UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di *Medicina animale, produzioni e salute (MAPS)*

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE VETERINARIE

CICLO XXVII

Further insight into the molecular epidemiology and evolutionary dynamics of rapidly evolving RNA and ssDNA viruses

Direttore della Scuola : Ch.mo Prof. Gianfranco Gabai

Supervisore : Ch.mo Prof. Michele Drigo

Dottorando : Dr. Giovanni Franzo

CONTENTS

ITALIAN SCENARIO

Agriculture plays an important role in the Italian economy. More than one million people are employed in activities directly related to agriculture and, in the last 10 years, this sector has produced goods with an average value of €50 billion. More specifically, animal productions account for more than €15 billion without taking into account all of the linked industries that further contribute to the richness of this sector. From an economic point of view, the swine and poultry industries (both meat and egg production), together with bovine farming, have been the most relevant animal production industries of the new millennium (Figure 1) (INEA, http://www.inea.it/banchedati).

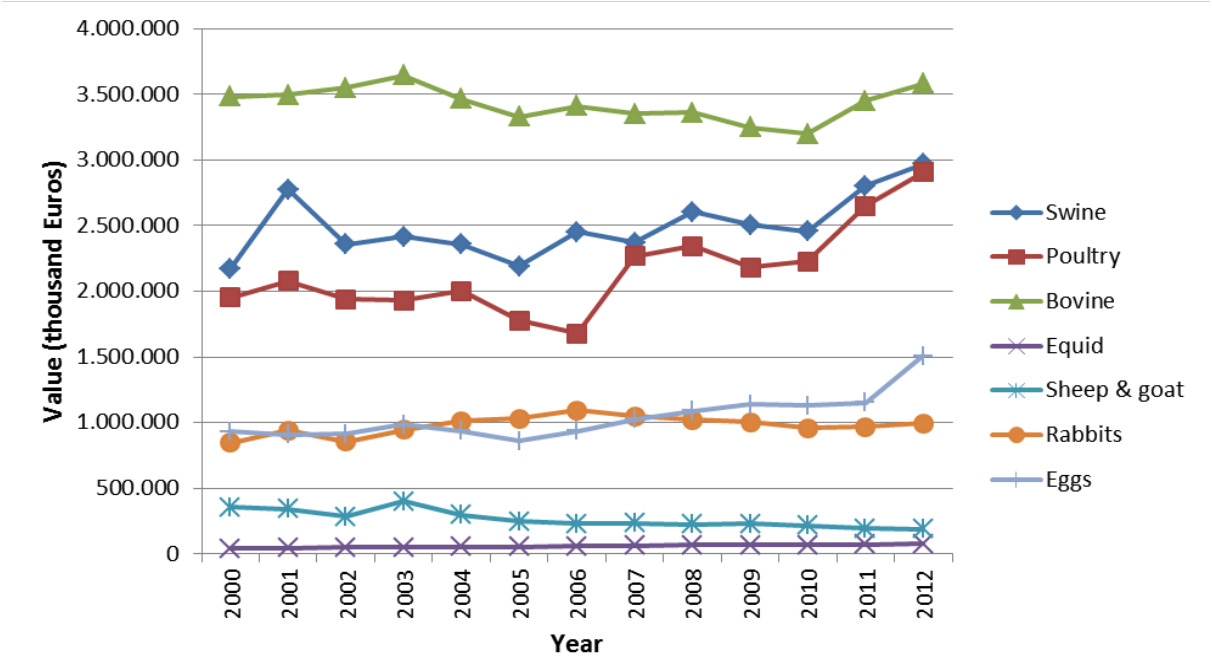

Figure 1. Economic value of animal productions in Italy.

Swine and poultry represent the species with the highest populations in Italy. An average of 9 million swine (Figure 2) are raised every year in our country, producing meat with an average value of approximately €3 billion that is intended for both consumption as fresh meat and for the production of Italian cured meats. The size of the poultry population is far higher. In 2013, more that 160 million poultry animals (Figure 2), including chickens and turkeys, were raised, leading to meat and egg productions with values of approximately 3 and €1 billion, respectively (FAOSTAT, http://faostat3.fao.org/home/E; INEA, http://www.inea.it/banchedati).

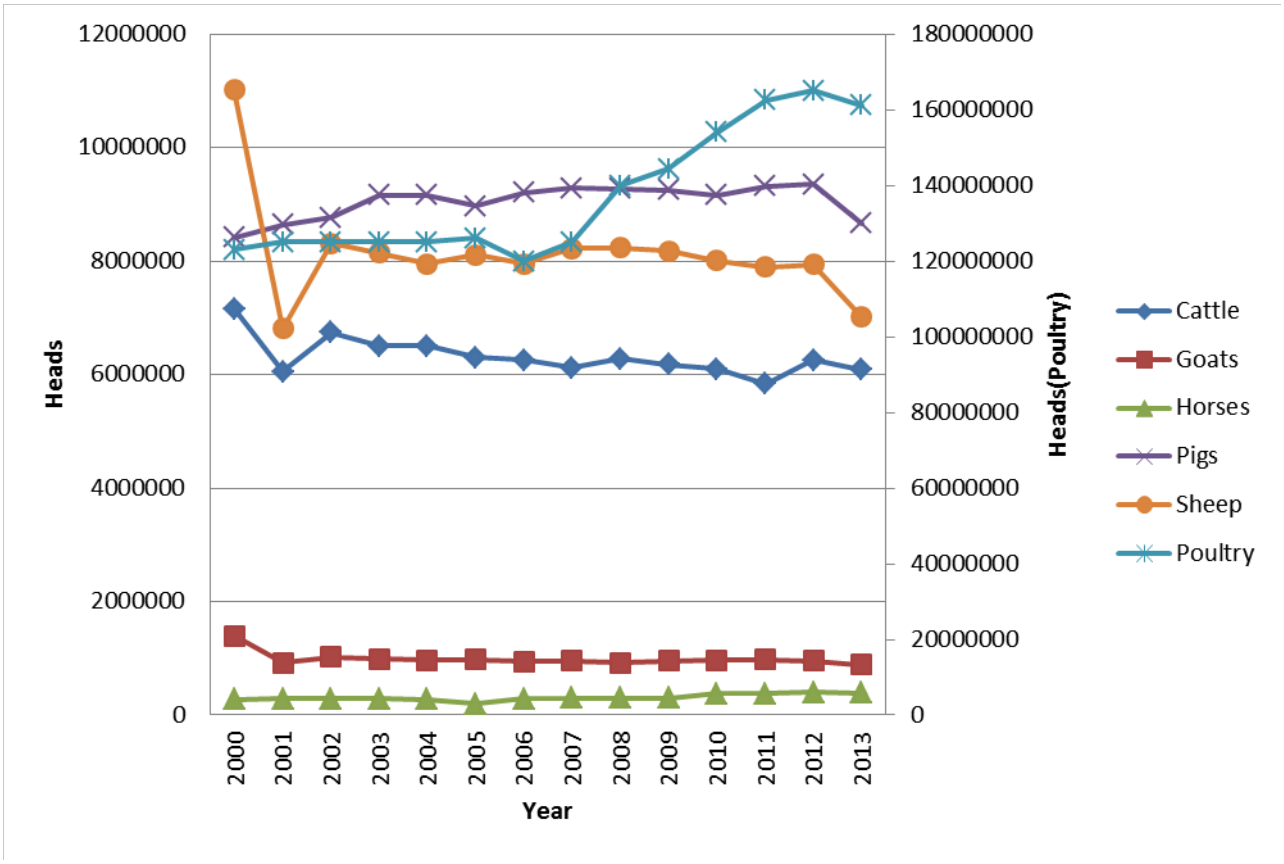

Figure 2. Number of live animals raised in Italy between 2000 and 2013 (FAOSTAT). Due to the incomparable order of magnitude, the poultry live animal value is reported on the right axis.

This marked production level, particularly considering the modest Italian territorial vocation to agriculture, is the consequence of a progressive industrialization and intensification of livestock production. The farming strategies have changed from traditional small-scale methods to industrial-scale, vertically integrated management. The need to maximize production efficiency and standardization has resulted in several changes in animal management and population composition. The current standard is represented by high-density, short-living, animal populations, typically concentrated in limited geographic areas and strongly interconnected by direct (e.g., multi-site integrated units) or indirect links. The rationale behind this organization can be identified by the easier optimization of the workforce, including both farm employees and specialized consultants, and in the possibility of obtaining a better and less-fragmented coordination of different farm productions. Similarly, the need to benefit from the facilities required to support an industrial production has led to the concentration of farms in the areas better

served by the major transport routes and the development of an industrial context that is able to provide raw materials and process livestock production. Obviously, the management of human and productive resources implies the existence of a dense network of animals, people and vehicles that move on and off farms (Leibler et al., 2010).

In addition, the international trade of live animals, germplasm and animal products has experienced a remarkable development (Commission on genetic resources for food and agriculture, 2007). The Italian importation and exportation of pigs and chickens and their products, even if with a fluctuating tendency, have shown a marked increase since the end of the Second World War (FAOSTAT, http://faostat3.fao.org/browse/T/TA/E). Globally, these aspects have surely facilitated the uniform and reliable production of consumer goods and contributed to the well-being and wealth of our country.

THE INFECTIOUS DISEASE THREAT

As will be further discussed in the following chapters, the previously described productive system has posed remarkable challenges for animal (and human) health (Graham et al., 2008) by creating favourable circumstances for the introduction and spread of new pathogens and their evolution. Infectious diseases, in addition to the pressing concern in terms of human and animal health and welfare (Mokdad et al., 2004; Buzby and Roberts, 2009), represent a major hazard for corporate incomes, causing direct losses (e.g., animal deaths, reduced production, reductions in food intake and conversion efficiency, etc.) and costs associated with disease control (McInerney et al., 1992; Morris and Marsh, 1992; McLeod, 1995; Bennett, 2003; Fonkwo, 2008).

Due to the variety of diseases and the different socio-economical contexts, general data on the impact of infectious diseases on the economy are limited. However, many studies have been conducted on several endemic and emerging diseases and in different countries, attesting the relevance of this issue (Neumann et al., 2005; Pritchett et al., 2005; Alarcon et al., 2013).

In fact, the major challenges for both avian and swine farming are represented by the so-called emerging diseases. Briefly, an emerging disease is one that "has appeared in a population for the first time or that may have existed previously but is rapidly increasing in incidence or geographic range" (WHO, accessed 2014- http://www.who.int/topics/emerging_diseases/en/). Currently, most emerging pathogens

affecting both humans and domestic animals are viruses (Jones et al., 2013). In particular, RNA viruses (and rapidly evolving ssDNA) viruses are mostly represented in this group (Woolhouse and Gowtage-Sequeria, 2005). Approximately 44% of emerging diseases in humans and 55% in livestock are caused by viruses, and RNA viruses have been reported to be associated with a higher relative risk (RR) to rise as an emerging disease than DNA viruses (Woolhouse and Gowtage-Sequeria, 2005).

Accordingly, state of art books dealing with swine and poultry diseases (Zimmerman et al., 2012; Swayne et al., 2013) dedicate particular relevance to RNA and ssDNA viruses, which representing approximately 80% of the described viral diseases. Such data, even if likely biased to pathogens that are more known and relevant in the context of developed/industrial farming, mirrors the overwhelming health and economic importance of this group of viruses. These pathogens represent a remarkable and rapidly evolving challenge due to their ability to adapt to new species, new host populations and new environments (Cleaveland et al., 2001; Domingo, 2010). The biological features of RNA and ssDNA viruses and the current global livestock farming scenario as well as more general environmental and social changes take part in defining the complex, and often unsolved, puzzle of swine and poultry diseases emergence and evolution.

VIRAL FACTORS

RAPIDLY EVOLVING RNA VIRUSES

RNA viruses have been demonstrated to possess an unrivalled evolutionary rate, which makes them both a great challenge for human and animal health and a fascinating opportunity to study the process of evolution in real time. It has been estimated that the mutation rates approach 10^{-2}-10^{-5} mutations·site^{-1}·replication^{-1}, corresponding to approximately 1 mutation·genome^{-1}·replication^{-1} (Duffy et al., 2008; Sanjuan et al., 2010), which is a rate that is 10-fold higher than that of retroviruses and 10000-fold greater than that of DNA viruses (Sanjuan et al., 2010) (Figure 3).

Figure 3. Average rates of spontaneous mutation in viruses adjusted to the rate per genome replication (Duffy et al., 2008).

The main causes of such a broad difference are mainly attributable to the lower fidelity of RNA-dependent RNA polymerase (RdRp) compared with DNA-dependent DNA polymerase (DdDp). Differently from RdRp, DNA polymerases are able to correct errors during DNA duplication, resulting in an intrinsically lower mutation rate. Complex organisms have also developed different DNA repair mechanisms, further reducing the extent of this phenomenon. In contrast, RNA viruses, with the partial exception of coronaviruses, are devoid of these systems (Minskaia et al., 2006). Retroviruses display an intermediate pattern because of their peculiar replication cycle. A relevant part of the viral cycle is spent

as a provirus, which is integrated in the host genome and subjected to the same mutation rate as the host genes (Kim et al., 1996; Svarovskaia et al., 2003). Additionally, it has been demonstrated that reverse transcriptase has a higher fidelity than RdRp, even if both enzymes are devoid of proofreading or base-excisions repair activities.

A great debate has developed regarding the evolutionary advantage behind the high mutation rates of RNA viruses. Because mutations are the ultimate source of genetic variation on which other forces, such as natural selection, act, it has been proposed that a high mutation rate is a strategy to rapidly explore a larger fitness landscape (Elena et al., 2000). This hypothesis is partially weakened by the evidence that the vast majority of mutations are deleterious. Studies on the Vesicular stomatitis virus have reported that approximately 40% of mutations are lethal, 30% are deleterious, 25% are neutral, and only 5% are beneficial (Sanjuan et al., 2004). A similar conclusion have been substantially confirmed for other pathogens both experimentally and using a bioinformatics approach (Escarmís et al., 1996; Yuste et al., 1999; Holmes, 2003b; Pybus et al., 2007; Bertolotti et al., 2008). However, the RNA virus population is often described as a quasispecies, i.e., an organized spectrum of genetic variants dominated by a master-consensus sequence (Domingo et al., 1996). Even if the actual existence of a quasispecies, as formulated by Eigen, remains a matter of discussion (Domingo, 2002; Holmes, 2009), it is well recognized that RNA viruses are characterized by huge populations subdivided in several sub-populations. This type of structure has been advocated to guarantee both mutational robustness and a continuous source of potential new solutions for adaption to new environments (Lauring et al., 2013). The high mutation rate of RNA virus could be considered a trade-off between generating sufficient beneficial mutations to adapt to an ever-changing environment (i.e., host immunity, new cells, news hosts, etc.) and avoiding the excess of deleterious mutations leading to the "mutational meltdown" (Manrubia et al., 2005; Holmes,. 2009).

Another, albeit not mutually exclusive, trade-off that potentially dominates RNA virus evolution is the competition between fidelity and replication efficiency (which is linked to replication fitness) (Wargo and Kurath, 2012). In other words, selective pressure would have simply favoured a rapid replication, which is also more error-prone. In fact, it has been demonstrated *in vitro* that an increase in RdRp fidelity is associated with areduced replication rates and is thus associated with a fitness cost (Furio et al., 2005). In addition, under experimental conditions, it is possible to select RNA polymerases with a lower

mutation rate. The absence of this phenomenon in natural populations strongly suggests that, even if a higher grade of fidelity is potentially possible, selective forces act against it (Belshaw et al., 2008).

Further support to this hypothesis originates from the cost-benefits associated with different replication strategies. When early genome copies are themselves used as templates for further replication (i.e., "geometric replication strategy"), mutations are predicted to accumulate geometrically. On the contrary, with "stamping –machine" replication, in which a single genome acts as a template for all replication cycles, the mutation load is expected to increase linearly. Nevertheless, these two stratagems are clearly correlated with the replication efficiency and speed such that the "geometric replication" strategy is anticipated to lead to a more rapid replication compared with the slower "stamping-machine" approach (Holmes, 2009).

Paradoxically, artificially elevating the mutation rate results in a fitness loss, suggesting that RNA viruses are close to their maximum tolerable mutation rates (Holmes, 2009). Chemical mutagens are currently used for the lethal mutagenesis-based therapy of viral diseases, such as Hepatitis C virus (HCV) and Respiratory syncytial virus (RSV). The rationale behind this approach is indeed to increase the mutation rates, reducing the viral population through the "error catastrophe" or the "extinction threshold" (Belshaw et al., 2008; Holmes, 2009; Lázaro, 2011). Similarly, some intrinsic host defences, such as APOBEC proteins, display a similar antiviral activity by increasing the mutational load (Mangeat et al., 2003). The inability of RNA viruses to tolerate a higher mutational load is also proved by the limited genome size of RNA viruses (i.e., lower than 30 kb). Even if a short genome has also been regarded as a strategy to allow faster replication (Elena and Sanjuan, 2005), it is mainly considered a constraint of the high mutation rate. Because the vast majority of mutations are harmful, viruses with longer genomes will suffer from higher deleterious mutations. This leads to the hypothesis, named the "Eigen paradox", that RNA viruses are in some ways trapped by their high mutation rates: the poor fidelity of RdRp poses that they are unable to evolve longer genomes, and the encoding of a polymerase with a higher fidelity would require longer genomes.

Interestingly, ssDNA viruses, which display the highest substitution rates among DNA viruses, are characterized by a genome length that is comparable to that of RNA viruses (Hoelzer et al., 2008; Firth et al., 2009). The understanding of the underlying causes of such rapid evolution is much more challenging for ssDNA viruses. These viruses rely on host

polymerases for their replication, and consequently, the lack of proofreading activity cannot be advocated. However, it has been proposed that both proofreading and excision repair mechanisms are less efficient in ssDNA viruses, potentially due to the different methylation patterns among the host and virus genomes (Duffy and Holmes, 2008). A deamination-dependent increase in the mutation rate has also been suggested (Duffy and Holmes, 2008).

This high mutation rates (i.e., rate at which mutations are generated in a genome) are substantially mirrored by high substitution rates (i.e., rates at which they are fixed at the population level) that characterizes RNA viruses and leads to their surprisingly rapid evolution and heterogeneity. A common misconception is to equate phenotypic and genotypic variability. Studies on RNA viruses have revealed a broad range of features in terms of phenotypic, particularly antigenic, variability. In fact, although many viruses are truly characterized by a great antigenic variability, which has resulted in an unresolved puzzle in terms of human and animal health, others are substantially stable. The possibility to control different diseases relatively easily through widespread vaccination activities, as in the case of distemper, rinderpest, rabies, poliomyelitis, just to mention some, has introduced the idea that some RNA viruses evolve slowly (Holmes, 2009). However, even if the rates of non-synonymous substitutions can vary largely among viruses (and among genes themselves), their impact on the global substitution rates is substantially limited because the vast majority of substitutions occur at synonymous sites (Holmes, 2003a).

In fact, despite some rare and still hardly debated exceptions, all RNA viruses are currently regarded to display a substitution rate between 10^{-2} and 10^{-5} substitution·site^{-1}·year^{-1} (Duffy et al., 2008; Pybus and Rambaut, 2009). It has been reported that this high value can be biased by several assumptions underlying the methods currently used for the estimation of substitution rates (Ho et al., 2007). For example, the assumptions of the absence of selective forces or that all of the observed nucleotide changes are actually fixed in a population hardly hold in the real-world scenario (Lemey et al., 2009), leading to an estimation that is potentially a combination of mutation and substitution rates. However, because the majority of deleterious mutations are rapidly purged from viral populations, their impact is thought to be substantially limited, and their estimation, at least in terms of their order of magnitude, is widely accepted by the scientific community (Holmes, 2009).

Understanding the multitude of forces and events as well as their interactions that shape substitution process represents one of the major challenges in modern virology. Even if

mutations are the ultimate source of genetic variation, the substitution rates are the product of their interaction with different factors, such as replication rate, effective population size (Ne), viral fitness and natural selection (Duffy et al., 2008). Viral populations are often described as a mutant swarm of related genotypes, i.e., a huge population with a spectrum of mutations dominated by a master sequence (Domingo et al., 1996). Such broad populations provide a double benefit. First, wide populations can be considered a strategy to obtain mutational robustness, meaning the capacity to display no phenotypic change despite genetic variability (Lauring et al., 2013). More complex organisms have developed several strategies to buffer mutation effects (e.g., gene duplication, protein function overlapping, presence of chaperone molecules, etc.). RNA viruses are, at the same time, characterized by the highest mutation rates (mainly deleterious) and unable to benefit from these strategies, mainly due to extreme genome compression. The generation of numerous progeny is thus regarded as a stratagem to guarantee that mutational robustness exists at the population level, despite the negative effects of mutations at the individual level (Elena, 2012). More simply, in a sufficiently broad population of related individuals, the maintenance of a certain percentage of unmuted progenies or mutants whose fitness is close to the major group is highly likely (Lauring et al., 2013). In addition, this population structure continuously provides variants that allow the rapid exploration of vast areas of the sequence space and consequently the fitness landscape. The destiny of these variants (i.e., of the mutation/s that they harbour) is subjected to two opposite forces: genetic drift (i.e., the stochastic fluctuation of the genotype frequency due to random sampling from generation to generation in a population) and natural selection (i.e., the differential survival and reproductive success of individuals with certain traits), which prevail on the other at different population sizes. Natural selection is expected to dominate when the product of the selection coefficient (s) and effective population size (Ne) is much greater than 1. The virus lifecycle is typically dominated by phases of the population size expansion alternated with severe bottleneck events, causing a marked reduction in the effective population size. Transmission events represent the most frequent source of genetic bottlenecks. Even if exceptions have been reported (Hughes et al., 2012), it is generally assumed that only few if not just one viral particle are transmitted during both vertical and horizontal transmission (Keele et al., 2008; Bull et al., 2011; Forrester et al., 2012). In this context, even if it is true that some viruses can display an actually higher transmission fitness (Wargo and Kurath, 2012;

Alizon and Fraser, 2013; Carlson et al., 2014), the stochastic sampling of genetic variants plays a major role in shaping the substitution dynamics. Because the vast majority of mutations are damaging, genetic drift allows fixation of many slightly deleterious (i.e., with a small selection coefficient) mutations that would have been purged by natural selection in larger population. Repetitive bottlenecks are associated with a progressive fitness loss due to the accumulation of deleterious mutations and the loss of the fittest mutants in a small asexual population, a phenomenon known as *Muller's Rachet* (Muller, 1932; Muller, 1964). This phenomenon has been repeatedly demonstrated to be true also for viral populations through *in vitro* and *in vivo* experiments (Chao, 1990; Escarmís et al., 1996; Carrillo et al., 2007), even if different patterns have been reported and viruses have probably evolved some strategies to counterbalance this phenomenon and withstand extinction (Lazaro et al., 2003; Manrubia et al., 2005).

