comp Pathol* 2000, **122**(1):9-24.
110. Lopez-Soria S, Segales J, Nofrarias M, Calsamiglia M, Ramirez H, Minguez A, Serrano JM, Marin O, Callen A: **Evidence of genetics influence on the expression of porcine circovirus disease (PCVD) in postweaning pigs**. *36th Annual Meeting of the American Association of Swine Veterinarians, Proceedings* 2005:391-393.
111. Ramamoorthy S, Meng XJ: **Porcine circoviruses: a minuscule yet mammoth paradox**. *Animal health research reviews / Conference of Research Workers in Animal Diseases* 2009, **10**(1):1-20.
112. Reiner G, Bronnert B, Hohloch C, Reinacher M, Willems H: **Distribution of ORF2 and ORF3 genotypes of porcine circovirus type 2 (PCV-2) in wild boars and domestic pigs in Germany**. *Vet Microbiol* 2011, **148**(2-4):372-376.
113. Turner AK, Begon M, Jackson JA, Bradley JE, Paterson S: **Genetic Diversity in Cytokines Associated with Immune Variation and Resistance to Multiple Pathogens in a Natural Rodent Population**. *PLoS Genet* 2011, **7**(10).
114. Savage AE, Zamudio KR: **MHC genotypes associate with resistance to a frog-killing fungus**. *Proc Natl Acad Sci U S A* 2011, **108**(40):16705-16710.
115. Coltman DW, Pilkington J, Kruuk LEB, Wilson K, Pemberton JM: **Positive genetic correlation between parasite resistance and body size in a free-living ungulate population**. *Evolution* 2001, **55**(10):2116-2125.
116. Acevedo-Whitehouse K, Vicente J, Gortazar C, Hofle U, Fernandez-De-Mera IG, Amos W: **Genetic resistance to bovine tuberculosis in the Iberian wild boar**. *Mol Ecol* 2005, **14**(10):3209-3217.
117. Spielman D, Brook BW, Briscoe DA, Frankham R: **Does inbreeding and loss of genetic diversity decrease disease resistance?** *Conserv Genet* 2004, **5**(4):439-448.
118. Curtis J, Bourne FJ: **Immunoglobulin quantitation in sow serum, colostrum and milk, and the serum of young pigs.** . *Biochimica et Biophysica Acta (BBA) - Protein Structure* 1971, **236**(1):319-332.