When the population size or selection coefficient is high, natural selection is predicted to be the dominant force affecting the pattern of nucleotide substitution. Huge viral populations are typically predicted during the within-host virus lifecycle (when potentially millions or billions of viral particles are produced) (Stilianakis and Schenzle, 2006) and probably at the epidemiological level, even if several factors can affect the actual relationship between N and Ne at the inter-host level (Lloyd-Smith et al., 2005; Kilpatrick et al., 2006; Bansal et al., 2007; Holmes, 2009). In fact, it is hypothesized that viral populations are sufficiently large for natural selection to act efficaciously, even at the epidemiological scale (Holmes, 2009), particularly when large and well-mixed host populations allow frequent and potentially multiple transmission events (Drew, 2011). In this scenario, selective forces can purge deleterious mutations (negative selection) and provide a reproductive advantage (positive selection) to the mutants that are more adapted to the particular environment, leading to a progressive fitness increase. Many forces can have an effect on viral populations by affecting the individual's fitness. Albeit far from comprehensive, the following overview aims to describe the complexity of the factors, which continuously interact among them and affect the survival and expansion of particular strains. Viruses are also continuously engaged in an arms race with the host and other pathogens (as well as strains of the same population), which lead to a progressive fitness increase in both the "winners" and the "losers" relative to their ancestor (i.e., the Red Queen Hypothesis) (Clarke et al., 1994; Elena et al., 2000). This competition for limited resources, which

ultimately leads to competitive exclusion, is the key to understanding the extreme refinement of viral biology.

Probably the most known and exemplifying selective pressure is the ability of viruses to escape from the host immune response or from antiviral therapy. Variations in the environment could favour an increase in variability (Borrow et al., 1997; Ross and Rodrigo, 2002; Chen et al., 2005) (diversifying selection) or direct evolution toward a peculiar AA (Seoighe et al., 2007; Kosakovsky Pond et al., 2008) (directional selection) in certain regions, such as epitopic sites or active sites, that can prevent recognition by the immune response or drug efficacy. Similarly, infection of a new host, cell type or ecological niche leads to a progressive selection of fitter mutants (Elena et al., 2000; Lalic et al., 2011).

In opposite ways, purifying pressure acts to purge harmful variations, typically located within regions with structural and functional constrains (Hughes and Hughes, 2007; Teoh et al., 2012). The vast majority of methods currently used to predict individual sites under selection are based on the non-synonymous/synonymous substitution rate (dN/dS) (Kosakovsky Pond and Frost, 2005; Murrell et al., 2012). These methods are substantially based on the assumption that synonymous mutations have a negligible effect on phenotype and thus on fitness. Even if mutations are assumed to randomly affect the genome, those causing amino acid changes at sites with major functional constraints will be purged by natural selection, and thus, an excess of synonymous mutations compared with the number expected will be observed. In contrast, when amino acid changes represent an evolutionary advantage (e.g., immune escape), an excess of non-synonymous changes is expected. Although these methods are extensively used and their power has been reported (Holmes, 2009; Lemey et al., 2009), natural selection is not limited only to proteins but also acts on other components of the viral phenotype that are often more difficult to model. Obviously, non-coding regions can possess several functions, such as regulation of gene transcription and translation, genome encapsidation, etc. (Pasternak et al., 2001; Novella et al., 2004). Furthermore, it is well known that RNA molecules can form a complex secondary structure. RNA viruses are not an exception, and their genome may form a set of secondary structures that play an apparently relevant role in virus biology (Pasternak et al., 2006). This evidence enforces the idea that not all synonymous mutations are evolutionarily neutral and are able to affect short- and long-range RNA interactions. Codon bias is another relevant phenomenon that can severely affect viral fitness, even in the presence of synonymous mutations. Codon usage (CU) refers to the frequency of the occurrence of each codon for at

least two-fold degenerate codons (Hershberg and Petrov, 2008); namely, it is an indication of the 'preference' of a genome for one or more codons if more than one codon is possible for the same amino acid (Hershberg and Petrov, 2008; Brandao, 2013). It appears that the major cause for selection on codon bias is that certain preferred codons are translated more accurately and/or efficiently (Hershberg and Petrov, 2008). Because viruses are dependent on the host replicative machinery, it is conceivable that codon bias and nucleotide composition have been optimized by selective forces to match those of the peculiar tissue or host (Rabadan et al., 2006; Holmes, 2009; Liu et al., 2010; Brandao, 2013; Chen et al., 2014).

As previously mentioned, the high mutational burden has imposed the evolution of different strategies to maximize genome compactness, allowing the storage of sufficient information to support the viral lifecycle (Holmes, 2009). The presence of overlapping open reading frames (ORFs) is one of the most widespread stratagems (Walewski et al., 2001; Belshaw et al., 2007). As a consequence, changes in one nucleotide can be synonymous in one protein while drastically changing the amino acid in other ones or altering the regulatory sequences (Robinson et al., 2013). Understandably, these constraints further limit the genome flexibility.

Recombination is a well-known phenomenon among both DNA and RNA viruses and has been demonstrated to display a great variability in terms of frequency and patterns, even within the members of same viral family (Simon-Loriere and Holmes, 2011). Considering the issues of this thesis, the focus of this chapter is rapidly evolving RNA and ssDNA viruses. Despite the completely different natures of these two viral groups, their recombination mechanisms are shared or at least display major analogies. The most common and widely accepted model of RNA/DNA intramolecular recombination is 'copy choice' recombination. In this process, the polymerase, or the reverse transcriptase (RT) in retroviruses, switches from one genome molecule (the donor template) to another (the acceptor template) during synthesis while remaining bound to the nascent nucleic acid chain, thereby generating a genome with mixed ancestry (Martin et al., 2011; Simon-Loriere and Holmes, 2011). Less frequently, non-replicative forms of recombination have been reported for RNA viruses when the recombining RNAs are cleaved at specific points and ligated by different enzymes to form hybrid molecules (Gmyl et al., 1999; Gallei et al., 2004). Similarly, ssDNA viruses can undergo recombination mediated by host DNA repair enzymes, which allow fixing of the broken double-strand DNA intermediates via a homology-dependent mechanism (Martin et al., 2011). Another common form of recombination, which is properly called reassortment, is based on the exchange of complete genome segments in segmented RNA viruses (Greenbaum et al., 2012; Marshall et al., 2013). Moreover, something similar can occur in multicomponent ssDNA (e.g., nanoviruses) (Katul et al., 1998) and RNA (e.g., Bromoviridae and Comoviridae) (Simon-Loriere and Holmes, 2011) viruses or those having satellite molecules (e.g., Begomoviruses) (Martin et al., 2011).

Recombination has traditionally been considered a form of sexual reproduction in viruses that has been evolutionary advantaged because it carries two main benefits: a faster removal of deleterious mutations and an increase in the rate at which advantageous genetic combinations are produced (Keightley and Otto, 2006; Martin et al., 2006). However, this hypothesis is still highly debated, and the validity of the assumptions of these theories has been questioned (Holmes, 2009; Simon-Loriere and Holmes, 2011). RNA viruses (and potentially rapidly evolving ssDNA viruses) afford both the maintenance of an un-mutated viral progeny and the gaining of advantageous traits mainly due to their high

mutation rates and broad population sizes. It is highly probable that the mutation rates are normally so high in RNA viruses and that their population sizes are often so large at both the host and epidemiological scales that advantageous combinations of mutations are regularly generated without the assistance of recombination (Simon-Loriere and Holmes, 2011). The generation of a broad progeny should guarantee the maintenance of fit descendants through back mutations or compensatory mutations (Elena et al., 2006). Based on these considerations, from an evolutionary point of view, Holmes et al. proposed recombination as a byproduct of genome organization and of the viral lifecycle (Holmes, 2009; Simon-Loriere and Holmes, 2011). Nevertheless, independently from the in-depth causes underlying this phenomenon, it could be considered an effective way to increase genetic diversity and rapidly move through vast regions of the sequence space. Several studies have reported its role in the origin of new virus species (Briddon et al., 1996; Gibbs and Weiller, 1999; Jackwood et al., 2010), new host and cell tropism (Hofer et al., 1997; Qin and Petty, 2001), variation in virulence (Li et al., 2009), resistance to treatment (Nora et al., 2007) and evasion of host immunity (Malim and Emerman, 2001).

Different viral families (but also different genera and even species) are characterized by dissimilar recombination rates. Many factors affect the recombination occurrence. The most obvious is the epidemiological factor: for recombination to occur, different viruses have to share the same geographic area and host and cell tropism (Martin et al., 2011). Extensive viral spread and frequent mixed infection obviously increase the recombination frequency (Franzo et al., 2014a). As previously mentioned, genome organization is deeply involved in the determination of recombination occurrence and localization. Most frequently, recombination mediated by a copy choice mechanism is favoured by the sequence similarity between the nascent and the acceptor nucleic acid molecules (van Vugt et al., 2001; Magiorkinis et al., 2003). The secondary structure is also considered to play a role in both RNA and ssDNA virus recombination by stalling the RNA/DNA polymerase during replication and facilitating the transfer of the polymerase-nascent nucleic acid molecule complex onto the acceptor RNA/DNA (Lai. 1992; Rowe et al., 1997; Leppik et al., 2007; Shen et al., 2009; Simon-Loriere et al., 2010). Accordingly, the origin of replication (v-ori) of circular ssDNA viruses has been demonstrated to be a recombination hot-spot because, in addition to being the site where viral genomes are released from genomic concatemers (Lefeuvre et al., 2009), it forms a hairpin structure.

The marked difference in the recombination frequency between negative- and positive-sense RNA viruses poses a focus on the role of the lifecycle and replication strategies. In fact, although (+)ss RNA viruses have been reported to have a high recombination rate (Lai, 1996; Savolainen-Kopra and Blomqvist, 2010), this phenomenon is typically rare in (-) ssRNA viruses. This difference has been reported as a byproduct of the different approaches for protein-translation-levels optimization (Simon-Loriere and Holmes, 2011). Because the (+)ssRNA genome itself is a functional mRNA, other means different from transcription regulation have evolved. *Coronaviridae*, as well as other (+)ssRNA viruses with analogous replication strategies, attest the linkage between the extremely high recombination frequency and viral lifecycle (Pasternak et al., 2006). The regulation of gene expression is mediated by the production of sub-genomic RNAs (sgRNAs). The synthesis of these sgRNAs is based on discontinuous transcription, which relays on the template switching capacity of RdRp (i.e., a copy choice mechanism) (Pasternak et al., 2006). It is plausible that the RdRp of these viruses was selected to efficiently mediate template-switching events and that the very high rates of recombination observed are a direct consequence of this particular strategy for controlling gene expression (Simon-Loriere and Holmes, 2011). Interestingly, transcription strategies appear to condition also the recombination rate of ssDNA viruses with an ambisense circular genome, such as Geminiviruses and Circoviruses. Clashes between transcription and replication complexes have been proposed to justify the evidence that members of these families tend to display a higher recombination frequency in their complementary sense genes than they do in their sense genes (Martin et al., 2011). The interaction between nucleic acids and proteins also alters the genome availability. The genomes of both (-)ssRNA and ssDNA viruses form complexes with nucleoproteins and host histones during replication, respectively, and are consequently preserved from recombination by lowering the probability of hybridization between complementary sequences (Simon-Loriere and Holmes, 2011). Accordingly, for members of the genus *Begomovirus*, recombination hot spots have been linked to an exposed region of the mini-chromosomes, potentially because these are hyper-sensitive to either physical breakage or host nuclease attack (Jeske et al., 2001; Pilartz and Jeske, 2003). Similarly to mutations, recombination events may lead to either advantageous or deleterious fitness effects, which determine the reproductive success over parental strains. Viral genome and proteins are the results of a continuous optimization of the interactions among different portions of the genome, different proteins and different domains of the

same protein. Many recombination events are associated with a major fitness cost due to the destruction of such an interaction network (Li et al., 2008). Accordingly, Lefeuvre et al. (2009) reported some trends that are in substantial agreement with this hypothesis. In ssDNA viruses, recombination is more frequent in the inter-genic regions or at the periphery of the genes, where they are less likely to disrupt protein folding. Similarly, the lower frequency of recombination within genes encoding structural proteins is probably related to the higher net of interactions between these proteins (e.g., protein constituting the capsid) compared with non-structural ones.

RAPIDLY EVOLVING VIRUSES IN MODERN FARMING

During the last century, farming has undergone a radical change that, starting at the so-called developed counties, is now involving developing ones with impressive energy (Thornton, 2010). This growth is driven by the rapidly increasing demand for livestock products determined by population growth and increasing incomes in developing countries (Delgado, 2005). Forecasts for the future estimate that the demand will increase to more than double in the following fifty years (Thornton, 2010). This livestock revolution has been compared to Henry Ford's mass production line based on the aim of increasing the total yield, efficiency and profitability through intensification and industrialization of the productive system (Pappas, 2013; Perry et al., 2013). The emerging farming strategies, in addition to the obvious socio-economic implications that are beyond the scope of this thesis, provide a totally new environment for pathogens, particularly RNA and ssDNA viruses, to evolve. Traditional livestock-dependent small-holders and pastoralist systems characterized western countries in the previous centuries and continue to exist in many underdeveloped as well as in vast regions of developing countries. These systems provide a picture dominated by low density/extensive farming (Perry et al., 2013). From a virus point of view, the limited contact network and the small number of animals involved favour a more severe bottleneck, reduce mixed infections and determine a low Ne, at least at the epidemiological level. Repeated bottlenecks, such as those that characterize small populations with rare transmission events, have been suggested to lead to the accumulation of mainly deleterious genetic substitutions through genetic drift, resulting in a fitness loss due to "Muller's ratchet" effect (Chao, 1990; Carrillo et al., 2007). Reduced viral fitness is typically mirrored by a reduced replicative capacity and, potentially, reduced virulence (Wargo and Kurath, 2012). In contrast, the current intensive farming productive system, in which thousands of animals are in strict contact, provides the basis for both multiple infections with variants of the same viruses (reducing the transmission-associated bottlenecks) and for the development of a huge viral effective population size in which natural selection can act, favouring a fitness increase (Elena et al., 2000; Drew, 2011). Similarly, the long-range spread of highly heterogeneous strains, often characterized by limited cross-protection, allows a high Ne at the epidemiological scale, whereas compresence and competition

among different strains can also be advocated to further foster a fitness increase (Elena et al., 2000).

In addition, the presence of multi-site integrate units coupled with flourishing national and international trades increases the frequency of the co-infection of different strains, which may lead to an increased likelihood of recombination (Franzo et al., 2014a). On the host side, germplasm trade is leading to a trend characterized by animal “monoculture” (Drew, 2011). Genetic homogenization poses relevant questions regarding how these populations may respond immunologically to a novel pathogen (Commission on genetic resources for food and agriculture, 2007; Pappas, 2013; Segales et al., 2013).

Even if undoubtedly simplistic and somewhat speculative, this dichotomy provides a fascinating overview of the challenges that institutions aimed at animal heath control have to and will have to face. All of these factors bring with them both opportunities and challenges as well as theoretical and practical issues that must be assessed to develop a more aware approach to the control of livestock infectious diseases and their associated losses.

The understanding of viral biology, evolution and epidemiology is fundamental to providing the basic knowledge required to plan adequate control strategies and to verify their effectiveness (Holmes and Grenfell, 2009). In this sense, the developing field of phylodynamics is acquiring great importance (Grenfell et al., 2004). This broad field is aimed to the linkage and merging of studies of immunodynamics, epidemiology and evolutionary biology that shape the spatiotemporal and phylogenetic patterns of infectious diseases. Understandably, to gain knowledge regarding these issues, different levels of viral evolution must be assessed: from the micro-level of intra-host evolution to the macro-level of pandemic spread (Norstrom et al., 2012) (Figure 4).

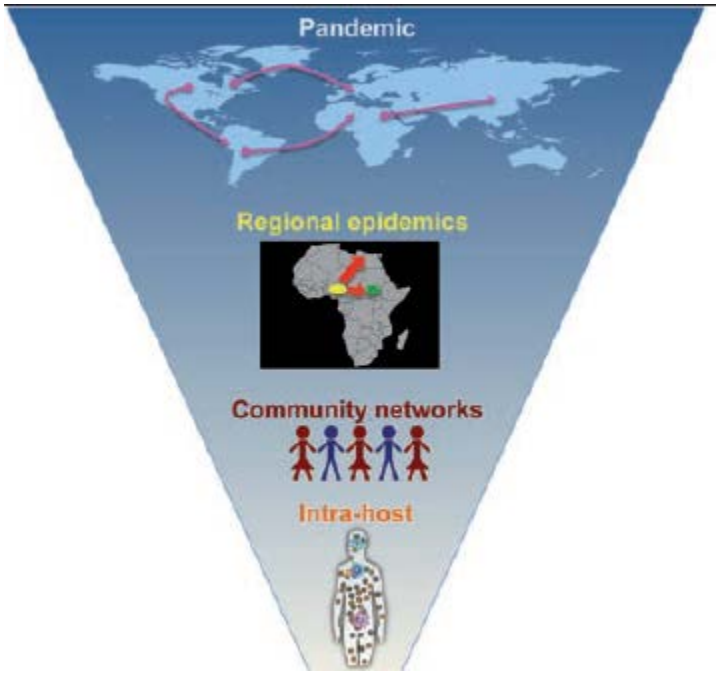

Figure 4. Different levels of phylodynamics inferences ranging from the intra-host viral evolution to the worldwide spread patterns. From Norstrom et al., 2012.

The increasing availability of sequences coupled with advances in the application of mathematical modelling and computational resources has greatly contributed to the progress in this field. The coalescent theory (Kingman, 1982) has greatly contributed to the linking of phylogenetics to the study of population structure and dynamics. Briefly, relations among *n* lineages from an *N*-sized population is viewed in terms of coalescent events from present to past between lineage pairs until only one lineage remains (i.e., Most Common Recent Ancestor; MRCA). The likelihood of two lineages to coalescence at

different time points is dependent on the population structure and on its variation (with small populations having a faster coalescence rate compared with bigger ones), a phenomenon that leaves traces in the tree topology (Volz et al., 2013). The progressive accumulation of genetic variation can be used to estimate, under the molecular clock assumption, the coalescent time since MRCA. Even if the initial formulation of Kingman's coalescent was founded on several assumptions (e.g., a panmictic haploid population, a constant population size, no selective forces acting on the population, absence of recombination, etc.) (Lemey et al., 2009), several extensions have been developed over time. One of the most relevant for its application to Measurably Evolving Populations (MEP) (i.e., a MEP is any population that permits the detection of a statistically significant accumulation of substitutions when samples are obtained at different times) (Drummond et al., 2003) is the serial-coalescent (Crandall, 1999; Rodrigo and Felsenstein, 1999; Lemey et al., 2009), able to deal with sequences obtained from samples collected at different times. This allow a simple estimation of mutation rates scaled in chronological time by estimating the number of mutations at each time interval and dividing by the amount of time between the samples. In addition, it allows the increase of lineages moving back in time, buffering the progressive effect of lineages reduction through coalescent process and increasing in this way our capability of making statements about population parameters and dynamics (Lemey et al., 2009).

Further extensions have allowed progressive reductions in the gap between model assumptions and real populations. The imposition of a fixed evolutionary rate among all lineages, an assumption that is hardly suitable for viral population, has been outdated by the development of relaxed molecular clocks, which allow rate variations among lineages (Drummond et al., 2006). Similarly, the assumption of a constant population size has been progressively relaxed. First, variations were described by parametric models (e.g., logistic, exponential growth, etc.), but more complex and flexible non-parametric "skyline plot"-like models have subsequently been developed, and these are able to address one (Drummond et al., 2005; Minin et al., 2008) or multiple unlinked loci (Heled and Drummond, 2008; Gill et al., 2013). These 'skyline plot' methods enable the estimation of historical patterns and variations in population size from a genealogy without the need of *a priori* restrictions (Ho and Shapiro, 2011). It should be mentioned that the actual value estimated by a skyline plot is a composite value, which is also called relative genetic diversity ($N_e \cdot t$, where "t" is the generation time). This has posed some doubts regarding the interpretation of this value for

infectious diseases and the relationship with the number of infected individuals (i.e., prevalence) (Frost and Volz, 2010). More sophisticated models linking classical epidemiology to phylogenesis-based models have been and are being developed, even if their implementation in publically distributed software is still at the beginning stages (Koelle and Rasmussen, 2012; Stadler et al., 2013; Kuhnert et al., 2014). Nevertheless, despite these limitations, coalescence-based population dynamics reconstruction has been successfully used to infer the epidemic patterns of different human (Lemey et al., 2004; Rambaut et al., 2008; Magiorkinis et al., 2009) and animal (Firth et al., 2009; Kerr et al., 2009; Shi et al., 2010) diseases, and this approach has been demonstrated to be reliable for the investigation of the potential causes of drastic changes in epidemiological patterns and the effectiveness of control strategies (Pybus et al., 2003; Murcia et al., 2011; Alfonso-Morales et al., 2013). The development of phylogeography reconstruction, based on robust mathematic modelling (Lemey et al., 2009; Lemey et al., 2010), has provided additional tools, alongside classical epidemiology, for inferring spatiotemporal links and trends in the history of infectious diseases (Faria et al., 2011; Lo Presti et al., 2012; Alfonso-Morales et al., 2013).

Similarly, many models and approaches have been developed to estimate the selective forces that shape gene evolution, and their application is commonplace in computational virology. Methods based on the estimation of the ratio between non-synonymous and synonymous substitution rates (dN/dS) are currently the most used, even if they can model only one component of the forces affecting viral fitness. Briefly, these methods are based on the assumption that selection acts less strongly on synonymous mutations than on non-synonymous ones. Values of dN/dS lower than 1 are attributable to a fitness cost associated with an amino acid change such that only synonymous mutations can spread and reach fixation, whereas non-synonymous mutations are purged from the population (purifying or negative selection). In contrast, dN>dS suggest an evolutionary advantage associated with diversification (diversifying or positive selection). The first mathematical models dealt in particular with pervasive diversifying/purifying selection, at both the global and single site levels (Kosakovsky Pond and Frost, 2005; Tamura et al., 2013). Other models have, in addition to improving computational efficiency (Murrell et al., 2013), progressively increased the range of the modelled selective patterns. In addition to diversifying selection, in which any nucleotide substitution that changes an amino acid is favoured, methods have been developed to estimate directional selection (Kosakovsky

Pond et al., 2008; Murrell et al., 2012), in which only substitutions toward a reduced number of amino acids are favoured. Another refinement is the implementation of models that are able to address pressures, including both diversifying (Murrell et al., 2012) and directional (Murrell et al., 2012), that change over time or act on a single or few lineages (i.e., episodic selection). The evidence that adaptive evolution often occurs in episodic bursts, i.e., on a limited amount of sites of a reduced number of different lineages, has prompted the formulation of "branch-site" models that are able to account for both of these components at the same time (Kosakovsky Pond et al., 2011; Murrell et al., 2012).