119. Keller LF, Waller DM: **Inbreeding effects in wild populations**. *Trends Ecol Evol* 2002, **17**(5):230-241.
120. Hammer R, Ritzmann M, Palzer A, Lang C, Hammer B, Pesch S, Ladinig A: **Porcine Reproductive and Respiratory Syndrome Virus and Porcine Circovirus Type 2 Infections in Wild Boar (Sus Scrofa) in Southwestern Germany**. *J Wildl Dis* 2012, **48**(1):87-94.
121. Segales J, Valero O, Espinal A, Lopez-Soria S, Nofrarias M, Calsamiglia M, Sibila M: **Exploratory study on the influence of climatological parameters on Mycoplasma hyopneumoniae infection dynamics**. *International journal of biometeorology* 2012, **56**(6):1167-1171.
122. Stark KD: **The role of infectious aerosols in disease transmission in pigs**. *Vet J* 1999, **158**(3):164-181.
123. Lipowski A: **European wild boar (Sus scrofa L.) as a reservoir of infectious diseases for domestic pigs**. *Med Weter* 2003, **59**(10):861-863.
124. Witte W: **Selective pressure by antibiotic use in livestock**. *Int J Antimicrob Ag* 2000, **16**:S19-S24.
125. Servanty S, Gaillard JM, Toïgo C, Brandt S, Baubet E: **Pulsed resources and climate-induced variation in the reproductive traits of wild boar under high hunting pressure**. *J Anim Ecol* 2009, **78**:1278-1290.
126. Wuensch U, Nitter G, Schueler L: **Genetic and economic evaluation of genetic improvement schemes in pigs I. Methodology with an application to a three-way crossbreeding scheme**. *Arch Tierz-Arch Anim Breed* 1999, **42**(6):571-582.
127. Cohen ML: **Epidemiology of Drug-Resistance - Implications for a Postantimicrobial Era**. *Science* 1992, **257**(5073):1050-1055.
128. Whiteman NK, Matson KD, Bollmer JL, Parker PG: **Disease ecology in the Galapagos Hawk (Buteo galapagoensis): host genetic diversity, parasite load and natural antibodies**. *Proc R Soc B-Biol Sci* 2006, **273**(1588):797-804.
129. Van Valen L: **Molecular Evolution as Predicted by Natural-Selection**. *J Mol Evol* 1974, **3**(2):89-101.
130. Ebert D, Hamilton WD: **Sex against virulence: The coevolution of parasitic diseases**. *Trends Ecol Evol* 1996, **11**(2):A79-A82.
131. Salathe M, Kouyos RD, Bonhoeffer S: **The state of affairs in the kingdom of the Red Queen**. *Trends Ecol Evol* 2008, **23**(8):439-445.
132. Whittlestone P: **Mycoplasma in Enzootic Pneumonia of Pigs**. *Ann NY Acad Sci* 1967, **143**(A1):271-280.
133. Caruso JP, Ross RF: **Effects of Mycoplasma-Hyopneumoniae and Actinobacillus (Hemophilus) Pleuropneumoniae Infections on Alveolar Macrophage Functions in Swine**. *Am J Vet Res* 1990, **51**(2):227-231.
134. Kishima M, Ross RF: **Suppressive Effect of Nonviable Mycoplasma-Hypopneumoniae on Phytohemagglutinin-Induced Transformation of Swine Lymphocytes**. *Am J Vet Res* 1985, **46**(11):2366-2368.
135. Meyns T, Maes D, Dewulf J, Vicca J, Haesebrouck F, de Kruif A: **Quantification of the spread of Mycoplasma hyopneumoniae in nursery pigs using transmission experiments**. *Prev Vet Med* 2004, **66**(1-4):265-275.
136. Dee S, Otake S, Oliveira S, Deen J: **Evidence of long distance airborne transport of porcine reproductive and respiratory syndrome virus and Mycoplasma hyopneumoniae**. *Vet Res* 2009, **40**(4).
137. Thacker EL, Thacker BJ, Boettcher TB, Jayappa H: **Comparison of antibody production, lymphocyte stimulation, and protection induced by four commercial Mycoplasma hyopneumoniae bacterins**. *Swine Health and Production* 1998, **6**(3):107-112.
138. Maes D, Segales J, Meyns T, Sibila M, Pieters M, Haesebrouck F: **Control of Mycoplasma hyopneumoniae infections in pigs**. *Vet Microbiol* 2008, **126**(4):297-309.