As briefly summarized in the previous chapters, large-scale evolutionary patterns are the final product of many interacting forces. Nevertheless, the starting point of this complex phenomenon is represented by the intra-host, or even intra-cell, dynamics, which provide the fundamental substrate on which evolutionary forces act. Remarkably, different viruses display highly variable intra-host patterns depending on the biological cycle and relation with the host. Understandably, most of the available information originates from chronic human diseases. Human immunodeficiency virus (HIV) infection represents an illuminating example of the various patterns that characterize different evolutionary scales. The HIV evolutionary rates are slower at the epidemiological than at the intra-host level, probably because of the higher transmission fitness of slower-evolving lineages compared with that of the more rapidly evolving ones. Potentially, the arms race between virus and host, which causes a progressive fitness gain within a specific host environment, comes at a major cost in terms of transmissibility to other individuals (Pybus and Rambaut, 2009). These predictions are confirmed by the phylogenetic structure of HIV at the population level, which reflects the demographic and spatial history instead of immune selection. In contrast, evidence of evolution driven by immune pressure is found in within-host phylogenies, which, not surprisingly, resemble those obtained from influenza epidemics (Grenfell et al., 2004). Far from being exhaustive, this simple example provides a foretaste of the implications of a thorough investigation of the viral population composition. The study of the intra-host mutant swarm has been traditionally based on the Sanger sequencing of cloned PCR products. This approach, which is still widely used, has the advantage of providing an almost perfect variant reconstruction and has allowed the study of many aspects of viral biology, ranging from bottleneck structure (Murcia et al., 2010; Bull et al., 2011; Stack et al., 2013) to the selection of drug-resistant mutants (Baxa et al., 2013), from antigenic escape (Honegger et al., 2013) to the role of viral diversity in an

increasing host range (Hoelzer et al., 2010; Ubol et al., 2011). Nevertheless, the relative paucity of data, which are substantially limited to some major human diseases, is mainly due to the practical difficulties associated with this type of studies. Clonal sequencing is not readily automatable and is cumbersome in terms of bench work and cost (Prosperi et al., 2013). Recent advances in nucleic acid sequencing technologies, which are broadly referred to as Next-Generation Sequencing, have opened new prospects for virological research and diagnostic activities (Capobianchi et al., 2013). Holmes and Grenfell stated that "Just as PCR and first-generation DNA sequencing ushered in the science of molecular epidemiology, so next-generation sequencing may herald the age of phylodynamics" (Holmes and Grenfell, 2009). The possibility of obtaining million to billions of reads of known and unknown samples in a more rapid and more inexpensive manner compared with Sanger sequencing is radically changing our approach to virology (Glenn, 2011). The most obvious transformation is the increased ease of obtaining full-genome sequences through both re-sequencing or *de novo* sequencing (Marston et al., 2013; Ladner et al., 2014). NGS has allowed viral metagenomics to be analysed at a previously unseen resolution, boosting the discovery of new viral candidate pathogens (Briese et al., 2009; Blomstrom, 2011; Sachsenroder et al., 2012; Belak et al., 2013; Tan le et al., 2013) and allowing remarkable advances in virus ecology (Edwards and Rohwer, 2005; Ng et al., 2011; Hurwitz et al., 2014). Transcriptome analysis trough NGS has been used to investigate the expression of both viral and host genes as well as their timing and interaction (Yang et al., 2010; Yang et al., 2011). All of these fields and many others will undoubtedly contribute to our understanding of viral ecology and evolution (Barzon et al., 2011). However, the field that potentially gained the biggest advantages from these new technologies is, precisely, the study of viral mutant clouds. It is generally accepted that NGS can detect the minority of variants present in approximately 1% of sequence reads (Radford et al., 2012), replacing and outperforming previous strategies. In addition, several studies are now considering the whole genome diversity, an approach that will likely lead to the generation of new insights into the pathogenetic role of the variability of previously neglected genes (Bimber et al., 2010; Henn et al., 2012). As most technologies, particularly those under development, NGS displays some limitations, which are mainly due to the short read length and the unavoidable presence of errors due to the sample preparation and sequencing procedures (Beerenwinkel and Zagordi, 2011). This issue is particularly relevant when the aim of the study is to investigate the minority single nucleotide variation

(SNV) present at a frequency comparable to the instrument error rate. Many approaches and algorithms have been developed to address this problem, and these allow a reduction in the overall error rate by a factor of at least 2 to 20 (Beerenwinkel and Zagordi, 2011; Yang et al., 2013). Far more challenging and still substantially unresolved is the attempt to assemble global haplotypes and estimate their frequency. As a matter of fact, genetic diversity is currently studied based on the detection of SNVs rather than on the reconstruction of linked mutations (Beerenwinkel et al., 2012). Notable efforts have been made in this direction, and several methods have been developed based on different approaches (Beerenwinkel et al., 2012). Nevertheless, even if encouraging results have been obtained in terms of the estimation of the global trends in the distribution of viral haplotypes, reliable methods providing both sensitivity and precision have not yet been developed, and the evaluation of the trade-off between these two aspects remains one of the major concerns in this emerging and fascinating field (Prosperi et al., 2013).

In addition to a mesmerizing opportunity to study evolutionary processes in real time, rapidly evolving RNA and ssDNA viruses are, above all, a great and often unsolved challenge for human and animal health (Cleaveland et al., 2001; Morens and Fauci, 2013).

Each control or prevention strategy must be founded on the availability of accurate and rapid diagnostic tools that allow the continuous surveillance and updating of the epidemiological scenario to guarantee the possibility of implementing effective interventions.

Even if there is a broad choice in the selection of diagnostic assays, most of the currently used diagnostic protocols are based on the detection of nucleic acids through polymerase chain reaction (PCR) and their variants or on the detection of antigen-antibody interactions. It is impossible to provide a detailed summary of these tests' features and performances, which are highly variable and dependent on the specific assay and pathogen. However, in an attempt to identify some common factors, their wide application is mainly associated with their sensitivity, low cost and ease of automation. Diagnostic assays based on molecular biology are now widely used in the diagnosis of infectious diseases in human and veterinary medicine (Belák, 2007; Balka et al., 2009; Hoffmann et al., 2009). High sensitivity and specificity, coupled with the rapidness of the diagnostic response, prompt their use as a diagnostic tool for both pathogen identification in single individuals and pathogen surveillance at the population level. Different strategies based on PCR and real-time PCR have been developed to simultaneously detect, quantify and genotype one or more viral pathogens (Brunborg, 2004; Cecchinato et al., 2013; Cella et al., 2013; Drigo et al., 2014a; Listorti et al., 2014). Nevertheless, even if the high specificity of primers and probes binding to target sequences has been widely used for subtyping, this feature also represents one of the main weaknesses of these assays. The high evolutionary rate coupled with intense national and international trades constantly expose diagnostic laboratories to the risk of false-negative results (Mackay et al., 2002; Hughes et al., 2004; Kim et al., 2006; Miller et al., 2010; Drigo et al., 2014b). Even if the frequent update of diagnostic oligonucleotides is mandatory, this mission has often been proven to be challenging due to the costs associated with frequent test updates and validation and to the lack of data regarding the genetic variability of circulating field strains (Toplak et al., 2012). As a consequence, continuous devotion should be reserved to the monitoring of the molecular

epidemiologic scenario and to the sharing of related information. Moreover, the not-negligible cost of these assays could hinder adequate epidemiological investigations, particularly in poorly developed countries (Mabey et al., 2004). Remarkable efforts have been conducted and must be carried out to develop instrumentations and techniques aimed to reduce the economic burden and make them available for large-scale investigations in developing countries (Hotez et al., 2004; Mabey et al., 2004; Mori and Notomi, 2009; Franzo et al., 2014b).

Despite all of the efforts to limit pathogen introduction and spread, the wide interconnection among farms, at both the national and international levels, makes the coexistence with several pathogens substantially unavoidable. Vaccination represents the most commonly applied strategy to reduce clinical signs and economic losses. Unfortunately, the manufacturing of vaccines that are both efficacious and affordable for farming systems is often puzzling. A failure to vaccinate, as well as vaccinee-related (e.g., immunodeficiency, age-related maturation and senescence of immune responsiveness, interference due to other infectious agents, etc.) or vaccine-related causes (e.g., manufacturing-related, not 100% efficacy against included antigens, incomplete coverage of strains, serotypes, genotypes, etc.), can be adducted as causes of vaccination failure (Heininger et al., 2012). Even if it is undeniable that many RNA (Novella et al., 1995; Tannous et al., 2014) and ssDNA (Fort et al., 2008; Kristensen et al., 2011) virus infections can be effectively controlled by vaccination because of the presence of relatively conserved epitopes, the usual antigenic heterogeneity of several rapidly evolving viruses continues to frustrate every attempt to develop a globally valid control tool. In fact, some viruses have been demonstrated to produce escape mutants during infection, counteracting the host immune response and hampering vaccine efficacy (Goulder and Watkins, 2004). In most instances, however, the underling intra-host variability plays a less evident role (Murtaugh and Genzow, 2011), and “escape mutants” arise through genetic variations in time and space among host populations. Vaccination with seasonal influenza vaccines, *Porcine reproductive and respiratory syndrome virus* (PRRSV) and *Infectious bronchitis virus* (IBV) vaccines, just to report a few examples, has been proven to be more effective for inducing an adequate immune response and for protecting against overt clinical signs when challenged with homologous strains compared with heterologous ones (Gupta et al., 2006; Carrat and Flahault, 2007; Mateu and Diaz, 2008; Sjaak de Wit et al., 2011). In fact, the extreme diversity of these viruses continuously provides a plethora of variants against

which cross-protection is often poor. The frequent updating and development of homologous vaccines as well as validation of different vaccination strategies are strongly limited by economic constrains, particularly in veterinary medicine. The inadequate tools available, the scarce epidemiological data on circulating genotypes/serotypes/protectotypes and the paucity of experimental and field studies on vaccine efficacy against the currently circulating strains often lead to strategies based more on the personal experience of veterinarians than on factual data.

This background is particularly arduous considering that the implementation of vaccination protocols comes at costs that are not limited to economical costs. The use of live vaccines, which has often been proven to be markedly more effective than inactivated ones, can severely complicate the interpretation of the epidemiological scenario. In addition to altering the serological results, live vaccines are able to replicate and spread in a partially immune population for extended time periods (Cavanagh et al., 1999). A wide circulation of vaccine or vaccine-like strains has been documented for several animal diseases (Stadejek et al., 2013; Franzo et al., 2014c). In the absence of markers that allow the differentiation of field and vaccine strains, the interpretation of laboratory diagnosis results may be tentative (Lupini et al., 2011; Cecchinato et al., 2014; Franzo et al., 2014c; Listorti et al., 2014). Episodes of clinical disease associated to vaccine or vaccine-like strains are typically too cumbersome to resolve. Even though their association with other undetected pathogens by chance alone is typically the most likely explanation, residual virulence (Nielsen et al., 2001) or reversion to virulence are well-described phenomena (Lupini et al., 2011; Cecchinato et al., 2014). Modern farming and vaccination protocols, particularly for avian species, are at particular risk for reversion to virulence. The non-absolute coverage of animal population during ordinary mass vaccination in the poultry industry allows a prolonged circulation of vaccine viruses in animal populations (Cavanagh et al., 1999; Toro et al., 2012). Several back-passages in a partially immune population can progressively select those variants of the mutant spectrum that are more adapted to live animals. How and especially where these variants are generated and then selected remain a matter of discussion. Low-frequency virulent populations may “survive” through the attenuation process (Naylor and Jones, 1994) and then be selected following *in vivo* passages. An alternative hypothesis could assess the progressive accumulation and selection of mutations *in vivo*, starting from a homogenous vaccine population. Considering the RNA virus population structure, intermediate scenarios that propose an initial within-

vaccine variability followed by progressive adaption to the host environment are equally highly probable (McKinley et al., 2008; van Santen and Toro, 2008).

Last but not least, a wide use of vaccines has been implied in conditioning viral evolution, both for increasing virulence (Gandon et al., 2001; Gandon et al., 2003) and for leading to a shift in epitope profiles (Catelli et al., 2010; Cecchinato et al., 2010). Moreover, vaccination with live vaccines results in an increased risk of recombination events between vaccine and field strains or between different vaccine strains (Lee et al., 2012; Wenhui et al., 2012), a phenomenon with hard-to-predict consequences.

References

Alarcon, P., Rushton, J., Wieland, B., 2013. Cost of post-weaning multi-systemic wasting syndrome and porcine circovirus type-2 subclinical infection in England - an economic disease model. Prev. Vet. Med. 110, 88-102.

Alfonso-Morales, A., Martinez-Perez, O., Dolz, R., Valle, R., Perera, C.L., Bertran, K., Frias, M.T., Majo, N., Ganges, L., Perez, L.J., 2013. Spatiotemporal Phylogenetic Analysis and Molecular Characterisation of Infectious Bursal Disease Viruses Based on the VP2 Hyper-Variable Region. PLoS One 8, e65999.

Alizon, S., Fraser, C., 2013. Within-host and between-host evolutionary rates across the HIV-1 genome. Retrovirology 10, 49-4690-10-49.

Balka, G., Hornyák, A., Bálint, A., Benyeda, Z., Rusvai, M., 2009. Development of a one-step real-time quantitative PCR assay based on primer-probe energy transfer for the detection of porcine reproductive and respiratory syndrome virus. J. Virol. Methods 158, 41-45.

Bansal, S., Grenfell, B.T., Meyers, L.A., 2007. When individual behaviour matters: homogeneous and network models in epidemiology. J. R. Soc. Interface 4, 879-891.

Barzon, L., Lavezzo, E., Militello, V., Toppo, S., Palu, G., 2011. Applications of next-generation sequencing technologies to diagnostic virology. Int. J. Mol. Sci. 12, 7861-7884.

Baxa, D.M., Thekdi, A.D., Golembieski, A., Krishnan, P.V., Sharif, O., Kizy, A., Shetron-Rama, L., Jovanovich, J., Chappell, B.J., Snow-Lampart, A., Borroto-Esoda, K., Gordon, S.C., 2013.

Evaluation of anti-HBV drug resistant mutations among patients with acute symptomatic hepatitis B in the United States. J. Hepatol. 58, 212-216.

Beerenwinkel, N., Zagordi, O., 2011. Ultra-deep sequencing for the analysis of viral populations. Current opinion in virology 1, 413-418.

Beerenwinkel, N., Gunthard, H.F., Roth, V., Metzner, K.J., 2012. Challenges and opportunities in estimating viral genetic diversity from next-generation sequencing data. Front. Microbiol. 3, 329.

Belák, S., 2007. Molecular diagnosis of viral diseases, present trends and future aspects. A view from the OIE Collaborating Centre for the Application of Polymerase Chain Reaction Methods for Diagnosis of Viral Diseases in Veterinary Medicine. Vaccine 25, 5444-5452.

Belak, S., Karlsson, O.E., Blomstrom, A.L., Berg, M., Granberg, F., 2013. New viruses in veterinary medicine, detected by metagenomic approaches. Vet. Microbiol. 165, 95-101.

Belshaw, R., Gardner, A., Rambaut, A., Pybus, O.G., 2008. Pacing a small cage: mutation and RNA viruses. Trends in Ecology & Evolution 23, 188-193.

Belshaw, R., Pybus, O.G., Rambaut, A., 2007. The evolution of genome compression and genomic novelty in RNA viruses. Genome Res. 17, 1496-1504.

Bennett, R., 2003. The 'direct costs' of livestock disease: the development of a system of models for the analysis of 30 endemic livestock diseases in Great Britain. Journal of Agricultural Economics 54, 55-71.

Bertolotti, L., Kitron, U.D., Walker, E.D., Ruiz, M.O., Brawn, J.D., Loss, S.R., Hamer, G.L., Goldberg, T.L., 2008. Fine-scale genetic variation and evolution of West Nile Virus in a transmission "hot spot" in suburban Chicago, USA. Virology 374, 381-389.

Bimber, B.N., Dudley, D.M., Lauck, M., Becker, E.A., Chin, E.N., Lank, S.M., Grunenwald, H.L., Caruccio, N.C., Maffitt, M., Wilson, N.A., Reed, J.S., Sosman, J.M., Tarosso, L.F., Sanabani, S., Kallas, E.G., Hughes, A.L., O'Connor, D.H., 2010. Whole-genome characterization of

human and simian immunodeficiency virus intrahost diversity by ultradeep pyrosequencing. J. Virol. 84, 12087-12092.

Blomstrom, A.L., 2011. Viral metagenomics as an emerging and powerful tool in veterinary medicine. Vet. Q. 31, 107-114.

Borrow, P., Lewicki, H., Wei, X., Horwitz, M.S., Peffer, N., Meyers, H., Nelson, J.A., Gairin, J.E., Hahn, B.H., Oldstone, M.B., Shaw, G.M., 1997. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. Nat. Med. 3, 205-211.

Brandao, P.E., 2013. The evolution of codon usage in structural and non-structural viral genes: the case of Avian coronavirus and its natural host Gallus gallus. Virus Res. 178, 264-271.

Briddon, R.W., Bedford, I.D., Tsai, J.H., Markham, P.G., 1996. Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. Virology 219, 387-394.

Briese, T., Paweska, J.T., McMullan, L.K., Hutchison, S.K., Street, C., Palacios, G., Khristova, M.L., Weyer, J., Swanepoel, R., Egholm, M., Nichol, S.T., Lipkin, W.I., 2009. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. PLoS Pathog. 5, e1000455.

Brunborg, I.M., 2004. Quantitation of porcine circovirus type 2 isolated from serum/plasma and tissue samples of healthy pigs and pigs with postweaning multisystemic wasting syndrome using a TaqMan-based real-time PCR. J. Virol. Methods 122, 171-178.

Bull, R.A., Luciani, F., McElroy, K., Gaudieri, S., Pham, S.T., Chopra, A., Cameron, B., Maher, L., Dore, G.J., White, P.A., Lloyd, A.R., 2011. Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. PLoS Pathog. 7, e1002243.

Buzby, J.C., Roberts, T., 2009. The economics of enteric infections: human foodborne disease costs. Gastroenterology 136, 1851-1862.

Capobianchi, M., Giombini, E., Rozera, G., 2013. Next-generation sequencing technology in clinical virology. Clinical Microbiology and Infection 19, 15-22.

Carlson, J.M., Schaefer, M., Monaco, D.C., Batorsky, R., Claiborne, D.T., Prince, J., Deymier, M.J., Ende, Z.S., Klatt, N.R., DeZiel, C.E., Lin, T.H., Peng, J., Seese, A.M., Shapiro, R., Frater, J., Ndung'u, T., Tang, J., Goepfert, P., Gilmour, J., Price, M.A., Kilembe, W., Heckerman, D., Goulder, P.J., Allen, T.M., Allen, S., Hunter, E., 2014. HIV transmission. Selection bias at the heterosexual HIV-1 transmission bottleneck. Science 345, 1254031.

Carrat, F., Flahault, A., 2007. Influenza vaccine: the challenge of antigenic drift. Vaccine 25, 6852-6862.

Carrillo, C., Lu, Z., Borca, M.V., Vagnozzi, A., Kutish, G.F., Rock, D.L., 2007. Genetic and phenotypic variation of foot-and-mouth disease virus during serial passages in a natural host. J. Virol. 81, 11341-11351.

Catelli, E., Lupini, C., Cecchinato, M., Ricchizzi, E., Brown, P., Naylor, C.J., 2010. Field avian metapneumovirus evolution avoiding vaccine induced immunity. Vaccine 28, 916-921.

Cavanagh, D., Mawditt, K., Britton, P., Naylor, C., 1999. Longitudinal field studies of infectious bronchitis virus and avian pneumovirus in broilers using type-specific polymerase chain reactions. Avian Pathol. 28, 593-605.

Cecchinato, M., Catelli, E., Lupini, C., Ricchizzi, E., Clubbe, J., Battilani, M., Naylor, C.J., 2010. Avian metapneumovirus (AMPV) attachment protein involvement in probable virus evolution concurrent with mass live vaccine introduction. Vet. Microbiol. 146, 24-34.

Cecchinato, M., Catelli, E., Lupini, C., Ricchizzi, E., Prosperi, S., Naylor, C.J., 2014. Reversion to virulence of a subtype B avian metapneumovirus vaccine: Is it time for regulators to require availability of vaccine progenitors? Vaccine 32, 4660-4664.

Cecchinato, M., Lupini, C., Munoz Pogoreltseva, O.S., Listorti, V., Mondin, A., Drigo, M., Catelli, E., 2013. Development of a real-time RT-PCR assay for the simultaneous identification, quantitation and differentiation of avian metapneumovirus subtypes A and B. Avian Pathol. 42, 283-289.

Cella, L.N., Blackstock, D., Yates, M.A., Mulchandani, A., Chen, W., 2013. Detection of RNA viruses: current technologies and future perspectives. Crit. Rev. Eukaryot. Gene Expr. 23, 125-137.

Chao, L., 1990. Fitness of RNA virus decreased by Muller's ratchet. Nature 348, 454-455.

Chen, Y., Sun, J., Tong, X., Xu, J., Deng, H., Jiang, Z., Jiang, C., Duan, J., Li, J., Zhou, P., Wang, C., 2014. First analysis of synonymous codon usage in porcine circovirus. Arch. Virol. 159, 2145-2151.

Chen, Z., Terai, M., Fu, L., Herrero, R., DeSalle, R., Burk, R.D., 2005. Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. J. Virol. 79, 7014-7023.

Clarke, D.K., Duarte, E.A., Elena, S.F., Moya, A., Domingo, E., Holland, J., 1994. The red queen reigns in the kingdom of RNA viruses. Proc. Natl. Acad. Sci. U. S. A. 91, 4821-4824.

Cleaveland, S., Laurenson, M.K., Taylor, L.H., 2001. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 356, 991-999.

Commission on genetic resources for food and agriculture, 2007. THE STATE OF THE WORLD's ANIMAL GENETIC RESOURCES FOR FOOD AND AGRICULTURE. Rome.

Crandall, K.A., 1999. The evolution of HIV. JHU Press, .

Delgado, C., 2005. Rising demand for meat and milk in developing countries: implications for grasslands-based livestock production. Grassland: a global resource , 29-39.

Domingo, E., 2010. Mechanisms of viral emergence. Vet. Res. 41, 38.

Domingo, E., 2002. Quasispecies theory in virology. J. Virol. 76, 463-465.

Domingo, E., Escarmis, C., Sevilla, N., Moya, A., Elena, S.F., Quer, J., Novella, I.S., Holland, J.J., 1996. Basic concepts in RNA virus evolution. FASEB J. 10, 859-864.

Drew, T.W., 2011. The emergence and evolution of swine viral diseases: to what extent have husbandry systems and global trade contributed to their distribution and diversity? Rev. Sci. Tech. 30, 95-106.

Drigo, M., Franzo, G., Belfanti, I., Martini, M., Mondin, A., Ceglie, L., 2014a. Validation and comparison of different end point and real time RT-PCR assays for detection and genotyping of porcine reproductive and respiratory syndrome virus. J. Virol. Methods 201, 79-85.

Drigo, M., Franzo, G., Gigli, A., Martini, M., Mondin, A., Gracieux, P., Ceglie, L., 2014b. The impact of porcine reproductive and respiratory syndrome virus genetic heterogeneity on molecular assay performances. J. Virol. Methods 202, 79-86.

Drummond, A.J., Pybus, O.G., Rambaut, A., Forsberg, R., Rodrigo, A.G., 2003. Measurably evolving populations. Trends in Ecology & Evolution 18, 481-488.

Drummond, A.J., Ho, S.Y., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22, 1185-1192.

Duffy, S., Shackelton, L.A., Holmes, E.C., 2008. Rates of evolutionary change in viruses: patterns and determinants. Nature Reviews Genetics 9, 267-276.

Duffy, S., Holmes, E.C., 2008. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus. J. Virol. 82, 957-965.

Edwards, R.A., Rohwer, F., 2005. Viral metagenomics. Nature Reviews Microbiology 3, 504-510.

Elena, S.F., 2012. RNA virus genetic robustness: possible causes and some consequences. Curr. Opin. Virol. 2, 525-530.

Elena, S.F., Carrasco, P., Daros, J.A., Sanjuan, R., 2006. Mechanisms of genetic robustness in RNA viruses. EMBO Rep. 7, 168-173.

Elena, S.F., Miralles, R., Cuevas, J.M., Turner, P.E., Moya, A., 2000. The two faces of mutation: extinction and adaptation in RNA viruses. IUBMB Life 49, 5-9.

Elena, S.F., Sanjuan, R., 2005. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. J. Virol. 79, 11555-11558.

Escarmís, C., Dávila, M., Charpentier, N., Bracho, A., Moya, A., Domingo, E., 1996. Genetic lesions associated with Muller's ratchet in an RNA virus. J. Mol. Biol. 264, 255-267.

Faria, N.R., Suchard, M.A., Rambaut, A., Lemey, P., 2011. Toward a quantitative understanding of viral phylogeography. Curr. Opin. Virol. 1, 423-429.

Firth, C., Charleston, M.A., Duffy, S., Shapiro, B., Holmes, E.C., 2009. Insights into the evolutionary history of an emerging livestock pathogen: porcine circovirus 2. J. Virol. 83, 12813-12821.

Fonkwo, P.N., 2008. Pricing infectious disease. EMBO Rep. 9, S13-S17.

Forrester, N.L., Guerbois, M., Seymour, R.L., Spratt, H., Weaver, S.C., 2012. Vector-borne transmission imposes a severe bottleneck on an RNA virus population. PLoS Pathog. 8, e1002897.

Fort, M., Sibila, M., Allepuz, A., Mateu, E., Roerink, F., Segales, J., 2008. Porcine circovirus type 2 (PCV2) vaccination of conventional pigs prevents viremia against PCV2 isolates of different genotypes and geographic origins. Vaccine 26, 1063-1071.

Franzo, G., Cecchinato, M., Martini, M., Ceglie, L., Gigli, A., Drigo, M., 2014a. Observation of high recombination occurrence of Porcine Reproductive and Respiratory Syndrome Virus in field condition. Virus Res. 194, 159-166.

Franzo, G., Drigo, M., Lupini, C., Catelli, E., Laconi, A., Listorti, V., Bonci, M., Naylor, C.J., Martini, M., Cecchinato, M., 2014b. A sensitive, reproducible, and economic real-time

reverse transcription PCR detecting avian metapneumovirus subtypes A and B. Avian Dis. 58, 216-222.

Franzo, G., Naylor, C.J., Lupini, C., Drigo, M., Catelli, E., Listorti, V., Pesente, P., Giovanardi, D., Morandini, E., Cecchinato, M., 2014c. Continued use of IBV 793B vaccine needs reassessment after its withdrawal led to the genotype's disappearance. Vaccine 32, 6765-6767.

Frost, S.D., Volz, E.M., 2010. Viral phylodynamics and the search for an 'effective number of infections'. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365, 1879-1890.

Furio, V., Moya, A., Sanjuan, R., 2005. The cost of replication fidelity in an RNA virus. Proc. Natl. Acad. Sci. U. S. A. 102, 10233-10237.

Gallei, A., Pankraz, A., Thiel, H.J., Becher, P., 2004. RNA recombination in vivo in the absence of viral replication. J. Virol. 78, 6271-6281.

Gandon, S., Mackinnon, M., Nee, S., Read, A., 2003. Imperfect vaccination: some epidemiological and evolutionary consequences. Proc. Biol. Sci. 270, 1129-1136.

Gandon, S., Mackinnon, M.J., Nee, S., Read, A.F., 2001. Imperfect vaccines and the evolution of pathogen virulence. Nature 414, 751-756.

Gibbs, M.J., Weiller, G.F., 1999. Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. Proc. Natl. Acad. Sci. U. S. A. 96, 8022-8027.

Gill, M.S., Lemey, P., Faria, N.R., Rambaut, A., Shapiro, B., Suchard, M.A., 2013. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. Mol. Biol. Evol. 30, 713-724.

Glenn, T.C., 2011. Field guide to next-generation DNA sequencers. Mol. Ecol. Resour. 11, 759-769.

Gmyl, A.P., Belousov, E.V., Maslova, S.V., Khitrina, E.V., Chetverin, A.B., Agol, V.I., 1999. Nonreplicative RNA recombination in poliovirus. J. Virol. 73, 8958-8965.

Goulder, P.J., Watkins, D.I., 2004. HIV and SIV CTL escape: implications for vaccine design. Nat. Rev. Immunol. 4, 630-640.

Graham, J.P., Leibler, J.H., Price, L.B., Otte, J.M., Pfeiffer, D.U., Tiensin, T., Silbergeld, E.K., 2008. The animal-human interface and infectious disease in industrial food animal production: rethinking biosecurity and biocontainment. Public Health Rep. 123, 282-299.

Greenbaum, B.D., Li, O.T., Poon, L.L., Levine, A.J., Rabadan, R., 2012. Viral reassortment as an information exchange between viral segments. Proc. Natl. Acad. Sci. U. S. A. 109, 3341-3346.

Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood, J.L., Daly, J.M., Mumford, J.A., Holmes, E.C., 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. Science 303, 327-332.

Gupta, V., Earl, D.J., Deem, M.W., 2006. Quantifying influenza vaccine efficacy and antigenic distance. Vaccine 24, 3881-3888.

Heininger, U., Bachtiar, N.S., Bahri, P., Dana, A., Dodoo, A., Gidudu, J., Santos, E.M., 2012. The concept of vaccination failure. Vaccine 30, 1265-1268.

Heled, J., Drummond, A.J., 2008. Bayesian inference of population size history from multiple loci. BMC Evol. Biol. 8, 289-2148-8-289.

Henn, M.R., Boutwell, C.L., Charlebois, P., Lennon, N.J., Power, K.A., Macalalad, A.R., Berlin, A.M., Malboeuf, C.M., Ryan, E.M., Gnerre, S., Zody, M.C., Erlich, R.L., Green, L.M., Berical, A., Wang, Y., Casali, M., Streeck, H., Bloom, A.K., Dudek, T., Tully, D., Newman, R., Axten, K.L., Gladden, A.D., Battis, L., Kemper, M., Zeng, Q., Shea, T.P., Gujja, S., Zedlack, C., Gasser, O., Brander, C., Hess, C., Gunthard, H.F., Brumme, Z.L., Brumme, C.J., Bazner, S., Rychert, J., Tinsley, J.P., Mayer, K.H., Rosenberg, E., Pereyra, F., Levin, J.Z., Young, S.K., Jessen, H., Altfeld, M., Birren, B.W., Walker, B.D., Allen, T.M., 2012. Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. PLoS Pathog. 8, e1002529.

Hershberg, R., Petrov, D.A., 2008. Selection on codon bias. Annu. Rev. Genet. 42, 287-299.

Ho, S.Y., Shapiro, B., 2011. Skyline-plot methods for estimating demographic history from nucleotide sequences. Mol. Ecol. Resour. 11, 423-434.

Ho, S.Y., Shapiro, B., Phillips, M.J., Cooper, A., Drummond, A.J., 2007. Evidence for time dependency of molecular rate estimates. Syst. Biol. 56, 515-522.

Hoelzer, K., Murcia, P.R., Baillie, G.J., Wood, J.L., Metzger, S.M., Osterrieder, N., Dubovi, E.J., Holmes, E.C., Parrish, C.R., 2010. Intrahost evolutionary dynamics of canine influenza virus in naive and partially immune dogs. J. Virol. 84, 5329-5335.

Hoelzer, K., Shackelton, L.A., Parrish, C.R., Holmes, E.C., 2008. Phylogenetic analysis reveals the emergence, evolution and dispersal of carnivore parvoviruses. J. Gen. Virol. 89, 2280-2289.

Hofer, P., Bedford, I.D., Markham, P.G., Jeske, H., Frischmuth, T., 1997. Coat protein gene replacement results in whitefly transmission of an insect nontransmissible geminivirus isolate. Virology 236, 288-295.

Hoffmann, B., Beer, M., Reid, S.M., Mertens, P., Oura, C.A.L., van Rijn, P.A., Slomka, M.J., Banks, J., Brown, I.H., Alexander, D.J., King, D.P., 2009. A review of RT-PCR technologies used in veterinary virology and disease control: Sensitive and specific diagnosis of five livestock diseases notifiable to the World Organisation for Animal Health. Vet. Microbiol. 139, 1-23.

Holmes, E.C., 2009. The Evolution and Emergence of RNA Viruses. Oxford University Press, USA.

Holmes, E.C., Grenfell, B.T., 2009. Discovering the phylodynamics of RNA viruses. PLoS computational biology 5, e1000505.

Holmes, E.C., 2003a. Molecular clocks and the puzzle of RNA virus origins. J. Virol. 77, 3893-3897.

Holmes, E.C., 2003b. Patterns of intra- and interhost nonsynonymous variation reveal strong purifying selection in dengue virus. J. Virol. 77, 11296-11298.

Holmes, E.C., 2009. The Evolutionary Genetics of Emerging Viruses. Annu. Rev. Ecol. Evol. Syst. 40, 353-372.

Honegger, J.R., Kim, S., Price, A.A., Kohout, J.A., McKnight, K.L., Prasad, M.R., Lemon, S.M., Grakoui, A., Walker, C.M., 2013. Loss of immune escape mutations during persistent HCV infection in pregnancy enhances replication of vertically transmitted viruses. Nat. Med. 19, 1529-1533.

Hotez, P.J., Remme, J.H., Buss, P., Alleyne, G., Morel, C., Breman, J.G., 2004. Combating tropical infectious diseases: report of the Disease Control Priorities in Developing Countries Project. Clin. Infect. Dis. 38, 871-878.

Hughes, A.L., Hughes, M.A., 2007. More effective purifying selection on RNA viruses than in DNA viruses. Gene 404, 117-125.

Hughes, G.J., Smith, J.S., Hanlon, C.A., Rupprecht, C.E., 2004. Evaluation of a TaqMan PCR assay to detect rabies virus RNA: influence of sequence variation and application to quantification of viral loads. J. Clin. Microbiol. 42, 299-306.

Hughes, J., Allen, R.C., Baguelin, M., Hampson, K., Baillie, G.J., Elton, D., Newton, J.R., Kellam, P., Wood, J.L., Holmes, E.C., Murcia, P.R., 2012. Transmission of equine influenza virus during an outbreak is characterized by frequent mixed infections and loose transmission bottlenecks. PLoS Pathog. 8, e1003081.

Hurwitz, B.L., Westveld, A.H., Brum, J.R., Sullivan, M.B., 2014. Modeling ecological drivers in marine viral communities using comparative metagenomics and network analyses. Proc. Natl. Acad. Sci. U. S. A. 111, 10714-10719.

Jackwood, M.W., Boynton, T.O., Hilt, D.A., McKinley, E.T., Kissinger, J.C., Paterson, A.H., Robertson, J., Lemke, C., McCall, A.W., Williams, S.M., 2010. Emergence of a group 3 coronavirus through recombination. Virology 398, 98-108.

Jeske, H., Lutgemeier, M., Preiss, W., 2001. DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. EMBO J. 20, 6158-6167.

Jones, B.A., Grace, D., Kock, R., Alonso, S., Rushton, J., Said, M.Y., McKeever, D., Mutua, F., Young, J., McDermott, J., Pfeiffer, D.U., 2013. Zoonosis emergence linked to agricultural intensification and environmental change. Proc. Natl. Acad. Sci. U. S. A. 110, 8399-8404.

Katul, L., Timchenko, T., Gronenborn, B., Vetten, H.J., 1998. Ten distinct circular ssDNA components, four of which encode putative replication-associated proteins, are associated with the faba bean necrotic yellows virus genome. J. Gen. Virol. 79 (Pt 12), 3101-3109.

Keele, B.F., Giorgi, E.E., Salazar-Gonzalez, J.F., Decker, J.M., Pham, K.T., Salazar, M.G., Sun, C., Grayson, T., Wang, S., Li, H., Wei, X., Jiang, C., Kirchherr, J.L., Gao, F., Anderson, J.A., Ping, L.H., Swanstrom, R., Tomaras, G.D., Blattner, W.A., Goepfert, P.A., Kilby, J.M., Saag, M.S., Delwart, E.L., Busch, M.P., Cohen, M.S., Montefiori, D.C., Haynes, B.F., Gaschen, B., Athreya, G.S., Lee, H.Y., Wood, N., Seoighe, C., Perelson, A.S., Bhattacharya, T., Korber, B.T., Hahn, B.H., Shaw, G.M., 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc. Natl. Acad. Sci. U. S. A. 105, 7552-7557.

Keightley, P.D., Otto, S.P., 2006. Interference among deleterious mutations favours sex and recombination in finite populations. Nature 443, 89-92.

Kerr, P.J., Kitchen, A., Holmes, E.C., 2009. Origin and phylodynamics of rabbit hemorrhagic disease virus. J. Virol. 83, 12129-12138.

Kilpatrick, A.M., Daszak, P., Jones, M.J., Marra, P.P., Kramer, L.D., 2006. Host heterogeneity dominates West Nile virus transmission. Proc. Biol. Sci. 273, 2327-2333.

Kim, L.M., Afonso, C.L., Suarez, D.L., 2006. Effect of probe-site mismatches on detection of virulent Newcastle disease viruses using a fusion-gene real-time reverse transcription polymerase chain reaction test. J. Vet. Diagn. Invest. 18, 519-528.

Kim, T., Mudry, R.A.,Jr, Rexrode, C.A.,2nd, Pathak, V.K., 1996. Retroviral mutation rates and A-to-G hypermutations during different stages of retroviral replication. J. Virol. 70, 7594-7602.

Kingman, J.F.C., 1982. The coalescent. Stochastic processes and their applications 13, 235-248.

Koelle, K., Rasmussen, D.A., 2012. Rates of coalescence for common epidemiological models at equilibrium. J. R. Soc. Interface 9, 997-1007.

Kosakovsky Pond, S.L., Frost, S.D., 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol. Biol. Evol. 22, 1208-1222.

Kosakovsky Pond, S.L., Murrell, B., Fourment, M., Frost, S.D., Delport, W., Scheffler, K., 2011. A random effects branch-site model for detecting episodic diversifying selection. Mol. Biol. Evol. 28, 3033-3043.

Kosakovsky Pond, S.L., Poon, A.F., Leigh Brown, A.J., Frost, S.D., 2008. A maximum likelihood method for detecting directional evolution in protein sequences and its application to influenza A virus. Mol. Biol. Evol. 25, 1809-1824.

Kristensen, C.S., Baadsgaard, N.P., Toft, N., 2011. A meta-analysis comparing the effect of PCV2 vaccines on average daily weight gain and mortality rate in pigs from weaning to slaughter. Prev. Vet. Med. 98, 250-258.

Kuhnert, D., Stadler, T., Vaughan, T.G., Drummond, A.J., 2014. Simultaneous reconstruction of evolutionary history and epidemiological dynamics from viral sequences with the birth-death SIR model. J. R. Soc. Interface 11, 20131106.

Ladner, J.T., Beitzel, B., Chain, P.S., Davenport, M.G., Donaldson, E.F., Frieman, M., Kugelman, J.R., Kuhn, J.H., O'Rear, J., Sabeti, P.C., Wentworth, D.E., Wiley, M.R., Yu, G.Y., Threat Characterization Consortium, Sozhamannan, S., Bradburne, C., Palacios, G., 2014. Standards for sequencing viral genomes in the era of high-throughput sequencing. MBio 5, e01360-14.

Lai, M., 1996. Recombination in large RNA viruses: coronaviruses. 7, 381-388.

Lai, M.M., 1992. RNA recombination in animal and plant viruses. Microbiol. Rev. 56, 61-79.

Lalic, J., Cuevas, J.M., Elena, S.F., 2011. Effect of host species on the distribution of mutational fitness effects for an RNA virus. PLoS Genet. 7, e1002378.

Lauring, A.S., Frydman, J., Andino, R., 2013. The role of mutational robustness in RNA virus evolution. Nat. Rev. Microbiol. 11, 327-336.

Lázaro, E., 2011. RNA Viruses: Control, Mutagenesis and Extinction. eLS .

Lazaro, E., Escarmis, C., Perez-Mercader, J., Manrubia, S.C., Domingo, E., 2003. Resistance of virus to extinction on bottleneck passages: study of a decaying and fluctuating pattern of fitness loss. Proc. Natl. Acad. Sci. U. S. A. 100, 10830-10835.

Lee, S.W., Markham, P.F., Coppo, M.J., Legione, A.R., Markham, J.F., Noormohammadi, A.H., Browning, G.F., Ficorilli, N., Hartley, C.A., Devlin, J.M., 2012. Attenuated vaccines can recombine to form virulent field viruses. Science 337, 188.

Lefeuvre, P., Lett, J.M., Varsani, A., Martin, D.P., 2009. Widely conserved recombination patterns among single-stranded DNA viruses. J. Virol. 83, 2697-2707.

Leibler, J.H., Carone, M., Silbergeld, E.K., 2010. Contribution of company affiliation and social contacts to risk estimates of between-farm transmission of avian influenza. PLoS One 5, e9888.

Lemey, P., Pybus, O.G., Rambaut, A., Drummond, A.J., Robertson, D.L., Roques, P., Worobey, M., Vandamme, A.M., 2004. The molecular population genetics of HIV-1 group O. Genetics 167, 1059-1068.

Lemey, P., Rambaut, A., Drummond, A.J., Suchard, M.A., 2009. Bayesian phylogeography finds its roots. PLoS Comput. Biol. 5, e1000520.

Lemey, P., Rambaut, A., Welch, J.J., Suchard, M.A., 2010. Phylogeography takes a relaxed random walk in continuous space and time. Mol. Biol. Evol. 27, 1877-1885.

Lemey, P., Salemi, M., Vandamme, A., 2009. The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Cambridge University Press.

Leppik, L., Gunst, K., Lehtinen, M., Dillner, J., Streker, K., de Villiers, E.M., 2007. In vivo and in vitro intragenomic rearrangement of TT viruses. J. Virol. 81, 9346-9356.

Li, B., Fang, L., Xu, Z., Liu, S., Gao, J., Jiang, Y., Chen, H., Xiao, S., 2009. Recombination in vaccine and circulating strains of porcine reproductive and respiratory syndrome viruses. Emerg. Infect. Dis. 15, 2032-2035.

Li, C., Hatta, M., Watanabe, S., Neumann, G., Kawaoka, Y., 2008. Compatibility among polymerase subunit proteins is a restricting factor in reassortment between equine H7N7 and human H3N2 influenza viruses. J. Virol. 82, 11880-11888.

Listorti, V., Lupini, C., Cecchinato, M., Pesente, P., Rossi, G., Giovanardi, D., Naylor, C.J., Catelli, E., 2014. Rapid detection of subtype B avian metapneumoviruses using RT-PCR restriction endonuclease digestion indicates field circulation of vaccine-derived viruses in older turkeys. Avian Pathol. 43, 51-56.

Liu, Y.S., Zhou, J.H., Chen, H.T., Ma, L.N., Ding, Y.Z., Wang, M., Zhang, J., 2010. Analysis of synonymous codon usage in porcine reproductive and respiratory syndrome virus. Infect. Genet. Evol. 10, 797-803.

Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., Getz, W.M., 2005. Superspreading and the effect of individual variation on disease emergence. Nature 438, 355-359.

Lo Presti, A., Ciccozzi, M., Cella, E., Lai, A., Simonetti, F.R., Galli, M., Zehender, G., Rezza, G., 2012. Origin, evolution, and phylogeography of recent epidemic CHIKV strains. Infect. Genet. Evol. 12, 392-398.

Lupini, C., Cecchinato, M., Ricchizzi, E., Naylor, C.J., Catelli, E., 2011. A turkey rhinotracheitis outbreak caused by the environmental spread of a vaccine-derived avian metapneumovirus. Avian Pathol. 40, 525-530.

Mabey, D., Peeling, R.W., Ustianowski, A., Perkins, M.D., 2004. Tropical infectious diseases: diagnostics for the developing world. Nature Reviews Microbiology 2, 231-240.

Mackay, I.M., Arden, K.E., Nitsche, A., 2002. Real-time PCR in virology. Nucleic Acids Res. 30, 1292-1305.

Magiorkinis, G., Magiorkinis, E., Paraskevis, D., Ho, S.Y., Shapiro, B., Pybus, O.G., Allain, J.P., Hatzakis, A., 2009. The global spread of hepatitis C virus 1a and 1b: a phylodynamic and phylogeographic analysis. PLoS Med. 6, e1000198.

Magiorkinis, G., Paraskevis, D., Vandamme, A.M., Magiorkinis, E., Sypsa, V., Hatzakis, A., 2003. In vivo characteristics of human immunodeficiency virus type 1 intersubtype recombination: determination of hot spots and correlation with sequence similarity. J. Gen. Virol. 84, 2715-2722.

Malim, M.H., Emerman, M., 2001. HIV-1 sequence variation: drift, shift, and attenuation. Cell 104, 469-472.

Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., Trono, D., 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature 424, 99-103.

Manrubia, S.C., Escarmis, C., Domingo, E., Lazaro, E., 2005. High mutation rates, bottlenecks, and robustness of RNA viral quasispecies. Gene 347, 273-282.

Marshall, N., Priyamvada, L., Ende, Z., Steel, J., Lowen, A.C., 2013. Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. PLoS pathogens 9, e1003421.

Marston, D.A., McElhinney, L.M., Ellis, R.J., Horton, D.L., Wise, E.L., Leech, S.L., David, D., de Lamballerie, X., Fooks, A.R., 2013. Next generation sequencing of viral RNA genomes. BMC Genomics 14, 444-2164-14-444.

Martin, D.P., Biagini, P., Lefeuvre, P., Golden, M., Roumagnac, P., Varsani, A., 2011. Recombination in eukaryotic single stranded DNA viruses. Viruses 3, 1699-1738.

Martin, G., Otto, S.P., Lenormand, T., 2006. Selection for recombination in structured populations. Genetics 172, 593-609.

Mateu, E., Diaz, I., 2008. The challenge of PRRS immunology. Veterinary Journal 177, 345-351.

McInerney, J., Howe, K., Schepers, J., 1992. A framework for the economic analysis of disease in farm livestock. Prev. Vet. Med. 13, 137-154.

McKinley, E.T., Hilt, D.A., Jackwood, M.W., 2008. Avian coronavirus infectious bronchitis attenuated live vaccines undergo selection of subpopulations and mutations following vaccination. Vaccine 26, 1274-1284.

McLeod, R., 1995. Costs of major parasites to the Australian livestock industries. Int. J. Parasitol. 25, 1363-1367.

Miller, P.J., Decanini, E.L., Afonso, C.L., 2010. Newcastle disease: evolution of genotypes and the related diagnostic challenges. Infect. Genet. Evol. 10, 26-35.

Minin, V.N., Bloomquist, E.W., Suchard, M.A., 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Mol. Biol. Evol. 25, 1459-1471.

Minskaia, E., Hertzig, T., Gorbalenya, A.E., Campanacci, V., Cambillau, C., Canard, B., Ziebuhr, J., 2006. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. Proc. Natl. Acad. Sci. U. S. A. 103, 5108-5113.

Mokdad, A.H., Marks, J.S., Stroup, D.F., Gerberding, J.L., 2004. Actual causes of death in the United States, 2000. JAMA 291, 1238-1245.

Morens, D.M., Fauci, A.S., 2013. Emerging infectious diseases: threats to human health and global stability. PLoS Pathog. 9, e1003467.

Mori, Y., Notomi, T., 2009. Loop-mediated isothermal amplification (LAMP): a rapid, accurate, and cost-effective diagnostic method for infectious diseases. Journal of infection and chemotherapy 15, 62-69.