139. Goodwin RFW: **Apparent Reinfection of Enzootic-Pneumonia-Free Pig Herds - Search for Possible Causes**. *Vet Rec* 1985, **116**(26):690-694.
140. Feulner PGD, Gratten J, Kijas JW, Visscher PM, Pemberton JM, Slate J: **Introgression and the fate of domesticated genes in a wild mammal population**. *Mol Ecol* 2013(early view, doi:10.111/mec.12378).
141. Anderson RM, Jackson HC, May RM, Smith AM: **Population-Dynamics of Fox Rabies in Europe**. *Nature* 1981, **289**(5800):765-771.
142. Smith DL, Lucey B, Waller LA, Childs JE, Real LA: **Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes**. *Proc Natl Acad Sci U S A* 2002, **99**(6):3668-3672.
143. Dwyer G, Levin SA, Buttel L: **A Simulation-Model of the Population-Dynamics and Evolution of Myxomatosis**. *Ecol Monogr* 1990, **60**(4):423-447.
144. Hawley DM, Altizer SM: **Disease ecology meets ecological immunology: understanding the links between organismal immunity and infection dynamics in natural populations**. *Funct Ecol* 2011, **25**(1):48-60.
145. Bishop SC, Stear MJ: **Modeling of host genetics and resistance to infectious diseases: understanding and controlling nematode infections**. *Vet Parasitol* 2003, **115**(2):147-166.
146. Yazdanbakhsh M, Kremsner PG, van Ree R: **Allergy, parasites, and the hygiene hypothesis**. *Science* 2002, **296**(5567):490-494.
147. Wills-Karp M, Santeliz J, Karp CL: **The germless theory of allergic disease: revisiting the hygiene hypothesis**. *Nature reviews Immunology* 2001, **1**(1):69-75.
148. Guarner F, Bourdet-Sicard R, Brandtzaeg P, Gill HS, McGuirk P, van Eden W, Versalovic J, Weinstock JV, Rook GA: **Mechanisms of disease: the hygiene hypothesis revisited**. *Nature clinical practice Gastroenterology & hepatology* 2006, **3**(5):275-284.
149. Castrucci MR, Donatelli I, Sidoli L, Barigazzi G, Kawaoka Y, Webster RG: **Genetic Reassortment between Avian and Human Influenza-a Viruses in Italian Pigs**. *Virology* 1993, **193**(1):503-506.
150. Brown IH, Harris PA, McCauley JW, Alexander DJ: **Multiple genetic reassortment of avian and human influenza A viruses in European pigs, resulting in the emergence of an H1N2 virus of novel genotype**. *The Journal of general virology* 1998, **79 (Pt 12)**:2947-2955.

Summary

Essentially all human and livestock infectious diseases ultimately originate from wildlife populations. Wildlife infectious diseases are therefore considered to be an important risk factor for biodiversity, livestock and public health. However, the driving forces of wildlife disease dynamics are poorly understood. Technological and analytical advances in molecular ecology increasingly provide the means to investigate wildlife disease dynamics, and the properties of the domestic pig and wild boar study systems offer interesting opportunities for epidemiological research.



In this thesis I report on the influence of a number of demographic, environmental and host genetic factors on infectious disease prevalence in European wild boar. I focus on two swine respiratory diseases: porcine circovirus type 2 (PCV2) and *Mycoplasma hyopneumoniae* (*Mhyo*).

In chapter 2 I start off by unexpectedly identifying a relatively high frequency (10%) of recent genetic introgression from domestic pigs among 88 wild boar samples from Northwest Europe.

This work included a novel approach for identifying and quantifying introgression based on a high-density genetic assay and hybridisation simulations.

The most likely route of genetic introgression is through release of farmed wild boar, which are bred for production of luxury meat and which are frequently crossed with domestic pigs to increase growth rates and reproduction.

From the high-density genetic assay used in chapter 2, I derived a low-density genetic assay, which was applied in chapter 3 to genotype a total of 645 wild boar samples. With these data I constructed the genetic population structure of the study area, comprising the Netherlands, Luxembourg and adjacent parts of Western Germany. Six geographically distinct wild boar populations were identified, and 25 recent domestic hybrids (1-5 generations ago) were found geographically scattered across the area, indicating wide-spread release of farmed wild boar. Genetic population structure was shaped by a landscape barrier (the river Rhine) and historical human reintroduction/translocation events.

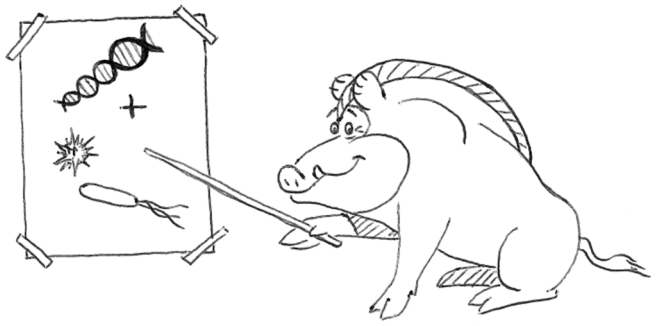
In chapter 4 I show that PCV2 prevalence varies between years, but also increases with wild boar age and is finally influenced by wild boar genetic heterozygosity. The age effect is a consequence of cumulative exposure to the virus and a typical PCV2 trait: silent intracellular persistence, which leads to increasing chances of PCV2 reinfection over time. The heterozygosity effect demonstrates that the most inbred wild boar have a lower chance to survive PCV2 infection in their first years of life. This shows that PCV2 infection can act as a selective force and that disease infection and mortality can be mediated by genetic heterogeneity.