Morris, R., Marsh, W., 1992. The relationship between infections, diseases and their economic effects .

Muller, H.J., 1964. The relation of recombination to mutational advance. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 1, 2-9.

Muller, H.J., 1932. Some genetic aspects of sex. Am. Nat., 118-138.

Murcia, P.R., Baillie, G.J., Daly, J., Elton, D., Jervis, C., Mumford, J.A., Newton, R., Parrish, C.R., Hoelzer, K., Dougan, G., Parkhill, J., Lennard, N., Ormond, D., Moule, S., Whitwham, A., McCauley, J.W., McKinley, T.J., Holmes, E.C., Grenfell, B.T., Wood, J.L., 2010. Intra- and interhost evolutionary dynamics of equine influenza virus. J. Virol. 84, 6943-6954.

Murcia, P.R., Wood, J.L., Holmes, E.C., 2011. Genome-scale evolution and phylodynamics of equine H3N8 influenza A virus. J. Virol. 85, 5312-5322.

Murrell, B., de Oliveira, T., Seebregts, C., Kosakovsky Pond, S.L., Scheffler, K., Southern African Treatment and Resistance Network-SATuRN Consortium, 2012. Modeling HIV-1 drug resistance as episodic directional selection. PLoS Comput. Biol. 8, e1002507.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S.L., Scheffler, K., 2013. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol. Biol. Evol. 30, 1196-1205.

Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., Kosakovsky Pond, S.L., 2012. Detecting individual sites subject to episodic diversifying selection. PLoS Genet. 8, e1002764.

Murtaugh, M.P., Genzow, M., 2011. Immunological solutions for treatment and prevention of porcine reproductive and respiratory syndrome (PRRS). Vaccine 29, 8192-8204.

Naylor, C.J., Jones, R.C., 1994. Demonstration of a virulent subpopulation in a prototype live attenuated turkey rhinotracheitis vaccine. Vaccine 12, 1225-1230.

Neumann, E.J., Kliebenstein, J.B., Johnson, C.D., Mabry, J.W., Bush, E.J., Seitzinger, A.H., Green, A.L., Zimmerman, J.J., 2005. Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States. J. Am. Vet. Med. Assoc. 227, 385-392.

Ng, T.F., Willner, D.L., Lim, Y.W., Schmieder, R., Chau, B., Nilsson, C., Anthony, S., Ruan, Y., Rohwer, F., Breitbart, M., 2011. Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. PLoS One 6, e20579.

Nielsen, H.S., Oleksiewicz, M.B., Forsberg, R., Stadejek, T., Botner, A., Storgaard, T., 2001. Reversion of a live porcine reproductive and respiratory syndrome virus vaccine investigated by parallel mutations. J. Gen. Virol. 82, 1263-1272.

Nora, T., Charpentier, C., Tenaillon, O., Hoede, C., Clavel, F., Hance, A.J., 2007. Contribution of recombination to the evolution of human immunodeficiency viruses expressing resistance to antiretroviral treatment. J. Virol. 81, 7620-7628.

Norstrom, M.M., Karlsson, A.C., Salemi, M., 2012. Towards a new paradigm linking virus molecular evolution and pathogenesis: experimental design and phylodynamic inference. New Microbiol. 35, 101-111.

Novella, I.S., Domingo, E., Holland, J.J., 1995. Rapid viral quasispecies evolution: implications for vaccine and drug strategies. Mol. Med. Today 1, 248-253.

Novella, I.S., Zarate, S., Metzgar, D., Ebendick-Corpus, B.E., 2004. Positive selection of synonymous mutations in vesicular stomatitis virus. J. Mol. Biol. 342, 1415-1421.

Pappas, G., 2013. Socio-economic, industrial and cultural parameters of pig-borne infections. Clin. Microbiol. Infect. 19, 605-610.

Pasternak, A.O., Spaan, W.J., Snijder, E.J., 2006. Nidovirus transcription: how to make sense...? J. Gen. Virol. 87, 1403-1421.

Pasternak, A.O., van den Born, E., Spaan, W.J., Snijder, E.J., 2001. Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. EMBO J. 20, 7220-7228.

Perry, B.D., Grace, D., Sones, K., 2013. Current drivers and future directions of global livestock disease dynamics. Proc. Natl. Acad. Sci. U. S. A. 110, 20871-20877.

Pilartz, M., Jeske, H., 2003. Mapping of abutilon mosaic geminivirus minichromosomes. J. Virol. 77, 10808-10818.

Pritchett, J., Thilmany, D., Johnson, K., 2005. Animal disease economic impacts: A survey of literature and typology of research approaches. International Food and Agribusiness management review 8, 23-45.

Prosperi, M.C., Yin, L., Nolan, D.J., Lowe, A.D., Goodenow, M.M., Salemi, M., 2013. Empirical validation of viral quasispecies assembly algorithms: state-of-the-art and challenges. Sci. Rep. 3, 2837.

Pybus, O.G., Rambaut, A., 2009. Evolutionary analysis of the dynamics of viral infectious disease. Nature Reviews Genetics 10, 540-550.

Pybus, O.G., Drummond, A.J., Nakano, T., Robertson, B.H., Rambaut, A., 2003. The epidemiology and iatrogenic transmission of hepatitis C virus in Egypt: a Bayesian coalescent approach. Mol. Biol. Evol. 20, 381-387.

Pybus, O.G., Rambaut, A., Belshaw, R., Freckleton, R.P., Drummond, A.J., Holmes, E.C., 2007. Phylogenetic evidence for deleterious mutation load in RNA viruses and its contribution to viral evolution. Mol. Biol. Evol. 24, 845-852.

Qin, Y., Petty, I.T., 2001. Genetic analysis of bipartite geminivirus tissue tropism. Virology 291, 311-323.

Rabadan, R., Levine, A.J., Robins, H., 2006. Comparison of avian and human influenza A viruses reveals a mutational bias on the viral genomes. J. Virol. 80, 11887-11891.

Radford, A.D., Chapman, D., Dixon, L., Chantrey, J., Darby, A.C., Hall, N., 2012. Application of next-generation sequencing technologies in virology. J. Gen. Virol. 93, 1853-1868.

Rambaut, A., Pybus, O.G., Nelson, M.I., Viboud, C., Taubenberger, J.K., Holmes, E.C., 2008. The genomic and epidemiological dynamics of human influenza A virus. Nature 453, 615-619.

Robinson, S.R., Abrahante, J.E., Johnson, C.R., Murtaugh, M.P., 2013. Purifying selection in porcine reproductive and respiratory syndrome virus ORF5a protein influences variation in envelope glycoprotein 5 glycosylation. Infect. Genet. Evol. 20, 362-368.

Rodrigo, A.G., Felsenstein, J., 1999. Coalescent approaches to HIV population genetics. The evolution of HIV , 233-272.

Ross, H.A., Rodrigo, A.G., 2002. Immune-mediated positive selection drives human immunodeficiency virus type 1 molecular variation and predicts disease duration. J. Virol. 76, 11715-11720.

Rowe, C.L., Fleming, J.O., Nathan, M.J., Sgro, J.Y., Palmenberg, A.C., Baker, S.C., 1997. Generation of coronavirus spike deletion variants by high-frequency recombination at regions of predicted RNA secondary structure. J. Virol. 71, 6183-6190.

Sachsenroder, J., Twardziok, S., Hammerl, J.A., Janczyk, P., Wrede, P., Hertwig, S., Johne, R., 2012. Simultaneous identification of DNA and RNA viruses present in pig faeces using process-controlled deep sequencing. PLoS One 7, e34631.

Sanjuan, R., Moya, A., Elena, S.F., 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc. Natl. Acad. Sci. U. S. A. 101, 8396-8401.

Sanjuan, R., Nebot, M.R., Chirico, N., Mansky, L.M., Belshaw, R., 2010. Viral mutation rates. J. Virol. 84, 9733-9748.

Savolainen-Kopra, C., Blomqvist, S., 2010. Mechanisms of genetic variation in polioviruses. Rev. Med. Virol. 20, 358-371.

Segales, J., Kekarainen, T., Cortey, M., 2013. The natural history of porcine circovirus type 2: from an inoffensive virus to a devastating swine disease? Vet. Microbiol. 165, 13-20.

Seoighe, C., Ketwaroo, F., Pillay, V., Scheffler, K., Wood, N., Duffet, R., Zvelebil, M., Martinson, N., McIntyre, J., Morris, L., Hide, W., 2007. A model of directional selection applied to the evolution of drug resistance in HIV-1. Mol. Biol. Evol. 24, 1025-1031.

Shen, W., Gao, L., Balakrishnan, M., Bambara, R.A., 2009. A recombination hot spot in HIV-1 contains guanosine runs that can form a G-quartet structure and promote strand transfer in vitro. J. Biol. Chem. 284, 33883-33893.

Shi, M., Lam, T.T., Hon, C., Murtaugh, M.P., Davies, P.R., Hui, R.K., Li, J., Wong, L.T., Yip, C., Jiang, J., Leung, F.C., 2010. Phylogeny-based evolutionary, demographical, and geographical dissection of north american type 2 porcine reproductive and respiratory syndrome viruses. J. Virol. 84, 8700-8711.

Simon-Loriere, E., Holmes, E.C., 2011. Why do RNA viruses recombine? Nat. Rev. Microbiol. 9, 617-626.

Simon-Loriere, E., Martin, D.P., Weeks, K.M., Negroni, M., 2010. RNA structures facilitate recombination-mediated gene swapping in HIV-1. J. Virol. 84, 12675-12682.

Sjaak de Wit, J.J., Cook, J.K., van der Heijden, H.M., 2011. Infectious bronchitis virus variants: a review of the history, current situation and control measures. Avian Pathol. 40, 223-235.

Stack, J.C., Murcia, P.R., Grenfell, B.T., Wood, J.L., Holmes, E.C., 2013. Inferring the inter-host transmission of influenza A virus using patterns of intra-host genetic variation. Proc. Biol. Sci. 280, 20122173.

Stadejek, T., Stankevicius, A., Murtaugh, M.P., Oleksiewicz, M.B., 2013. Molecular evolution of PRRSV in Europe: current state of play. Vet. Microbiol. 165, 21-28.

Stadler, T., Kuhnert, D., Bonhoeffer, S., Drummond, A.J., 2013. Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). Proc. Natl. Acad. Sci. U. S. A. 110, 228-233.

Stilianakis, N.I., Schenzle, D., 2006. On the intra-host dynamics of HIV-1 infections. Math. Biosci. 199, 1-25.

Svarovskaia, E.S., Cheslock, S.R., Zhang, W.H., Hu, W.S., Pathak, V.K., 2003. Retroviral mutation rates and reverse transcriptase fidelity. Front. Biosci. 8, d117-34.

Swayne, D.E., Glisson, J.R., McDougald, L.R., Nolan, L.K., Suarez, D.L., Nair, V.L., 2013. Diseases of Poultry, 13th Edition. Wiley-Blackwell, .

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30, 2725-2729.

Tan le, V., van Doorn, H.R., Nghia, H.D., Chau, T.T., Tu le, T.P., de Vries, M., Canuti, M., Deijs, M., Jebbink, M.F., Baker, S., Bryant, J.E., Tham, N.T., BKrong, N.T., Boni, M.F., Loi, T.Q., Phuong le, T., Verhoeven, J.T., Crusat, M., Jeeninga, R.E., Schultsz, C., Chau, N.V., Hien, T.T., van der Hoek, L., Farrar, J., de Jong, M.D., 2013. Identification of a new cyclovirus in cerebrospinal fluid of patients with acute central nervous system infections. MBio 4, e00231-13.

Tannous, L.K., Barlow, G., Metcalfe, N.H., 2014. A short clinical review of vaccination against measles. JRSM Open 5, 2054270414523408.

Teoh, B., Sam, S., Tan, K., Danlami, M., Johari, J., Shu, M., Pang, S., Nor-Amdan, N., Hashim, S., Tiong, V., 2012. Purifying selection in the evolution of dengue virus type 1 in Malaysia. International Journal of Infectious Diseases 16, e273.

Thornton, P.K., 2010. Livestock production: recent trends, future prospects. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365, 2853-2867.

Toplak, I., Rihtarič, D., Hostnik, P., Grom, J., Štukelj, M., Valenčak, Z., 2012. Identification of a genetically diverse sequence of porcine reproductive and respiratory syndrome virus in Slovenia and the impact on the sensitivity of four molecular tests. J. Virol. Methods 179, 51-56.

Toro, H., Pennington, D., Gallardo, R.A., van Santen, V.L., van Ginkel, F.W., Zhang, J., Joiner, K.S., 2012. Infectious bronchitis virus subpopulations in vaccinated chickens after challenge. Avian Dis. 56, 501-508.

Ubol, S., Suksatu, A., Modhiran, N., Sangma, C., Thitithanyanont, A., Fukuda, M., Juthayothin, T., 2011. Intra-host diversities of the receptor-binding domain of stork faeces-derived avian H5N1 viruses and its significance as predicted by molecular dynamic simulation. J. Gen. Virol. 92, 307-314.

van Santen, V.L., Toro, H., 2008. Rapid selection in chickens of subpopulations within ArkDPI-derived infectious bronchitis virus vaccines. Avian Pathol. 37, 293-306.

van Vugt, J., Storgaard, T., Oleksiewicz, M.B., Bøtner, A., 2001. High frequency RNA recombination in porcine reproductive and respiratory syndrome virus occurs preferentially between parental sequences with high similarity. J. Gen. Virol. 82, 2615-2620.

Volz, E.M., Koelle, K., Bedford, T., 2013. Viral phylodynamics. PLoS Comput. Biol. 9, e1002947.

Walewski, J.L., Keller, T.R., Stump, D.D., Branch, A.D., 2001. Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. RNA 7, 710-721.

Wargo, A.R., Kurath, G., 2012. Viral fitness: definitions, measurement, and current insights. Current opinion in virology 2, 538-545.

Wenhui, L., Zhongyan, W., Guanqun, Z., Zhili, L., JingYun, M., Qingmei, X., Baoli, S., Yingzuo, B., 2012. Complete genome sequence of a novel variant porcine reproductive and respiratory syndrome virus (PRRSV) strain: evidence for recombination between vaccine and wild-type PRRSV strains. J. Virol. 86, 9543-12.

Woolhouse, M.E., Gowtage-Sequeria, S., 2005. Host range and emerging and reemerging pathogens. Emerg. Infect. Dis. 11, 1842-1847.

Yang, X., Chockalingam, S.P., Aluru, S., 2013. A survey of error-correction methods for next-generation sequencing. Brief Bioinform 14, 56-66.

Yang, Z., Bruno, D.P., Martens, C.A., Porcella, S.F., Moss, B., 2010. Simultaneous high-resolution analysis of vaccinia virus and host cell transcriptomes by deep RNA sequencing. Proc. Natl. Acad. Sci. U. S. A. 107, 11513-11518.

Yang, Z., Reynolds, S.E., Martens, C.A., Bruno, D.P., Porcella, S.F., Moss, B., 2011. Expression profiling of the intermediate and late stages of poxvirus replication. J. Virol. 85, 9899-9908.

Yuste, E., Sanchez-Palomino, S., Casado, C., Domingo, E., Lopez-Galindez, C., 1999. Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. J. Virol. 73, 2745-2751.

Zimmerman, J.J., Karriker, L.A., Ramirez, A., Schwartz, K.J., Stevenson, G.W., 2012. Diseases of Swine, 10th Edition.

AIM OF THE THESIS

As described in the introductory chapter, rapidly evolving RNA and ssDNA viruses represent a fascinating field characterized by a strict interconnection between the "speculative" study of viral evolution and its practical implications in everyday veterinary medicine.

This thesis has been thought as a collection of manuscripts that aimed to investigate different aspects and levels of viral evolution while still maintaining a focus on practical repercussions. Even if different infectious diseases and etiological agents are considered, they are all functional to the study of different aspects and implications of rapid virus evolution.

Considering the heterogeneous nature of the manuscripts, they have been organized according to a "crescent" scale, starting from the lowest scale of viral evolution and progressing to broader scales.

The manuscript "**Viral subpopulations in aMPV vaccines are unlikely to be responsible for reversion to virulence**." addresses a fine-level analysis of the population structure of the AMPV subtype B live attenuated vaccine and its potential role in the previously demonstrated phenomenon of reversion to virulence.

The widespread administration of live attenuated vaccines, despite their obvious advantages in terms of reducing disease prevalence, clinical signs and economic losses, is associated with costs that are not limited to the risk of reversion to virulence or to direct economic costs. Based on a wide collection of Italian samples, "**Continued use of IBV 793B vaccine needs reassessment after its withdrawal led to the genotype's disappearance**" reports the impact of these vaccines in complicating the diagnostic process and, as a consequence, the interpretation of the epidemiological scenario in the absence of known vaccine markers.

Obviously, updated knowledge of the strains currently circulating in a particular area is of great relevance for the implementation of proper control strategies. With this aim in mind, a field survey, which is published in "**Field survey of Avian Metapneumovirus in Northern Italy**", was conducted on hundreds of Italian farms to estimate and characterize the AMPV strains circulating in our country. To further support frequent and extensive surveys, an assay that is able to detect, quantify and genotype the two AMPV subtypes currently circulating in Italy was developed and validated. Because economic constraints

often represent a major limit, efforts were made to reduce the assay costs compared with other real-time RT-PCR methods while still guaranteeing comparable or superior performances ("**A Sensitive, Reproducible, and Economic Real-Time Reverse Transcription PCR Detecting Avian Metapneumovirus Subtypes A and B**").

Unfortunately, the diagnosis of rapidly evolving RNA viruses is itself an arduous task that requires a continuous evaluation and updating of diagnostic tools, even in laboratory that receive samples from limited geographic areas. The manuscripts entitled "**Observation of high recombination occurrence of Porcine reproductive and respiratory syndrome virus in field condition**" and "**Phylodynamic analysis of Porcine reproductive and respiratory syndrome virus (PRRSV) in Italy: action of selective pressures and interactions between different clades.**" address the study of the molecular epidemiology of *Porcine reproductive and respiratory syndrome virus* (PRRSV) in Italy considering the evolutionary forces driving PRRSV evolution at the local scale (i.e., high substitution rate, recombination, interaction between different clades and action of selective pressures). The high heterogeneity of PRRSV in the national contest was then evaluated with respect to the challenges that it poses in the development and validation of RT-PCR- and qRT-PCR-based diagnostic methods ("**Validation and comparison of different end point and real time RT-PCR assays for detection and genotyping of porcine reproductive and respiratory syndrome virus**") and assessing its impact on diagnostic accuracy ("**The impact of porcine reproductive and respiratory syndrome virus genetic heterogeneity on molecular assay performances**"). Similarly, "**International trades, local spread and viral evolution: the case of Porcine circovirus type 2 (PCV2) strains heterogeneity in Italy**" investigates the genetic variability of PCV2 within national borders and compares it with the knowledge of its molecular epidemiology available from other countries. This study provides evidence regarding the role of both "*in loco*" evolution and importation of different genotypes and strains from foreign countries in determining the Italian PCV2 genetic heterogeneity. The crescent amount of PCV2 sequences deposited in publically available databases has revealed its marked variability and challenged the current classification criteria. Nevertheless, at least a superficial knowledge of the PCV2 molecular epidemiology is mandatory for the planning and evaluation of control strategies. "**Revisiting the Taxonomical classification of PCV2: still a real challenge**" proposes new criteria for the classification of PCV2 into different genotypes. Our aim was to provide a scheme that both accounts for the constraint imposed by the biological proprieties of this

virus and allows a rapid, practical and easy way to classify PCV2 strains even during routine diagnostic activity. Last, "**Genetic characterisation of porcine circovirus type 2 (PCV2) strains from feral pigs in the Brazilian Pantanal: an opportunity to reconstruct the history of PCV2 evolution**" investigates more speculative issues inherent to the PCV2 origin. The discovery of a PCV2c genotype, which, to date, was believed to be extinct, in a feral pig population characterized by a peculiar population history and by a complex, and still partially known, relationship network with other PCV2-susceptible species opens exciting scenarios concerning the history and origin of PCV2.

Subpopulations in aMPV vaccines are unlikely to be the major cause of reversion to virulence.

Submitted to Vaccine

Franzo G[1]., Naylor C.J.[2], Drigo M.[1], Croville G.[3], Ducatez M.F.[3], Catelli E.[4], Cecchinato M.[1]

[1] Department of Animal Medicine, Production and Health, University of Padua, Viale dell'università, 16, 35020, Legnaro (PD), Italy;

[2] Department of Infection Biology, Institute of Infection and Global Health, Faculty of Health and Life Sciences, University of Liverpool, Leahurst Campus, Neston, Cheshire, CH64 7TE – United Kingdom.

[3]INRA and Université de Toulouse, INP, ENVT, UMR 1225 IHAP, F-31076, Toulouse, France

[4]Department of Veterinary Medical Sciences, *Alma Mater Studiorum* – University of Bologna,

Via Tolara di Sopra, 50, 40064 Ozzano Emilia (BO), Italy;

**Corresponding author. Tel: +390498272968; fax: +390498272973;*

email address: giovanni.franzo@studenti.unipd.it / giovanni.franzo1@gmail.com

Abstract

Avian metapneumovirus infects respiratory and reproductive tracts of domestic poultry, often involving secondary infection, and leads to serious economic losses in most parts of the world. While in general disease is effectively controlled by live vaccines, reversion to virulence of those vaccines has been demonstrated on several occasions. Consensus sequence mutations involved in the process have been identified in more than one instance. In one previous subtype A AMPV candidate vaccine study, small subpopulations were implicated. In the current study, the presence of subpopulations in a subtype B

vaccine was investigated by deep sequencing. Of the 19 positions where previously determined vaccine and progenitor consensus sequences differed, subpopulations were found to have sequence matching progenitor sequence in 4 positions. However none of these mutations occurred in a virulent revertant of that vaccine, thereby demonstrating that the majority progenitor virus population had not survived the attenuation process, hence were not obviously involved in any return to virulence. However within the vaccine, a single nucleotide variation was found which agreed with consensus sequence of a derived virulent revertant virus, hence this and other undetected, potentially virulent subpopulations, cannot totally discounted from being involved in reversion. Much deeper sequencing of progenitor, vaccine and revertant may clarify whether problematic virulent subpopulations are present and therefore whether these need to be routinely removed during AMPV vaccine preparation prior to registration and release.

Keywords

aMPV; Vaccine; Reversion; Subpopulation, NGS.

Introduction

Avian metapneumovirus (AMPV) is an important pathogen of poultry which is responsible for major economic losses in throughout most of the world due to both respiratory and reproductive syndromes [1,2]. While generally controlled by wide administration of live-attenuated vaccines, reversion to virulence has been previously demonstrated and consensus associated mutations have been identified for both subtype A and B AMPVs [3,4]. While subpopulations, detected by relative cell tropisms, were previously found involved in reversion of a subtype A AMPV vaccine candidate [5],this did not preclude evolution of the major population as a concurrent cause. In the current study, a subtype B vaccine was deep sequenced using next generation sequencing (NGS) in an attempt to identify virulent subpopulations potentially playing a role in reversion to virulence.

Material and methods

A vial of live subtype B AMPV vaccine (strain VCO3/50) (lot number 81312080) was resuspended in 2mL of sterile nanopure water. DNA was digested and RNA was extracted

from 200µL of virus suspension using a High Pure Viral RNA Kit (Roche Diagnostics, Monza, Italy), according to the manufacturer's instructions.

NGS sequencing

Two library preparation approaches were developed to evaluate the effectiveness of different strategies and to reduce the bias due to sample processing.

Protocol 1: A two-step RT-PCR was developed to amplify the full genome of aMPV as two overlapping amplicons. Primers were designed based on already published sequence (Cecchinato et al., 2014). Reverse transcription was performed (Maxima H Minus Reverse kit, Thermo Scientific) and PCR was performed (Phusion Hot Start II High-Fidelity DNA Polymerase kit ,Thermo Scientific) following manufacturer instructions. Briefly, primer aMPVB FullF (5'-AAAACGCATACAAGTCACAATAGAA-3') was used for reverse transcription while primer pairs aMPVB_FullF (5'-AAAACGCATACAAGTCACAATAGAA-3')-aMPVB_InternalR (5'-TATGTTGCAGGTTCCCGATT-3') and aMPVB_InternalF (5'-AAACTTCTGCCGATGCACTT-3')-aMPVB_FullR(5'-AACTAATGCACCCCTTTTTCC-3') were selected for PCR. PCR products were purified (NucleoSpin® Gel and PCR Clean-up kit, MACHEREY-NAGEL, Italy) according to manufacturer instructions. Purified PCR products were quantified using NanoDrop (Thermo Scientific), diluted to 25ng/µL and sent to MACROGEN for sequencing. The library was constructed using the Nextera DNA XT and products were sequenced on an Illumina HiSeq using a paired-end 100 bp run.

*Protocol 2:*Viral RNA was retro-transcribed with addition of conserved sequences [6] at 5' extremities of cDNA. For reverse transcription, 7.5 µL of viral RNA were mixed with 0.1 nmol of tagged random hexamer primers and incubated for 5 min at 65°C. Then, 4 µL of 5X reaction buffer (ThermoFisher Scientific, Carlsbad, CA) were added to this mixture at the same time as 0.5 µL of RNase Out (Lifetechnologies, Carlsbad, CA), 20 nmol of deoxynucleoside triphosphate (dNTP) solution (Finnzymes, Espoo, Finland) and 1 µL of RevertAid reverse transcriptase (ThermoFisher Scientific). The reaction volume was completed to 20 µL with distilled water. The RT reaction was performed at 42°C for 60 min and terminated by heating at 70°C for 10 min.

The second strand cDNA was synthesized by mixing 10 µL of cDNA with 50 pmol of the same tagged random hexamer primers as previously used, 3 µL of 10X React 2 Buffer (Lifetechnologies), 0.5 nmol of dNTP solution, 0.5 U of Large Fragment of DNA Polymerase

I (Lifetechnologies) and water to make a total volume of 30 µL. The incubation step was performed at room temperature for 60 min.

The PCR mixture was prepared and contained 33 µL of distilled water, 10 µL of 5X buffer (ThermoFisher Scientific), 10 nmol of dNTP solution (Finnzymes), 10 pmol of primer (corresponding only to the tag previously used), 5 µL of the double stranded cDNA and 2 units Phusion High-Fidelity DNA polymerase (ThermoFisher Scientific). The amplification program consisted of a first 30-s step at 98°C followed by 40 cycles with the following conditions: 98°C for 10 s, 65°C for 20 s, and 72°C for 30 s. The program ended with 1 step at 72°C for 10 min.

The amplification products were separated by electrophoresis using 1% agarose gel. Observed bands around 500 bp were excised from the gel and purified using the Nucleospin Gel and PCR clean-up kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's protocol. These PCR products were eluted into 50 µL of distilled water and subsequently used for the MiSeq sequencing. Sequencing has been performed on a MiSeq (Illumina, San Diego, CA) sequencer using a paired-end read length of 2x300 pb with the Illumina MiSeq reagent Kits v3 (Illumina).

Data analysis

Both data sets obtained from the two sequencing protocols were subjected to two analysis approaches. Settings were adjusted to account for varying reads length and sequencing strategies. The overlapping reads obtained through protocol 2 were assembled into contigs and treated as single reads.

Data quality was preliminary inspected using FastQC (http://www.bioinformatics.babraham.ac.uk /projects/fastqc/). Adaptor contamination and poor quality bases were trimmed using Trimmomatic [7]. Sequences shorter than 40bp were discarded. Reads were aligned to the reference genome using Bowtie2 [8]. Presence of single nucleotide variations was investigated using two open-source software implementing different approaches: ShoRAH [9]and QUASR [10]. The former corrected sequencing errors by clustering all reads that overlap the same region of the genome of length approximately equal to the read length, adjusted according to author recommendations. The consensus sequence of each cluster was assumed to be the original haplotype from which the erroneous reads were obtained. The number of reads associated

with the cluster estimates the prevalence of the haplotype in the population. The latter software used an approach based on base quality and variant coverage to asses the accuracy of the variant. A conservative consensus approach based on the combination of different sample preparation and data analysis protocols was applied to investigate the presence of viral subpopulations. Results of single methods were compared to known sites involved in vaccine attenuation process and reversion to virulence [4].

Complete genome sequences (70% consensus) were also obtained using QUASR and compared with the reference genome obtained using Sanger sequencing.

Results

Library preparation and NGS run

Protocol 1: RT-PCR products of 7108 and 6988bp were made. Approximately fifteen million reads were obtained. Ninety nine point five percent of paired-end sequences, were correctly aligned to the reference genome leading to a mean coverage of approximately 100,000.

Protocol 2. A total of 220,000 reads were obtained. Thirty one percent of single-end sequences were correctly aligned to the reference genome leading to a mean coverage of 3000.

Data analysis

Average coverage after duplicates removal was 200 and 272 for protocols 1 and 2, respectively. Both consensus sequences confirmed the previously obtained consensus genome sequence with one exception. A coverage higher than 10 was obtained from base 14 to 13427using protocol one and from 6 to 13433 using two. Comparison of the fully overlapping regions revealed an optimal agreement between NGS and Sanger obtained data. NGS revealed just one discrepancy with the previously published consensus sequence whereby NGS gave T at position 3566 which was previously reported as C.

Single nucleotide variations (SNV) analysis

Most SNVs reported by the different methods were not in agreement. For protocol 1, QUASR identified 220 SNVs present in at least 1% of the reads while 908 SNVs were

identified with protocol 2. ShoRAH identified 49 and 39 SNVs using protocol 1 and 2 databases, respectively. Nine sites were in agreement by all methods (Table 1). Of these, SNVs detected at 2082, 9571, 12290 and 12978 were previously identified as vaccine progenitor sequence markers. Nucleotides with higher frequency corresponded to the vaccine strain while those at lower frequency agreed with the vaccine progenitor. The SNV at position 6487 agreed with the sequence detected previously as a marker for a vaccine revertant. SNVs identified by single methods were compared to sites known to be markers of attenuation or reversion. These revealed other 4 and 2 locations, respectively proposed as involved in attenuation and reversion process [4], displaying sub-population variability (Table 2 and 3).

Discussion

The results of the 2 different sequence protocols and 2 downstream analysis techniques produced a range of sometimes conflicting data. Nonetheless, at about 10 genome positions all four sets of results agreed and these form the basis of following discussion. Subpopulations were detected in the vaccine where sequences agreed with the consensus sequence of the wild type vaccine progenitor [11] at four positions. In poultry these populations might have been expected to have had greater fitness and hence preferentially replicated and caused disease. However the consensus sequence of the sequenced revertant vaccine did not agree with vaccine progenitor sequence [4] in the positions detected as variable in the present study, thus indicating that these subpopulations were not responsible for the reversion to virulence.

This apparently conflicts with a subpopulation study of subpopulations in a subtype A candidate vaccine [5] which showed a small subpopulation (1 in 10^{6}), at levels undetectable using the sequence methods used in the current study, were responsible for the return to virulence. This virus was then plaque purified to remove such subpopulations prior to vaccine registration and release. In experimental conditions, this resultant product proved to much reduce the tendency to revert, but it did not prevent this entirely because after release of this plaque purified vaccine, virulent reverts were detected in the field. This unwanted finding confirmed that subpopulation removal by plaque purification did not prevent vaccine reversion. Consensus sequencing of progenitor, vaccine and reverts failed to detect progenitor virulence associated sequence markers matching those found in the revertant virus [3]. Furthermore when 2 field revertants of this vaccine were

independently collected and sequenced, no mutations associated with one reversion were shared in the reversion of the other [12]. These findings taken as a whole indicate that virus evolution was likely to have been the cause of field virulence; at least for subtype A vaccine in question.

For the subtype B virus of this study, it remains possible to suggest that undetectable subpopulations containing revertant virus sequence were already present in the progenitor viruses, and that these were preferentially selected as the virulence of the vaccine in the field increased. Interestingly, two SNVs matching the consensus sequences of the revertant were detected, though there is currently no evidence that these were present in the vaccine progenitor. Deeper sequencing of the progenitor might detect these if present.

It would be interesting to sequence several virulent revertants arising from the subtype B vaccine, much as has been started concerning the discussed subtype A type vaccine studies. The authors are not aware that the subtype B vaccine under investigation here had been plaque purified, hence the effect of small subpopulations remain a possibility. Consensus sequencing of viruses arising during a number of reversion events greater than the likely number of subpopulations, might start to address this issue. If some revertant viruses were found to share SNV motifs, this would appear to indicate that a limited number of virulent subpopulations were present in the vaccine. Unfortunately the detection of identical sequences during independent reversion events might also be considered the result of the virus system only permitting a limited number of mutation combinations in the return to virulence. As a way forward, sufficiently deep sequencing of viruses generated at all stages of a reversion event should show whether virulent motifs occurred simultaneously or serially, and thus confirm which mechanism to be responsible. Simultaneous detection of SNVs would suggest a subpopulation mechanism while sequential acquisition of such sequences would indicate evolution.

In conclusion, consideration of previous findings leads to the view that subpopulations are not required to encounter reversion to virulence in AMPV vaccines. The current study concerning a subtype B vaccine detected subpopulations matching some revertant virus sequence that were apparently not present in the virulent virus progenitor. If the latter are present, much deeper sequencing will be needed to detect them. Additionally, the evidence that SNVs detected in the vaccine, matching the progenitor consensus sequence, have not

been identified in the revertant virus sequence, indicates that these subpopulations were not responsible for the reversion to virulence .While this leaves the situation currently unresolved, it is likely to be of significance to note the ease of detecting revertant viruses arising from a cloned subtype A type vaccine, as this suggests that evolution rather than sub populations is likely to be playing a major role. Nonetheless, until a further deeper sequencing study has been performed, it would be prudent to routinely plaque purify all AMPV candidate vaccines so as to preclude that route to reversion to virulence.

References

[1] Gough RE, Jones RC. Diseases of Poultry. : Wiley-Blackwell, 2008.

[2] Cecchinato M, Lupini C, Ricchizzi E, Falchieri M, Meini A, Jones RC et al. Italian field survey reveals a high diffusion of avian metapneumovirus subtype B in layers and weaknesses in the vaccination strategy applied. Avian Dis 2012;56(4):720-4.

[3] Catelli E, Cecchinato M, Savage CE, Jones RC, Naylor CJ. Demonstration of loss of attenuation and extended field persistence of a live avian metapneumovirus vaccine. Vaccine 2006;24(42-43):6476.

[4] Cecchinato M, Catelli E, Lupini C, Ricchizzi E, Prosperi S, Naylor CJ. Reversion to virulence of a subtype B avian metapneumovirus vaccine: Is it time for regulators to require availability of vaccine progenitors? Vaccine 2014;32(36):4660-4.

[5] Naylor CJ, Jones RC. Demonstration of a virulent subpopulation in a prototype live attenuated turkey rhinotracheitis vaccine. Vaccine 1994;12(13):1225-30.

[6] Victoria JG, Kapoor A, Li L, Blinkova O, Slikas B, Wang C et al. Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. J Virol 2009;83(9):4642-51.

[7] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30(15):2114-20.

[8] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9(4):357-9.

[9] Zagordi O, Bhattacharya A, Eriksson N, Beerenwinkel N. ShoRAH: estimating the genetic diversity of a mixed sample from next-generation sequencing data. BMC Bioinformatics 2011;12:119,2105-12-119.

[10] Watson SJ, Welkers MR, Depledge DP, Coulter E, Breuer JM, de Jong MD et al. Viral population analysis and minority-variant detection using short read next-generation sequencing. Philos Trans R Soc Lond B Biol Sci 2013;368(1614):20120205.

[11] Sugiyama M, Ito H, Hata Y, Ono E, Ito T. Complete nucleotide sequences of avian metapneumovirus subtype B genome. Virus Genes 2010;41(3):389-95.

[12] Lupini C, Cecchinato M, Ricchizzi E, Naylor CJ, Catelli E. A turkey rhinotracheitis outbreak caused by the environmental spread of a vaccine-derived avian metapneumovirus. Avian Pathol 2011;40(5):525-30.

	QUASR						ShoRAH					
	Protocol1			Protocol2			Protocol1			Protocol2		
Position	Reference	Variant	Coverage	Reference	Variant	Coverage	Reference	Variant	Coverage	Reference	Variant	Coverage
2082	T	C(12%)	198	T	C(14%)	469	T	C(15%)	100	T	C(16%)	281
4921	A	G(4%)	197	A	G(4%)	246	A	G(7%)	101	A	G(4%)	190
4934	C	T(17%)	197	C	T(4%)	247	C	T(25%)	101	C	T(5%)	186
6346	T	A(10%)	196	T	A(14%)	227	T	A (24%)	101	T	A(16%)	122
6487	T	G(22%)	197	T	G(31%)	427	T	G (38%)	101	T	G(24%)	238
9571	A	T(9%)	194	A	T(11%)	124	A	T(22%)	101	A	T(23%)	71
9610	T	C(7%)	194	T	C(5%)	162	T	C(9%)	101	T	C(14%)	77
12290	A	T(14%)	193	T	A(39%)	166	T	A(68%)	99	T	A(39%)	111
12978	A	G(6%)	197	A	G(38%)	124	A	G(30%)	101	A	G(40%)	73

Table 1. Sites detected as SNV by all combinations of NGS protocols and data analysis. Positions are numbered according to the reference strain (Cecchinato et al., 2014). The coverage, reference nucleotide and variant are reported for all methods, together with the frequency of the less represented sequence (between brackets).

	QUASR						ShoRAH					
	Protocol1			Protocol2			Protocol1			Protocol2		
Position	Reference	Variant	Coverage	Reference	Variant	Coverage	Reference	Variant	Coverage	Reference	Variant	Coverage
1461	A	C (1%) T(4%)	190	A	G (1%)	280						
3508	-	-	-									
3566	C	T(100%)		C	T(100%)		C	T(100%)		C	T(100%)	
4348				A	G (5%)	342						
5614												
6487	T	G(22%)	197	T	G(31%)	427	T	G (38%)	101	T	G(24%)	238
12277												
12847												

Table 2. Sites involved in reversion to virulence (Cecchinato et al., 2014) detected by at least one method. Positions are numbered according to the reference strain (Cecchinato et al., 2014). The coverage, the reference nucleotide and the variant are reported for all methods together with the frequency of the less represented sequence (between brackets).

	QUASR						ShoRAH					
	Protocol1			Protocol2			Protocol1			Protocol2		
Position	Reference	Minor	Coverage	Reference	Minor	Coverage	Reference	Minor	Coverage	Reference	Minor	Coverage
1220												
2082	T	C(12%)	198	T	C(14%)	469	T	C(15%)	100	T	C(16%)	281
2422												
2554	T	C (3%)	192									
2654												
3210				T	G (3%)	281						
3391												
5434	G	A (11%)	194				G	A (9%)	101			
5777	T	A (4%)	199				T	A (4%)	101			
6104												
6712												
9571	A	T(9%)	194	A	T(11%)	124	A	T(22%)	101	A	T(23%)	71
9817												
12290	A	T(14%)	193	T	A(39%)	166	T	A(68%)	99	T	A(39%)	111
12978	A	G(6%)	197	A	G(38%)	124	A	G(30%)	101	A	G(40%)	73
13463												
13464												

Table 3. Sites involved in the attenuation process (Cecchinato et al., 2014) detected as SNV by at least one method. Positions are numbered according to the reference strain (Cecchinato et al., 2014). The coverage, the reference nucleotide and the variant nucleotide are reported for all methods together with the frequency of the less represented sequence (between brackets).

Vaccine 32 (2014) 6765–6767

Contents lists available at ScienceDirect

Vaccine

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

Continued use of IBV 793B vaccine needs reassessment after its withdrawal led to the genotype's disappearance

Giovanni Franzo[a,*], Clive John Naylor[b], Caterina Lupini[c], Michele Drigo[a], Elena Catelli[c], Valeria Listorti[c], Patrizia Pesente[d], Davide Giovanardi[d], Emilio Morandini[e], Mattia Cecchinato[a]

[a] Department of Animal Medicine, Production and Health, University of Padua, Viale dell'università, 16, 35020 Legnaro, PD, Italy
[b] Department of Infection Biology, Faculty of Health and Life Sciences, University of Liverpool, Leahurst Campus, Neston CH64 7TE, United Kingdom
[c] Department of Veterinary Medical Sciences – University of Bologna, Via Tolara di Sopra, 50, 40064 Ozzano Emilia, BO, Italy
[d] Tre Valli Laboratory, Viale A. Veronesi, 5, 37132 San Michele Extra, VR, Italy
[e] Field veterinary, 37026, Pescantina, VR, Italy

ARTICLE INFO

Article history:
Received 2 September 2014
Received in revised form 1 October 2014
Accepted 6 October 2014
Available online 22 October 2014

Keywords:
Infectious bronchitis virus
Broiler farm survey
S1 sequence
793B vaccination discontinued
793B genotype disappeared

ABSTRACT

Over a period of almost two years, broilers chickens on several hundred Italian farms, were monitored for infectious bronchitis virus. Detections were genotyped using a hypervariable region of the gene coding for the S1 segment of the spike protein. A range of genotypes were detected which comprised QX, Q1, Mass, D274 and 793B. Sequences of 793B viruses detected in chickens, vaccinated with either of the two commonly used 793B type vaccines were almost identical to sequences of one or other of these vaccines. This strong indication of vaccine association led to the withdrawal of live 793B vaccine use on all of the farms of the study. Except for one sample collected soon after 793B vaccination ceased, it was no longer possible to detect 793B vaccine on these farms. It appears that field 793B strains have disappeared from the region of Italy tested thus obviating any need for current vaccine protection against 793B.

1. Introduction

Infectious bronchitis virus (IBV) is highly contagious and causes serious economic losses for the worldwide chicken industry. It replicates in the respiratory, urogenital and gastroenteric tracts and can cause lesions in these and other regions. As for other coronaviruses, IBV evolves via high mutation and recombination rates [1–4]. Infection with one IBV protectotype generally only confers protection again that same protectotype. Other IBVs breaking through that protection are considered to have a different protectotype. Sequencing of the IBV RNA genome has shown these protectotypes roughly correlate to sequence differences between their spike surface proteins, in particular hyper-variable regions of S1 which have also been implicated in virus attachment and neutralization epitopes [5–7]. Numerous genotypes have been detected, some of which have quickly disappeared while others have caused major worldwide disease with economic relevance [8–10].

Vaccination is generally considered essential for the protection of commercial chickens [11]. The limited cross-protection between genotypes poses a great challenge in the control of this disease, so knowledge of field genotypes can be invaluable in any effective vaccination strategy. However, the presence of live vaccines confuses most such epidemiological studies due to the impossibility of differentiating vaccine from field strains. Our study aimed to investigate the circulation in broilers of different genotypes before and after the discontinuing of live 793B-based vaccination.

2. Materials and methods

2.1. Samples

Between November 2012 and June 2014 on 513 farms located mainly in Northern Italy, pools of tracheal swabs were collected from broilers in response to clinical episodes thought potentially attributable to IBV. This was part of systematic diagnostic activity.

On the day of collection, IBV real time RT-PCR was performed on tracheal swabs.

At the time of sampling, age, clinical status and geographical location of chickens were recorded. All were vaccinated with live

* Corresponding author. Tel.: +39 0498272968; fax: +39 0498272973.
E-mail address: giovanni.franzo@studenti.unipd.it (G. Franzo).

http://dx.doi.org/10.1016/j.vaccine.2014.10.006
0264-410X/

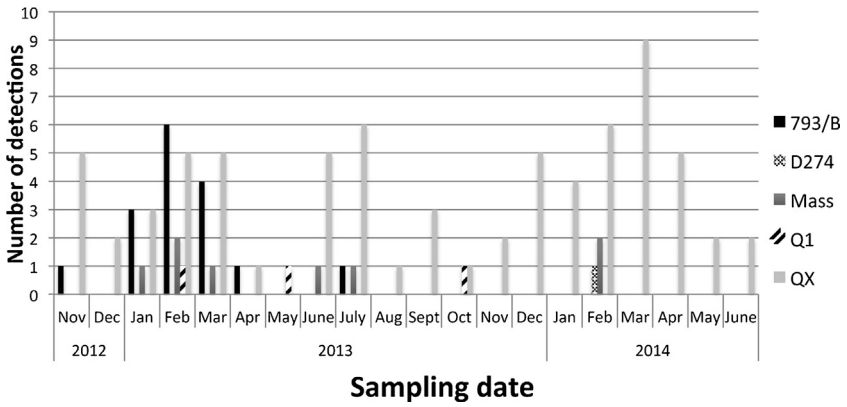

Fig. 1. Distribution over time of detected genotypes in Italian broiler farms.

H120 vaccine at the hatchery and, until late spring 2013, a live 793B booster (A and B) via drinking water. The 793B vaccination was introduced in Italy in the late '90 after the detection of this genotype in the field.

2.2. IBV detection by real time RT-PCR

Pools of ten tracheal swabs from the same flock were vortexed in 1 ml of PBS. The solution was used for RNA extraction using NucleoSpin® 8/96 RNA (Macherey-Nagel, Düren, Germany). Samples were tested for IBV using a real time RT-PCR commercial kit (Quantification of Avian Infectious Bronchitis Virus-IBV-kit; Genesig, Southampton, UK) following manufacturer instruction.

2.3. Sequencing of partial S1 gene

RT-PCRs covering a hypervariable region of the S1 gene were carried out on all real time RT-PCR positive samples using primer pairs SX3+ (5′-TAATACTG GC/T AATTTTTCAGA-3′) and SX2− (5′-TCCACCTCTATAAACACC C/T TT-3′). Nested PCRs were performed where results were weak or negative [12]. Amplicons were sequenced in both directions with the same primers used for PCR. Chromatograms were evaluated with FinchTV (http://www.geospiza.com) and consensus sequences were obtained using CromasPro (CromasPro Version 1.5)

2.4. IBV strain genotyping

IBV strains were initially genotyped at first performing a BLAST analysis (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Samples were more rigorously genotyped by comparison with a collection of reference sequences. For 793B the full range of database available S1 sequences were used. Briefly sequences were aligned using the MUSCLE method implemented in MEGA6 [13] and phylogenetic analysis was performed using the same software. To clarify any relationship to 793/B vaccines, the number of mismatches between samples and 793B vaccine in the sequenced region were calculated.

3. Results

3.1. IBV genotypes and distribution

One hundred and twenty-nine samples were positive to IBV by real time RT-PCR. Partial S1 gene sequences were obtained from 100 IBV strains.