In chapter 5 I report that *Mhyo* prevalence shows the same between year variation as PCV2 prevalence. This similarity is most likely caused by climatic conditions, which influence airborne disease transmission through aerosols. In addition, domestic hybrid wild boar intriguingly displayed a higher *Mhyo* antibody prevalence than wild-type animals. This suggests a higher susceptibility to this disease for animals with partial domestic ancestry. I hypothesise that this is caused by genetic degradation of immune functions in domestic pigs due to a history of veterinary care and the use of antibiotics in the livestock industry.

Finally I review the findings of this thesis in the Synthesis (chapter 6). Here, I reflect shortly on the evolutionary effects of hygiene and medicine on human immune capacity. With regard to wildlife disease ecology, I provide strong support for the hypothesis that genetic heterogeneity may explain the issue of pathogen aggregation, which is a central issue in epidemiology. Moreover, I propose that our current understanding of wildlife epidemiology and disease ecology can be improved by integrating population genetic and epidemiological models. In conclusion, this thesis shows that combining genetic and antibody data is a powerful approach and that host genetic factors and individual heterogeneity are important aspects for wildlife disease ecology research.

Samenvatting

Vrijwel alle menselijke infectieziekten zijn uiteindelijk afkomstig van wilde dieren. Infectieziekten van wilde dieren worden daarom beschouwd als een belangrijke risicofactor voor de volksgezondheid, maar ook voor biodiversiteit en voor de veehouderij. Echter, de drijvende krachten achter de dynamiek van ziekten bij wilde dieren zijn veelal onbekend. Technologische en analytische vooruitgang in de moleculaire ecologie bieden steeds meer middelen om de dynamiek van ziekten bij wilde dieren te bestuderen, en de eigenschappen van tamme varkens en wilde zwijnen bieden interessante mogelijkheden voor epidemiologisch onderzoek.



In dit proefschrift laat ik zien dat een aantal demografische, klimatologische en genetische factoren een belangrijke invloed hebben op de prevalentie van ziekten bij Europese wilde zwijnen. Ik heb gekeken naar twee luchtweginfecties: porcine circovirus type 2 (PCV2) en *Mycoplasma hyopneumoniae* (*Mhyo*).

In hoofdstuk 2 identificeer ik onverwacht een relatief hoge frequentie (10%) van recente genetische introgressie van tamme varkens naar wilde zwijnen in 88 zwijnen monsters.

Dit werk omvatte een nieuwe benadering voor het identificeren en kwantificeren van genetische introgressie op basis van een genetische analyse met een hoge dichtheid aan merkers en hybridisatie simulaties. De meest waarschijnlijke route van genetische introgressie is door introductie van gefokte wilde zwijnen, die gehouden worden voor de productie van luxe vlees en vaak worden gekruist met tamme varkens om de groei- en voortplantingssnelheid te verhogen.

Van de genetische analyse die gebruikt is in hoofdstuk 2, heb ik een genetische analyse met een lage dichtheid aan merkers afgeleid, die in hoofdstuk 3 is toegepast om een totaal van 645 wilde zwijnen te genotyperen. Met de gegevens hiervan heb ik de genetische populatiestructuur geconstrueerd van het onderzoeksgebied, welke bestaat uit Nederland, Luxemburg en de aangrenzende delen van West-Duitsland. Zes geografisch verschillende wilde zwijnen populaties werden geïdentificeerd, en 25 recente hybriden (1-5 generaties geleden) van tamme varkens werden gevonden, verspreid over het hele onderzoeksgebied. Dit geeft aan dat introductie van gefokte wilde zwijnen wijdverbreid plaatsvind. De genetische populatiestructuur werd verder gevormd door een landschappelijke barrière (de Rijn) en historische herintroductie / translocatie projecten van wilde zwijnen door de mens.