Five genotypes were identified. The genotype with the highest prevalence was QX (72%), followed by 793/B (16%), Mass (8%), Q1 (3%) and D274 (1%). Genotype distribution from November 2012 is given in Fig. 1 where it can be seen that QX, Mass, D274 and Q1 were detected throughout, whereas 793/B was no longer detected after July 2013. 793B viruses were numbered according to chronological detection and classified as N (north) or C (central) according to the region where viruses were detected. Two 793/B clusters were identified, with the larger of 14 sequences being related to vaccine A and the smaller being related to vaccine B (Fig. 2).

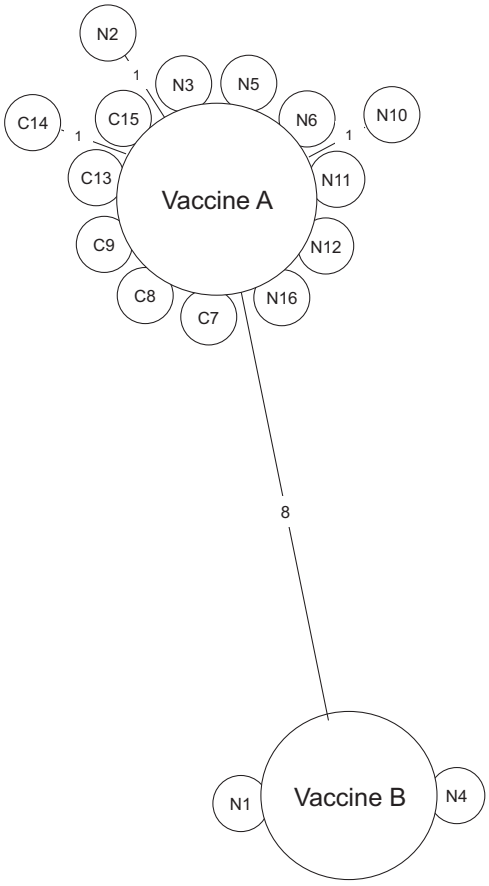

Fig. 2. Identities between 793B viruses detected on Italian broiler farms. Touching circles indicate identical sequences. Numbers indicate nucleotide differences between viruses in the tested region. Strains are named progressively according to their collection date and geographical locations are codified as N (North Italy) or C (Central Italy).

4. Discussion

For about half a century, countries with developed poultry industries have applied one or more live vaccines to all broiler, layer and breeder chickens to control disease caused by infectious bronchitis virus. More recently the availability of molecular diagnostics able to distinguish between genotypes has allowed those IBVs actually present on farms to be readily identified, and this has allowed the vaccine genotype to match that of any current field threat. Furthermore the technique has allowed the detection of new genotypes of which some might cause disease. Of these a proportion can become largely undetectable within a short period while others can become a dominant disease causing genotype persisting for a considerable period. Italy 02 would be a recent example of the former in most countries while 793B would be generally considered to fall into the latter category and present a threat over the longer term. This has either led to the development of new live vaccines of matching genotype or the deployment of established vaccines, sometimes in specific combination, to counter that disease threat. In general such an approach has allowed the poultry industry to control the effects of infection from both established and more recently detected IBV genotypes.

A weakness of the system has been the inability to unequivocally distinguish between field and vaccine viruses. For almost two decades 793B detection has been associated with disease and for most of that period one or more live 793B vaccines have been in use. In that time many 793B detections have been made. Currently our study as well as another study [14] following a large number of unrelated farms in Italy, has found consistent detections of 793B. However it has not been possible to know whether encountered disease has been caused by inadequately controlled 793B field viruses or persisting, and possibly unstable, vaccine viruses. This is in contrast with, for example, AMPV, another common virus affecting poultry worldwide, where the deduction of genetic vaccine markers has shown that live vaccine derived virus can persist on farms and cause disease [15–18]. While it should be possible to similarly identify markers for IBV vaccines, this has not to date been reported, probably due to the unavailability of the IBV vaccine progenitor viruses needed in the critical vaccine-progenitor sequence comparisons.

Results very strongly suggest that 793B detected on the farms currently studied, derived from the two 793B vaccines in use. This conclusion rests on four pieces of evidence of ascending importance. Firstly, in the Massi study [14] where chickens were given more doses of vaccine, a greater proportion of 793B detections were made. Secondly, all 793B detections in the current study had S1 hypervariable region sequences almost identical to either of the two vaccines. Uninvolved assessment might correctly raise the possibility that many 793B field viruses have equal identity in this region, and to confuse the picture further, it would be expected that mass use of vaccines A and B for some two decades, might have led to vaccine derived detections being added to the world sequence databases as field strains. Unfortunately this sequence issue cannot be unequivocally resolved without the availability of vaccine markers. Thirdly the two vaccines differed by eight nucleotides in the tested region and all detections matched the sequence of the vaccine which had been applied. Nonetheless the main evidence arises from 793B becoming undetectable after the withdrawal of the two 793B vaccines on several hundred Italian farms. It is reasonable to conclude that 793B vaccine use is no longer required for 793B virus control at the current time on these farms.

The paper confirms the importance to live vaccination strategies of distinguishing genuine field viruses from applied vaccines and their derivatives. For example for in human polio, vaccination strategies have needed to consider the occurrence of vaccine-associated paralytic poliomyelitis, which in turn has required the technical expertise for vaccine identification [19]. In another avian viral disease of poultry AMPV, the ability to clearly distinguish between vaccine and field virus has become practical and this would be expected to have benefits in future vaccination strategies [18]. For IBV it would be clearly undesirable to be adding millions or billions of doses of live vaccine virus to birds if, as results here suggest, they served little benefit. Furthermore the presence of unnecessary vaccine viruses, especially where multiple IBV genotypes might be involved, might be expected to drive the generation of recombinant viruses; indeed this notion has restricted certain concurrent IBV vaccination strategies in France. While to date there is no evidence of field recombinants with vaccine virus contribution causing major disease in the worldwide chicken industry, the possible mechanism is there, hence vaccine use should be ideally restricted to situations where there is a known protective benefit.

Acknowledgement

The work here described was supported by University of Padua grant 60A08-8005/14 (ex 60%, 2014).

References

[1] Jia W, Karaca K, Parrish CR, Naqi SA. A novel variant of avian infectious bronchitis virus resulting from recombination among three different strains. Arch Virol 1995;140(2):259–71.
[2] Cavanagh D, Mawditt K, Adzhar A, Gough RE, Picault JP., Naylor CJ, et al. Does IBV change slowly despite the capacity of the spike protein to vary greatly? Adv Exp Med Biol 1998;440:729.
[3] Dolz R, Pujols J, Ordonez G, Porta R, Majo N. Molecular epidemiology and evolution of avian infectious bronchitis virus in Spain over a fourteen-year period. Virology 2008;374(1):50–9.
[4] McKinley ET, Jackwood MW, Hilt DA, Kissinger JC, Robertson JS, Lemke C, et al. Attenuated live vaccine usage affects accurate measures of virus diversity and mutation rates in avian coronavirus infectious bronchitis virus. Virus Res 2011;158(1–2):225–34.
[5] Kant A, Koch G, van Roozelaar DJ, Kusters JG, Poelwijk FA, van der Zeijst BA. Location of antigenic sites defined by neutralizing monoclonal antibodies on the S1 avian infectious bronchitis virus glycopolypeptide. J Gen Virol 1992;73(Pt 3):591–6.
[6] Koch G, Hartog L, Kant A, van Roozelaar DJ. Antigenic domains on the peplomer protein of avian infectious bronchitis virus: correlation with biological functions. J Gen Virol 1990;71(Pt 6):1929–35.
[7] Ignjatovic J, Sapats S. Identification of previously unknown antigenic epitopes on the S and N proteins of avian infectious bronchitis virus. Arch Virol 2005;150(9):1813–31.
[8] Cavanagh D. Coronavirus avian infectious bronchitis virus. Vet Res 2007;38(2):281–97.
[9] Sjaak de Wit JJ, Cook JK, van der Heijden HM. Infectious bronchitis virus variants: a review of the history, current situation and control measures. Avian Pathol 2011;40(3):223–35.
[10] Jackwood MW. Review of infectious bronchitis virus around the world. Avian Dis 2012;56(4):634–41.
[11] Cook J. Coronaviridae. In: Pattison M, Bradbury A, editors. Poultry Diseases. Amsterdam: Saunders Elsevier; 2008. p. 340–9.
[12] Jones RC, Worthington KJ, Capua I, Naylor CJ. Efficacy of live infectious bronchitis vaccines against a novel European genotype, Italy 02. Vet Rec 2005;156(20):646–7.
[13] Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 2013;30(12):2725–9.
[14] Massi P. Situazione epidemiologica della bronchite infettiva in Italia. Rivista di Medicina Veterinaria, Speciale 2013:13–20.
[15] Catelli E, Cecchinato M, Savage CE, Jones RC, Naylor CJ. Demonstration of loss of attenuation and extended field persistence of a live avian metapneumovirus vaccine. Vaccine 2006;24(42–43):6476.
[16] Listorti V, Lupini C, Cecchinato M, Pesente P, Rossi G, Giovanardi D, et al. Rapid detection of subtype B avian metapneumoviruses using RT-PCR restriction endonuclease digestion indicates field circulation of vaccine-derived viruses in older turkeys. Avian Pathol 2014;43(1):51–6.
[17] Lupini C, Cecchinato M, Ricchizzi E, Naylor CJ, Catelli E. A turkey rhinotracheitis outbreak caused by the environmental spread of a vaccine-derived avian metapneumovirus. Avian Pathol 2011;40(5):525–30.
[18] Cecchinato M, Catelli E, Lupini C, Ricchizzi E, Prosperi S, Naylor CJ. Reversion to virulence of a subtype B avian metapneumovirus vaccine: is it time for regulators to require availability of vaccine progenitors? Vaccine 2014;32(36):4660–4.
[19] Kew OM, Sutter RW, de Gourville EM, Dowdle WR, Pallansch MA. Vaccine-derived polioviruses and the endgame strategy for global polio eradication. Annu Rev Microbiol 2005;59:587–635.

Diffusione dell'infezione da metapneumovirus aviare in allevamenti di tacchini e broiler nel Nord Italia

M. CECCHINATO[1], M. DRIGO[1], C. LUPINI[2], M. MARTINI[1], V. LISTORTI[2], G. FRANZO[1], M. BONCI[2], A. LACONI[2], E. MORANDINI[3], E. CATELLI[2]

[1] Dipartimento di Medicina Animale, Produzioni e Salute (MAPS), Università degli Studi di Padova Agripolis - Viale dell'Università, 16, 35020, Legnaro (PD), Italia
[2] Dipartimento di Scienze Mediche Veterinarie, Alma Mater Studiorum - Università di Bologna - Via Tolara di Sopra, 5 - 40064 Ozzano Emilia (BO) - Italia
[3] Veterinario avicolo

RIASSUNTO

Il metapneumovirus aviare (AMPV) è noto come agente eziologico della Rinotracheite del Tacchino (TRT), un'infezione del tratto respiratorio superiore del tacchino caratterizzata da elevata morbilità e da mortalità variabile, dovuta comunemente a complicazioni batteriche secondarie. Dati sistematici relativi alla diffusione e tipizzazione di AMPV in Italia sia nel pollo che nel tacchino risalgono ormai al 2006.

Allo scopo di colmare tale carenza e di fornire utili dati di campo, è stata eseguita un'indagine epidemiologica per AMPV tramite metodiche diagnostiche biomolecolari. Tutti gli allevamenti campionati (122 di tacchini da carne e 48 di broiler), ad eccezione di un allevamento di tacchini da carne in cui è stato evidenziato AMPV sottotipo A, sono risultati positivi ad AMPV sottotipo B confermando la prevalenza di questo sottotipo in Italia.

Positività per AMPV sono state riscontrate in tacchini da carne dalla prima alla quattordicesima settimana di età, con maggior frequenza (73/122; 59,8%) dalla nona alla undicesima. Tredici ceppi di origine vaccinale (ceppo VCO3) sono stati evidenziati in animali di età compresa tra le due e le dodici settimane. La maggior parte di questi (7/13; 53,8%) è stata rilevata nelle prime sei settimane di vita. In nove casi la presenza di ceppi di origine vaccinale era associata a sintomatologia respiratoria. Nel broiler, AMPV-B è stato riscontrato dalla quarta alla ottava settimana di vita. Il maggior numero di positività (24/48; 50%) si è rilevato alla sesta settimana di vita. Sulla base dei dati ottenuti da questa indagine vengono fatte alcune considerazioni sulla circolazione di virus vaccinali in campo e sulla presenza di sintomatologia respiratoria da imputare a infezioni da AMPV di campo in soggetti vaccinati.

PAROLE CHIAVE

Avian Metapneumovirus, broiler, tacchini da carne, ceppi di origine vaccinale, Nord Italia.

INTRODUZIONE

Il metapneumovirus aviare (AMPV) è noto come agente eziologico della Rinotracheite del Tacchino (*Turkey Rhinotracheitis*, TRT), un'infezione del tratto respiratorio superiore del tacchino caratterizzata da elevata morbilità e da mortalità variabile, dovuta comunemente a complicazioni batteriche secondarie. Nei riproduttori il virus è in grado di causare una riduzione della produzione e qualità delle uova. Oltre al tacchino, sono sensibili all'infezione anche altre specie aviari tra cui principalmente il pollo, in cui AMPV è coinvolto nella Sindrome della Testa Gonfia (*Swollen Head Syndrome*, SHS). Sono stati sino ad ora identificati, in base alle sequenze nucleotidiche di AMPV, quattro sottotipi virali A, B, C e D[1].

La TRT viene segnalata per la prima volta alla fine degli anni '70 in Sud Africa per poi diffondersi rapidamente in tutti i continenti, ad eccezione dell'Oceania. Oggi rappresenta una delle principali forme respiratorie ad eziologia virale dell'allevamento del tacchino[1].

In Italia l'infezione è stata osservata a partire dal 1987[2] e successivamente si è diffusa ed è diventata endemica nelle regioni a maggior vocazione avicola quali Lombardia, Veneto ed Emilia Romagna, acquisendo notevole importanza fra le problematiche sanitarie del tacchino[3,4]. Il sottotipo B è il prevalente in Nord Italia[3,4,5,10,16] mentre il sottotipo A, dopo la sua comparsa nel 2003[11], è stato segnalato sporadicamente[5,12,13].

Per la profilassi della TRT si è precocemente ricorsi all'impiego di vaccini e, fra quelli disponibili nel nostro Paese, quello maggiormente utilizzato appartiene al sottotipo B (ceppo VCO3). Nei primi anni, allo scopo di determinare la migliore modalità di somministrazione, il vaccino fu impiegato a "spot", per via oculare o spray, a dosi variabili e non standardizzate. Solo dalla fine del 2001, la vaccinazione ha cominciato ad essere praticata a tappeto, ad un giorno di vita, via spray, in incubatoio, ed ha contribuito a tenere in gran parte la malattia sotto controllo. Tuttavia dati epidemiologici recenti hanno mostrato come si evidenzino ancora con una certa frequenza focolai di malattia in animali vaccinati. Dati sistematici relativi alla diffusione e tipizzazione di AMPV in Italia sia nel pollo che nel tacchino risalgono ormai al 2006[5].

Autore per la corrispondenza:
Michele Drigo (michele.drigo@unipd.it).

Per colmare tale carenza e per fornire utili dati di campo, è stata eseguita un'indagine epidemiologica per AMPV allo scopo di evidenziare la diffusione dei due sottotipi di AMPV presenti in Italia nell'allevamento sia del tacchino da carne che del broiler, la distribuzione delle positività per fasce di età e, esistendo un test diagnostico discriminatorio[6], la diffusione di ceppi di origine vaccinale.

MATERIALI E METODI

Allevamenti

Sono stati utilizzati per la presente indagine epidemiologica 48 gruppi di broiler e 122 di tacchini da carne del Nord Italia (in prevalenza provenienti dal Veneto e dalla Lombardia) risultati positivi ad AMPV nell'ambito di attività diagnostica routinaria svolta nel periodo gennaio 2011-febbraio 2013. Per ciascun gruppo sono stati raccolti dati relativi all'età dei soggetti, alla sintomatologia e alla vaccinazione eseguita per il controllo delle infezioni da AMPV. Non vengono quindi esposti dati di prevalenza ma dati relativi alla circolazione dei diversi sottotipi di AMPV in allevamenti dell'area geografica considerata.

Campionamento

Per ciascun gruppo sono stati campionati, mediante tampone rino-faringeo, 10 volatili. I tamponi, dopo prelievo, sono stati lasciati asciugare all'aria per 30 minuti e conservati a temperatura ambiente sino alla processazione eseguita in pool da 10 per ciascun gruppo. Ogni allevamento è stato oggetto di campionamento una sola volta per ciclo.

Nested RT-PCR e qRT-PCR

Per evidenziare e tipizzare AMPV è stato impiegato, su 121 campioni, un protocollo di nested RT-PCR, disegnata sulla sequenza del gene G ed in grado di differenziare i sottotipi A e B. L'estrazione dell'RNA virale dai tamponi a secco e la retrotrascrizione sono state eseguite secondo il metodo descritto da Cavanagh *et al.*[7] La nested PCR è stata eseguita secondo quanto descritto da Naylor *et al.*[8] Il prodotto dell'amplificazione, dopo corsa elettroforetica in gel di agarosio al 2%, è stato visualizzato mediante colorazione con bromuro di etidio e successivamente, in caso di positività per AMPV sottotipo B, sottoposto al protocollo di restrizione enzimatica riportato di seguito.

Poiché nella routine diagnostica è stata introdotta una metodica di qRT-PCR[9] nell'ultimo periodo cui si riferisce il presente lavoro, quarantanove campioni sono stati processati con tale tecnica, anch'essa in grado di evidenziare e differenziare i sottotipi A e B. Tali campioni non sono stati sottoposti ad analisi di restrizione enzimatica.

Analisi di restrizione enzimatica

Per differenziare i ceppi di campo da quelli vaccinali (ceppo VCO3) è stato utilizzato un protocollo di restrizione enzimatica messo a punto da Listorti *et al.*[6] e brevemente descritto di seguito.

L'enzima di restrizione *Tru 9I* (sito di riconoscimento 5'T↓TAA3') è stato utilizzato per digerire i prodotti ottenuti mediante RT nested-PCR. Il protocollo di restrizione ha previsto la digestione a 65°C per 2 ore di 5 µl di prodotto di amplificazione con 5 unità di enzima di restrizione in un volume finale di reazione di 20 µl. Dopo la digestione, i frammenti di DNA ottenuti sono stati separati in gel di agarosio al 2% e visualizzati come precedentemente descritto. La presenza di una adenina (A) in posizione 91 della sequenza nucleotidica del gene G del ceppo vaccinale fa sì che il prodotto di RT nested-PCR venga tagliato dall'enzima in due frammenti di 300 e 40 pb rispettivamente. L'avvenuta digestione enzimatica e quindi il riconoscimento dei ceppi di origine vaccinale, è stata verificata mediante confronto col prodotto di PCR non sottoposto a taglio enzimatico e marker di riferimento (Φ X 174 RF DNA/ Hae III Fragments, INVITROGEN®, Carlsbad, California, USA).

RISULTATI

Tutti i gruppi, ad eccezione di un gruppo di tacchini della provincia di Verona in cui è stato evidenziato AMPV sottotipo A, sono risultati positivi ad AMPV sottotipo B utilizzando le due metodiche molecolari applicate nel presente studio.

I gruppi di tacchini erano stati tutti vaccinati a 1 gg di vita via spray in incubatoio con AMPV sottotipo B (ceppo VCO3) mentre i broiler non avevano ricevuto nessuna vaccinazione nei confronti di AMPV.

La distribuzione delle positività nelle diverse settimane di vita nei tacchini e nei broiler sono riportate rispettivamente nelle Figure 1 e 2.

AMPV è stato evidenziato nei tacchini dalla prima alla quattordicesima settimana di vita, con maggior frequenza dalla nona all'undicesima (73/122; 59,8%). I 90 campioni di tacchino risultati positivi alla nested RT-PCR per AMPV-B sono stati sottoposti a digestione enzimatica, ciò ha permesso di evidenziare 13 ceppi di origine vaccinale in animali di età compresa tra le due e le dodici settimane. La maggior parte di questi (7/13; 53,8%) è stata evidenziata nelle prime sei settimane di vita. In nove casi la presenza di ceppi di origine vaccinale era associata a sintomatologia respiratoria. I ceppi AMPV di campo erano evidenziati in gruppi con sintomatologia respiratoria ed aumento di mortalità.

Nei gruppi di broiler AMPV-B era evidenziato dalla quarta all'ottava settimana di vita, con massima concentrazione delle positività nella sesta settimana di vita (24/48; 50%). Un ceppo di AMPV-B rinvenuto in corso di sintomatologia respiratoria alla quinta settimana di vita è risultato essere di origine vaccinale. In alcuni casi AMPV era associato a Sindrome della testa gonfia. Tale forma clinica è stata riscontrata sia in corso di infezioni pure da AMPV che di infezioni miste (*E. coli*, Virus della bronchite infettiva e *Mycoplasma synoviae*).

DISCUSSIONE

L'indagine svolta nel presente lavoro ha confermato l'ampia diffusione di infezione da AMPV nel nostro Paese e in particolare in regioni a maggior vocazione avicola quali Lombardia e Veneto. Il sottotipo B è risultato essere quello prevalente, confermando quanto già riportato in letteratura[3,5,10], mentre il sottotipo A è stato evidenziato una sola volta. Quest'ultimo sottotipo, segnalato in Italia per la prima volta nel 2003[11], è riportato solo sporadicamente[5,12,13]. La distribuzione geografica di queste segnalazioni fa riferimento solamen-

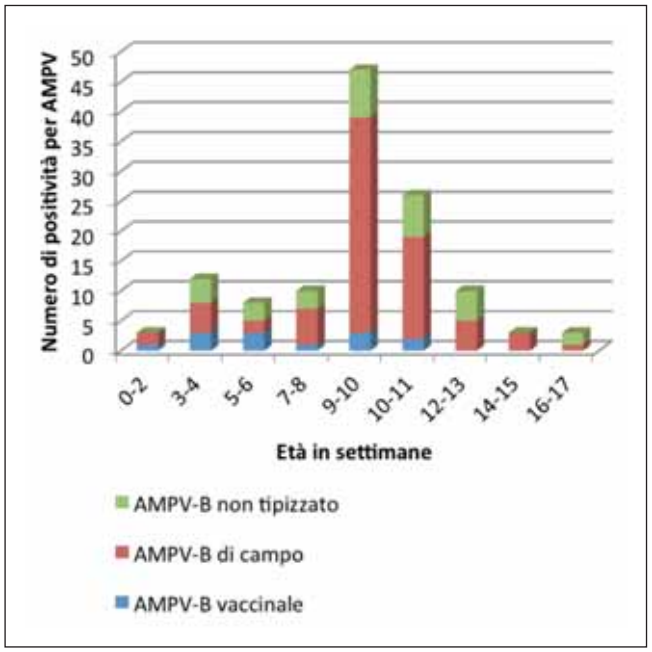

Figura 1 - Distribuzione per età della positività per AMPV sottotipo B in allevamenti tacchini da carne.