In hoofdstuk 4 laat ik zien dat de PCV2 prevalentie bij wilde zwijnen per jaar varieert, maar ook stijgt met de leeftijd van wilde zwijnen en beïnvloed wordt door de genetische heterozygotie van de wilde zwijnen. Het leeftijdseffect is een gevolg van cumulatieve blootstelling aan het virus en een typische PCV2 eigenschap: niet detecteerbare intracellulaire persistentie van het virus, wat leidt tot een steeds grotere kans op PCV2 (her)infectie over de tijd. Het effect van genetische heterozygotie toont aan dat die wilde zwijnen die het meest een product zijn van inteelt, een kleinere kans hebben om PCV2 infectie overleven in hun eerste levensjaren. Dit laat zien dat PCV2 infectie kan fungeren als een

natuurlijke selectiekracht en dat infectie en sterfte door ziekte beïnvloed kan worden door genetische heterogeniteit.

In hoofdstuk 5 laat ik zien dat *Mhyo* prevalentie dezelfde jaarlijkse variatie vertoont als PCV2 prevalentie. Deze overeenkomst wordt waarschijnlijk veroorzaakt door klimatologische omstandigheden die de overdracht van ademhalingsziekten via aerosolen beïnvloedt. Daarnaast is opvallend dat hybride wilde zwijnen met een genetische invloed van het tam varken een hogere *Mhyo* antilichaam prevalentie hebben dan pure wild-type zwijnen. Dit geeft aan dat de hybride dieren vatbaarder zijn voor deze infectieziekte dan de wild-type zwijnen. Ik stel dat dit effect wordt veroorzaakt door overerving van genetisch gedegradeerde immuunfuncties van tamme varkens als gevolg van een geschiedenis van veterinaire zorg en gebruik van antibiotica.

Uiteindelijk breng ik de bevindingen van dit proefschrift samen in de Synthese (hoofdstuk 6). Hier reflecteer ik kort op de evolutionaire effecten van hygiëne en geneeskunde op menselijke immuunsysteem functies. Met betrekking tot de ziekte ecologie van wilde dieren ondersteunt dit onderzoek de hypothese dat genetische heterogeniteit van de gastheer een verklaring vormt voor observaties van aggregatie van ziekteverwekkers, een centraal thema binnen de epidemiologie. Bovendien, stel ik dat ons huidige begrip van de epidemiologie en ziekte ecologie van wilde dieren kan worden verbeterd door een integratie

van de populatie genetische en epidemiologische modellen. Samenvattend, laat dit proefschrift zien dat de combinatie van genetische en antilichaam-gegevens een krachtige aanpak is en dat genetische factoren en individuele heterogeniteit essentiële aspecten zijn voor onderzoek aan de epidemiologie en ziekte ecologie van wilde dieren.

Acknowledgements

Doing a PhD is a journey in the sense that you don't know where you will end up when you start. At least that was certainly true for me. During this process it is essential to get some guidance from other people in order to find the right track. And I am very thankful to all of them. For starters, off course, there are the supervisors and promotors, and I was happy to have two of both at the Resource Ecology Group (REG). I remember very well how co-supervisor Sip van Wieren helped me at the beginning to negotiate with disease monitoring institutes, and to build a sampling network in Germany. I received the most frequent and content-specific support from my supervisor Pim van Hooft, who always managed to come up with yet another obscure additional analysis for a particular genetic problem. Co-promotor Ron Ydenberg, although often overseas, provided good conversation and feedback every time I drafted a manuscript or joined a boat excursion in British Columbia. And finally, there was my promotor and godfather of all his extended phenotypes, Herbert Prins, who made the coffee breaks unforgettable with rampant discussions on politics, history or any other random topic.

I would also like to mention Hendrik-Jan Megens from the Animal Breeding Genomics Group at this point, who supported me immensely in learning the bioinformatics behind the universe of high density SNP data and who never failed to be contagiously enthusiastic about my work.

Then there is the family of PhD candidates and postdocs, as well as remaining staff that created the pleasant atmosphere at REG. As a PhD candidate 'in between generations' I have seen many, and will provide a surely non-comprehensible list here: Alfred, Anil, Audrie, Bas, Benson, Cornelis, Dulce, Edson, Eduardo, Edward, Emmanuel, Farshid, Frank, Fred, Gerda, Helen, Henjo, Herman, Ignas, Iris, Jasper, Jente, Joost, Kyle, Lennart, Mikhail, Milena, Ntuthuko, Patricia, Patrick, Priya, Ralf, Robert, Rudy, Tessema, Tim, Tom, Tsewang, Vincent, Xavier, Yolanda, Yong and Zheng. Thanks to all. Robert Kraus deserves special mentioning as my first roommate. He provided a well-appreciated German sense of humour, many practical tips for me as a starting PhD candidate and left me the famous duck-machine. And then Lennart

Suselbeek, the ‘nuchtere Drent’, who was a worthy successor as my roommate with his perpetual smile.

I would also like to mention my MSc Students: Maaike de Jong, Thijs Bosch and Katharina Langenbeck. Thanks to all of you for your discussions, your dedication and the perseverance to master all the required lab and computer skills as ecology students. I wish you good luck with your further careers.

My parents, Jan Goedbloed and Marloes Wilbers, have always been a most loving and sincere inspiration for me. Your loving care and the freedom you created for me have largely made me what I am today. It was my father Jan, who was the first to teach me with deep respect about the beautiful world of life forms called biology. An eye-opening experience that will always to be part of my life. My comrade Jannes, who survived all those years with me as an older brother, is very dear to me.

And finally, I come to my wife Danielle. You supported and motivated me more than anyone else the last 10 years. You enabled me to survive the difficult episodes and somehow managed to put up with my shortcomings. No words can do right to what I feel for you. Thank you for everything you did and for simply being who you are. I love you and I want to make you shine, in freedom, together.

A heartfelt thanks to all of you.

Curriculum Vitae

Biography

Daniël Johannes Goedbloed was born on June 12, 1982 in Middelburg, the Netherlands. He attended VWO secondary education at CSW Middelburg, from which he graduated in 2000. He subsequently enrolled in the study Biology at Wageningen University. Extracurricular activities included security coordination and board membership at youth association Unitas as well as field surveys for Dutch Butterfly Conservation. He soon started to specialise in molecular ecology, completing MSc research projects on mitochondrial recombination in fungi, habitat requirements of a Dutch endangered butterfly and phylogeography of an endemic South African lizard. In November 2007 he received his MSc degree, and immediately started to work as a research assistant at the Laboratory of Genetics in Wageningen, where he conducted independent molecular and biochemical research on the model organism Podospora anserina as part of the EU IP project 'the role of Mitochondria in conserved Mechanisms of Ageing' (MiMAge). In May 2008 he seized the opportunity to start a PhD project on disease ecology of European wild boar at the Resource Ecology Group in Wageningen, which led to this thesis. This work included proposal writing, setting up a sampling network, molecular laboratory analysis, bioinformatics, statistical analysis, academic teaching assistance and guest lectures, supervision of MSc student projects and publication of results in peer-reviewed journals. During his PhD Daniel was a committee member of the monthly seminar series 'Wageningen Ecology and Evolution Seminars' (WEES) and attended several international conferences as well as giving various invited talks. He will continue to work in science as a postdoc researcher at the University of Braunschweig, Germany, where he will investigate the role of gene expression in habitat adaptation by fire salamanders.

Publications

P Alexandri, H-J Megens, RPMA Crooijmans, MAM Groenen, **DJ Goedbloed**, JM Herrero-Medrano, L Rund, LB Schook, E Chatzinikos, K Triantafyllidis, A Triantafyllidis. (2013) Distinguishing historical from recent migration: insights from combined microsatellite and genome wide single nucleotide polymorphism data in a large mammal population. (submitted)

L Iacolina, M Scandura, **DJ Goedbloed**, RPMA Crooijmans, M Apollonio, G Larson, P Alexandri, A Archibald, LB Schook, MAM Groenen, H-J Megens. (2013) Effects of isolation and human-mediated introgression in shaping the genomic distinctiveness of an insular large mammal. (submitted)

DJ Goedbloed, P van Hooft, H-J Megens, K Langenbeck, W Lutz, RPMA Crooijmans, SE van Wieren, RC Ydenberg, HHT Prins. (2013) Reintroductions and genetic introgression from domestic pigs have shaped the genetic population structure of Northwest European wild boar. BMC Genetics doi:10.1186/1471-1-2156-14-43

DJ Goedbloed, H-J Megens, P van Hooft, JM Herrero-Medrano, W Lutz, P Alexandri, RPMA Crooijmans, MAM Groenen, SE van Wieren, RC Ydenberg, HHT Prins. (2013) Genome-wide SNP analysis reveals recent genetic introgression from domestic pigs into Northwest European wild boar populations. Molecular Ecology doi:10.1111/j.1365-294X.2012.05670.x

AD van Diepeningen, **DJ Goedbloed**, SM Slakhorst, AM Koopmanschap, MFPM Maas, RF Hoekstra, AJM Debets. (2010) Mitochondrial recombination increases with age in Podospora anserina. Mechanisms of Ageing and Development doi:10.1016/j.mad.2010.03.001

AD van Diepeningen, MFPM Maas, DHEW Huberts, **DJ Goedbloed**, DJP Engelmoer, SM Slakhorst, AB Koopmanschap, F Krause, NA Dencher, CH Sellem, A Sainsard-Chanet, RF Hoekstra, AJM Debets. (2010) Calorie restriction causes healthy life span extension in the filamentous fungus Podospora anserina. Mechanisms of Ageing and Development doi:10.1016/j.mad.2009.12.002

PE&RC PhD Training Certificate

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of literature (6 ECTS)

- Literature review for wild boar population genetics project (2008)

Writing of project proposal (4.5 ECTS)

- Parasite transmission in the European wild boar, *Sus scrofa*: associations with population structure and immunogenetic diversity (2008)

Post-graduate courses (2.7 ECTS)

- Introduction to R for statistical analysis; PE&RC (2008)
- Bayesian statistics; WGS (2009)
- Consumer resource interactions (DDD); PE&RC (2010)

Deficiency, refresh, brush-up courses (3 ECTS)

- Laboratory Animal Science (2009)

Competence strengthening / skills courses (2.4 ECTS)

- PhD Competence assessment; WGS (2008)
- Presentation skills; WGS (2009)
- Scientific writing; WGS (2011)

PE&RC Annual meetings, seminars and the PE&RC weekend (2.1 ECTS)

- PE&RC Days (2008-2012)
- PE&RC Weekend (2008)

Discussion groups / local seminars / other scientific meetings (7.5 ECTS)

- Ecological Theory & Application (2008-2012)
- Wageningen Evolution and Ecology Seminars (2009-2012)
- Ecogenomics discussion group (2009-2012)

International symposia, workshops and conferences (3 ECTS)

- International wild boar symposium (2008)
- European wildlife disease association conference (2009)
- European Society for Evolutionary Biology (2011)

Lecturing / supervision of practical's/ tutorials; (3 ECTS)

- Wildlife Resource Management (2009)
- Ecological methods (2009-2010)
- Ecology I (2009-2010)

Supervision of 3 MSc students

- Wild boar habitat suitability analysis
- Wild boar landscape genetic analysis
- Wild boar spatial epidemiology

Cover picture by Tim Hofmeester.
Summary cartoons by Daniel Johannes Goedbloed.

The research described in this thesis was financially supported by the KNJV (Royal Dutch Hunters Association) and Stichting De Eik.

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Printed by Ipskamp Drukkers.