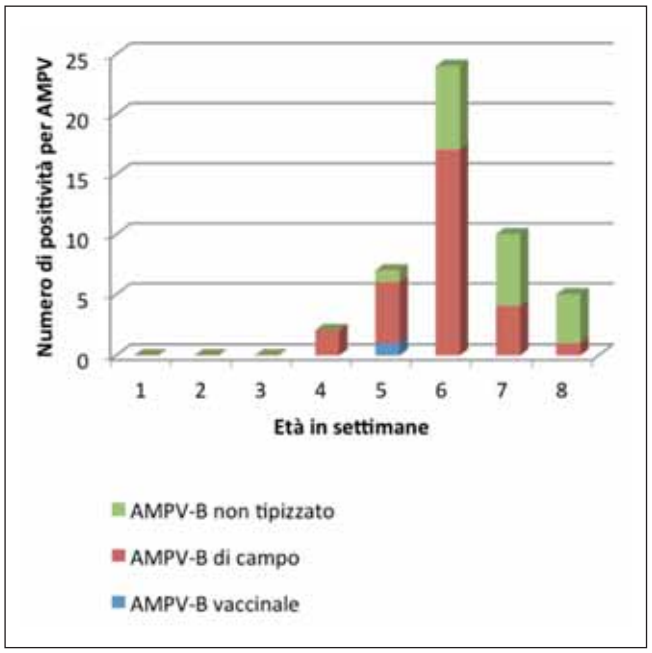

Figura 2 - Distribuzione per età della positività per AMPV sottotipo B in allevamenti di broiler.

te a Toscana ed Emilia Romagna anche se è plausibile una circolazione anche nelle aree cui fa riferimento il presente lavoro. Studi sulla sequenza nucleotidica di due di questi ceppi hanno permesso di dimostrare la loro derivazione vaccinale[12,13]. Ciò fa ipotizzare che la circolazione di tale sottotipo sia da ascrivere solamente a stipiti di origine vaccinale. Sfortunatamente non è stato possibile confermare con il sequenziamento la natura del ceppo dai noi evidenziato.

La maggior parte dei ceppi di AMPV sottotipo B riscontrati sono risultati, alla digestione enzimatica, essere ceppi di campo. Essi erano prevalenti in tacchini di 9-12 settimane di età con sintomatologia tipica da TRT. L'osservazione di focolai di TRT in soggetti vaccinati potrebbe essere imputata a varie cause quali il metodo e il piano di vaccinazione, la scarsa durata dell'immunità o l'elusione dell'immunità vaccinale conseguente ad evoluzione genomica del virus di campo[14,15]. A supporto di quest'ultima ipotesi, studi più recenti hanno dimostrato che la vaccinazione col ceppo AMPV-B VCO3 non riesce a dare protezione completa nei confronti di infezioni sperimentali con i ceppi attualmente circolanti nel nostro Paese[16] essendo questi evoluti in regioni antigeniche fondamentali[4] tali da determinare l'elusione della risposta immunitaria vaccinale.

La variabilità e in alcuni casi la gravità della sintomatologia osservata è sicuramente da imputare a fattori gestionali di allevamento quali la ventilazione, l'igiene e la densità animale che condizionano l'insorgenza di infezioni batteriche secondarie, da noi riscontrate in alcuni allevamenti positivi ad AMPV (dati non riportati).

I ceppi AMPV-B di origine vaccinale sono stati evidenziati nel tacchino con una più alta frequenza nelle prime settimane di vita e quindi a breve distanza dalla vaccinazione eseguita in incubatoio. Questo dato è in accordo con quanto osservato da altri autori che hanno messo in evidenza ceppi vaccinali di AMPV sottotipo A e B a 4-5 settimane di età[12,15,16]. Non sono però mai state riscontrate positività di origine vaccinale in tacchini piuttosto avanti nel ciclo produttivo (84 giorni di età) o broiler non vaccinati, come successo in questo studio. Nel primo caso appare inverosimile come un ceppo vaccinale possa persistere per un tempo così lungo dopo la vaccinazione. La spiegazione più plausibile, recentemente riportata per un ceppo vaccinale del sottotipo A[13], è che il vaccino circoli nell'ambiente ed entri successivamente in allevamento eludendo le misure di biosicurezza.

Nove allevamenti risultati positivi ad AMPV-B di origine vaccinale hanno mostrato sintomatologia respiratoria e/o aumento di mortalità. Questi allevamenti, testati per la presenza del virus dell'influenza aviaria, della malattia di Newcastle, di *Mycoplasma gallisepticum* e *Mycoplasma synoviae* hanno dato esito negativo (dati non riportati). Tali risultati, associati a quanto descritto in bibliografia riguardo alla possibilità che un vaccino vivo attenuato per AMPV sottotipo A possa persistere in allevamento ed andare incontro in condizioni di campo a reversione di virulenza[12], lasciano ipotizzare una situazione analoga anche per vaccini AMPV-B, già sospettata da altri autori[15,17,18]. Per verificare l'ipotesi di una riacquisizione di virulenza da parte del vaccino vivo attenuato, ed escludere l'influenza di fattori manageriali e/o la concomitante presenza di altri patogeni nelle manifestazioni cliniche osservate in campo, sarà necessario isolare i ceppi sospetti e, mediante infezioni sperimentali in tacchini di 1 giorno di vita in condizioni di isolamento biologico, valutarne la patogenicità.

In questo studio AMPV di campo è stato evidenziato solo a partire dalla quarta settimana di vita, benché sperimentalmente anche i soggetti più giovani siano sensibili all'infezione. Questo dato conferma le osservazioni riportate da Cavanagh *et al.*[7], che riferiscono come possibile spiegazione l'interferenza del coronavirus della Bronchite infettiva, sia di campo che vaccinale, nell'attecchimento e replicazione nelle prime vie respiratorie di AMPV. L'osservazione di positività per AMPV associate o meno a Sindrome della Testa Gonfia, in assenza di altri patogeni, conferma che in particolari condizioni di campo AMPV svolge ruolo di patogeno primario anche nel broiler.

CONCLUSIONI

Il presente lavoro ha permesso di evidenziare come le infezioni da AMPV, prevalentemente da sottotipo B, siano diffuse sia nell'allevamento del broiler che in quello del tacchino da carne nel Nord Italia dove insiste la maggior parte della produzione avicola nazionale. Dall'unico ceppo virale appartenente al sottotipo A evidenziato non è stato possibile ottenere la sequenza nucleotidica e quindi definirne l'origine vaccinale o di campo. Ciò lascia ancora aperta la domanda relativa alla presenza in Italia di veri ceppi di campo appartenenti a tale sottotipo.
L'osservazione di focolai di TRT in tacchini vaccinati conferma la problematicità del controllo di tale infezione imputabile in gran parte all'evoluzione genomica ed antigenica dei ceppi di campo circolanti in Italia. Da ultimo la circolazione di ceppi di origine vaccinale in corso di focolai respiratori lascia ipotizzare che anche vaccini AMPV-B possano riacquisire patogenicità e persistere nell'ambiente.

Field survey of Avian Metapneumovirus in Northern Italy

SUMMARY

Avian Metapneumovirus (AMPV) is the casual agent of Turkey Rhinotracheitis (TRT), and also causes a respiratory infection in chickens, which can result in Swollen Head Syndrome (SHS). A survey of AMPV infection in 122 turkey and 48 broiler farms, in a highly densely populated area of Northern Italy, performed from January 2011 to February 2013, is reported. AMPV positive samples were detected using or RT-nested PCR or qRT-PCR both able to detect and differentiate AMPV subtype A and B. All samples but one resulted positive for AMPV subtype B confirming the high prevalence of this subtype in Italy. Only one AMPV subtype A was detected in a turkey farm located in Verona province.
The majority of AMPV detections were of field origin, circulating mainly from 9 to 12 weeks of age in turkeys and from 5 to 7 weeks of age in broilers, and associated with respiratory symptoms. Turkeys were all vaccinated at 1 day of age in the hatchery. The reasons for vaccine failure could be field virus changes in key antigenic regions that allow replication and leading to disease in well vaccinated birds. Our study revealed that VCO3 vaccine strain can be detected in turkeys up to many weeks after vaccination, with a high prevalence in the first weeks after hatch. This confirmed the pattern seen previously for subtype A and B live vaccine which has been shown to persist on farm for 4-5 weeks. Uniquely in this study, apparently vaccine derived strains were observed in older turkeys of up to 84 days of age. It is unlikely that a vaccine persists on the same farm for such a long period, so it is likely that as observed in a previous study for AMPV subtype A, a vaccine-derived virus was present in the environment being able to circulate and affect turkey farms.

KEY WORDS

Avian Metapneumovirus, broiler, meat turkey, vaccine derived strains, Northern Italy.

Bibliografia

1. Gough R.E., Jones R.C. (2008). Avian Metapneumoviruses. In Diseases of Poultry, Eds. Saif Y.M., Fadly A.M., Glisson J.R., McDougald L.R., Nolan L.K., Swayne D.E., 12th ed., 100-110. Ames, Blackwell Publishing.
2. Fabris G., D'Aprile P.N. (1990). Rinotracheite infettiva del tacchino: osservazioni sul campo ed indagini di laboratorio. Zootecnica International, 6: 36-40.
3. Catelli, E., Cecchinato M., Delogu M., De Matteo P., Ortali G., Franciosi C., De Marco M.A., Naylor, C.J. (2004). Avian Pneumovirus infection in turkey and broiler farms in Italy: a virological, molecular and serological field survey. Ital. J. Anim. Sci., 3: 287-292.
4. Cecchinato M., Catelli E., Lupini C., Ricchizzi E., Clubbe J., Battilani. M., Naylor C.J. (2010). Avian metapneumovirus (AMPV) attachment protein involvement in probable virus evolution concurrent with mass live vaccine introduction. Vet Microbiol, 146: 24-34.
5. Catelli, E. (2006) Dati epidemiologici sulle infezioni da Pneumovirus Aviare in Italia. Giornata di studio INTERVET su "Malattie respiratorie e problemi di produzione", Bologna, 7 giugno 2006, 16-23.
6. Listorti V., Lupini C., Cecchinato M., Pesente P., Rossi G., Naylor C. J., Catelli E., (2012). Rapid differentiation of vaccine and field strains of Avian Metapneumovirus of subtype B by restriction enzyme digestion of PCR products. In: Proceedings of VII International Symposium on Avian Corona and Pneumovirus and Complicating Pathogens, Rauischholzhausen, Germany, 18-21 June 2012, 313-317, Druckerei Schröder, Wetter, Germany, ISBN 978-3-9805907-9-8.
7. Cavanagh D., Mawditt K., Britton P., Naylor, C.J. (1999). Longitudinal field studies of infectious bronchitis virus and avian pneumovirus in broiler using type-specific Polymerase chain reactions. Avian Pathol, 28: 593-605.
8. Naylor C., Shaw K., Britton P., Cavanagh D. (1997). Appearance of type B avian pneumovirus in Great Britain. Avian Pathol, 26:327-338.
9. Cecchinato M., Lupini C., Munoz Pogoreltseva O.S., Listorti V., Mondin A., Drigo M., Catelli E. (2013). Development of a Real-Time RT-PCR assay for the simultaneous identification, quantitation and differentiation of avian metapneumovirus subtypes A and B. Avian Pathol, 42: 283-9.
10. Cecchinato M., Lupini C., Ricchizzi E., Falchieri M., Meini A., Jones R.C., Catelli E. (2012). Italian field survey reveals a high diffusion of avian metapneumovirus subtype B in layers and weaknesses in the vaccination strategy applied. Avian Dis, 56: 720-4.
11. Cecchinato M., Catelli E., Savage C.E., De Matteo P., Faenzi M., Naylor C.J. (2003). Evidenza di pneumovirus aviare sottotipo A in corso di un focolaio di TRT in tacchini da carne in Italia. Atti del XLII Convegno della Società Italiana di Patologia Aviare, Forlì 2-3 ottobre 2003. Large Animal Review, 9: 121-122.
12. Catelli E., Cecchinato M., Savage C.E., Jones R.C., Naylor C.J. (2006). Demonstration of loss of attenuation and extended field persistence of a live avian metapneumovirus vaccine. Vaccine, 24: 6476-6482.
13. Lupini C., Cecchinato M., Ricchizzi E., Naylor C.J., Catelli E. (2011). A turkey rhinotracheitis outbreak caused by the environmental spread of a vaccine-derived avian metapneumovirus. Avian Pathol, 40, 525-530.
14. Pattison M. (1998). TRT in the field: field situation and control. European perspective of TRT. In: Proceedings & Technical Supplements of the Roche Avian Pneumovirus Workshop, 43-49.
15. Banet-Noach C., Simanov L., Laham-Karam N., Perk S., Bacharach E. (2009). Longitudinal survey of avian metapneumoviruses in poultry in Israel: infiltration of field strains into vaccinated flocks. Avian Dis, 53: 184-189.
16. Catelli E., Lupini C., Cecchinato M., Ricchizzi E., Brown P., Naylor, C.J. (2010). Field avian metapneumovirus evolution avoiding vaccine induced immunity. Vaccine, 28: 916-921.
17. Ongor H., Karahan M., Kalin R., Bulut H., Cetinkaya B. (2010). Detection of avian metapneumovirus subtypes in turkeys using RT-PCR. Vet Rec, 166: 363-366.
18. Chacón J.L., Mizuma M., Vejarano M.P., Toquín D., Eterradossi N., Patnayak D.P., Goyal S.M., Ferreira A.J. (2011). Avian metapneumovirus subtypes circulating in Brazilian vaccinated and non-vaccinated chicken and turkey farms. Avian Dis, 55: 82-89.

Field survey of Avian Metapneumovirus in Northern Italy (Translation)

M. Cecchinato[1], M. Drigo[1], C. Lupini[2], M. Martini[1], V. Listorti[2], G. Franzo[1],M. Bonci[2], A. Laconi[2], E. Morandini[3], E. Catelli[2]

1 Dipartimento di Medicina Animale, Produzioni e Salute (MAPS), Università degli Studi di Padova
Agripolis - Viale dell'Università, 16, 35020, Legnaro (PD), Italia
2 Dipartimento di Scienze Mediche Veterinarie, Alma Mater Studiorum - Università di Bologna -
Via Tolara di Sopra, 5 - 40064 Ozzano Emilia (BO) - Italia
3 Veterinario avicolo

Summary

Avian metapneumovirus (AMPV) is the causal agent of Turkey Rhinotracheitis (TRT) and also causes respiratory infection in chickens, which can result in Swollen Head Syndrome (SHS). A survey of AMPV infection in 122 turkey and 48 broiler farms in a highly densely populated area of Northern Italy, which was performed from January 2011 to February 2013, is reported. AMPV-positive samples were detected using RT-nested PCR or qRT-PCR, which are both able to detect and differentiate AMPV subtypes A and B. All of the samples with the exception of one were found to be positive for AMPV subtype B, confirming the high prevalence of this subtype in Italy. Only one AMPV subtype A was detected in a turkey farm located in the Verona province. The majority of AMPV detections were of field origin, circulating mainly from 9 to 12 weeks of age in turkeys and from 5 to 7 weeks of age in broilers, and associated with respiratory symptoms. The turkeys were all vaccinated at one day of age in the hatchery. The reasons for vaccine failure could be field virus changes in key antigenic regions that allow replication and lead to disease in well-vaccinated birds. Our study revealed that the VCO3 vaccine strain can be detected in turkeys up to many weeks after vaccination, and a high prevalence was detected in the first weeks after hatching. This finding confirmed the pattern observed previously for subtype A and B live vaccines, which have been shown to persist in farms for four to five weeks. This study provides the unique demonstration that apparently-vaccine-derived strains were observed in older turkeys of up to 84 days of age. It is unlikely that a vaccine persists on the same farm for such a long period; thus, it is likely that, as observed in a previous study for AMPV

subtype A, a vaccine-derived virus present in the environment is able to circulate and affect turkey farms.

Introduction

Avian metapneumovirus (AMPV) is the etiological agent of Turkey Rhinotracheitis (TRT), an infection of the upper respiratory tract that causes high mobility and mortality in turkeys, mainly related to secondary bacterial infection. In breeders, it is also associated to a decrease in egg production and quality. Other avian species have been described to be susceptible to AMPV infection, and in particular, in chickens, infection is associated with Swollen Head Syndrome (SHS). Currently, four AMPV subtypes (i.e., A, B, C and D) have been identified based on a genome comparison[1]. TRT was first reported in South Africa in the late 70's and then rapidly spread all over the world with the exception of Oceania. Currently, it represents one of the most important respiratory syndromes in the turkey industry[1]. In Italy, TRT was first described in 1987[2] and then became endemic in highly densely populated poultry areas, such as Lombardia, Veneto and Emilia Romagna, and emerged as one of the major challenges for poultry health[3,4]. AMPV subtype B is the most prevalent in Northern Italy[3,4,5,10,16], whereas subtype A, after its first detection in 2003[11], has been reported only sporadically[5,12,13]. The control of TRT was attempted through the widespread use of vaccination, and the vaccine based on subtype B (strain VCO3) is currently the most widely used. In the past, to refine the vaccination strategies, the vaccine was provided only in limited geographic areas and with different administration protocols (i.e., different routes and doses). Only since 2001, vaccination has begun to be widely implemented via spray administration at hatcheries (i.e., at one day of age) and contributes to the control of clinical signs and economic losses. However, recent epidemiological investigations have reported episodes of clinical overt disease in vaccinated flocks. Nevertheless, systematic studies on AMPV distribution and characterization have not been performed since 2006[5]. To fill this lack of knowledge, a field survey has been performed to study the distribution of the two AMPV subtypes circulating in turkeys and broilers Italian flocks and their association with the animal ages. Additionally, the actual circulation of vaccine-derived strains in the field has been investigated using a previously validated assay[6].

Materials and methods

Farms

A total of 122 turkeys and 48 broilers located in Northern Italy were selected based on the results of routine diagnostic activities performed between 2011 and 2013. For each group, information regarding age, clinical signs and vaccination was collected. Considering the study design, the results are not intended to describe the actual prevalence of the pathogen but to provide an overview of the different subtypes circulating in the geographic area under investigation.

Sampling

A nasopharyngeal swab was collected from 10 birds in each group. The swabs were air dried for 30 min and stored at room temperature until processing. Each farm was sampled only once during each productive cycle.

Nested RT-PCR and qRT-PCR

A nested RT-PCR protocol, which was designed based on the G gene sequence and that is able to differentiate AMPV A from AMPV B, was used to analyse 121 pooled samples (10 samples representative of one group) to simultaneously detect and subtype AMPV. RNA extraction and reverse transcription were performed according to Cavanagh et al.[7]. The nested protocol used in the current study was previously described by Naylor at al.[8]. The amplicons were demonstrated on an ethidium bromide-stained 2% agarose gel after electrophoresis. Samples positive for AMPV subtype B were further processed through a restriction fragment length polymorphism (RLFP) protocol (see below). Because a qRT-PCR assay[9] was validated and used in the diagnostic routine in the last period of the study, 49 samples were processed using this method, which is able to simultaneously detect and differentiate AMPV subtypes A and B. These samples were not subjected to the RLFP protocol.

Restriction fragment length polymorphism analysis

To differentiate field strains from vaccine strains (strain VCO3), a RLFP protocol was implemented (Listorti et al.[6]). Briefly, the restriction enzyme *Tru 9I* (restriction site 5'T↓TAA3') was used to digest the nested RT-PCR products. Five microliters of the PCR

products were digested at 65°C for 2 h using 5 units of the enzyme in a final reaction volume of 20 μL. After digestion, the fragments were displayed on an ethidium bromide-stained 2% agarose gel after electrophoresis. Only the vaccine strains have the appropriate restriction site and PCR products are cleaved in two fragments of 300 and 40 bp. The specificity of the bands length was verified through comparison with a molecular weight marker (Φ X 174 RF DNA/Hae III Fragments, INVITRO-GEN®, Carlsbad, CA, USA).

Results

All of the groups with the exception of one in the Verona province (tested positive for AMPV subtype A) were classified as positive for AMPV subtype B using the two molecular assays previously described. The turkey groups had been vaccinated through a spray administration of AMPV subtype B (strain VCO3) at the hatcheries (at 1 day of age), whereas the broiler flocks had not been subjected to any vaccination protocol. The relationship between the animal age and AMPV positivity is reported in Figures 1 and 2 for turkey and broilers, respectively. AMPV was detected in turkeys from the first to the fourteenth week of age, and a higher frequency was encountered between the ninth and the eleventh weeks (i.e., 73/112; 59.8%). The 90 turkey samples positive for AMPV subtype B were further processed with RLFP, demonstrating that 13 strains of vaccine origin were collected from animals between 2 and 12 weeks of age. The majority of these were evidenced in the first six weeks of life. In nine cases, vaccine strains were identified in association with respiratory signs. In contrast, the detection of all field strains was associated with clinical signs and an increase in mortality. In the broiler groups, AMPV-B was detected from the fourth to the eighth week of age, with a peak in the sixth week (24/48; 50%). One strain originating from a group analysed at 5 weeks of age with respiratory signs was classified to be of vaccine origin. In some instances, AMPV was associated with SHS. This syndrome was observed in AMPV infection alone as well as in mixed infection (*E. coli, Infectious bronchitis virus,* and *Mycoplasma synoviae*).

Discussion

The present study confirmed the widespread presence of AMPV infection in our country, particularly in highly densely populated poultry areas, such as Lombardia and Veneto. Subtype B was demonstrated to be the most prevalent, as previously reported by other researchers[3,5,10], whereas subtype A was detected only once. The latter subtype, which was

recognized for the first time in Italy in 2003[11], has been reported only sporadically[5,12,13] in Toscana and Emilia Romagna. Additionally, analysis of the nucleotide sequence demonstrated its vaccine origin[12,13]. As a consequence, it is possible to hypothesize that the presence of this subtype is due only to vaccine strains. Unfortunately, it was not possible to confirm this evidence through sequencing of the strain detected in the present study. The vast majority of AMPV subtype B strains were proven to be field strains. These affect mainly turkey of 9-12 weeks of age and cause typical TRT signs. The occurrence of clinical signs in vaccinated animals could be due to several causes, such as the method and vaccination strategies, the limited duration of immunity, and immune escape due to evolution of the field virus[14,15]. Supporting this hypothesis, recent studies have demonstrated that vaccination with strain VCO3 is not fully protective against challenges in experimental infections using strains that are currently circulating in our country[16] because their evolution in antigenic regions[4] allowed the evasion of vaccine-induced immune response. The variability and in some cases the severity of the clinical signs are also imputable to managerial factors, such as farm ventilation, hygiene and animal density, which notoriously favour the establishment of secondary bacterial infection, as has been identified in our study (data not shown). Vaccine-derived AMPV-B has been identified in turkeys, mainly in the first weeks of age, shortly after vaccination at the hatchery. These data are in agreement with other reports that highlighted the persistence of AMPV subtype A and B vaccine strains until four to five weeks of age[12,15,16]. Nevertheless, differently from this study, positivity due to vaccine strains has never been reported in older animals (i.e., 84 days) or in unvaccinated broilers. It is unlikely that a vaccine strain can persist this long after vaccination. One probable explanation, which has been previously reported for AMPV-A[13], is that a vaccine actually circulates in other flocks and then infects the farm, thereby circumventing biosecurity measures. Respiratory signs and increased mortality were reported in nine farms that were found to be positive for the AMPV-B vaccine. No other pathogens (i.e., *Newcastle disease virus*, *Mycoplasma gallisepticum,* and *Mycoplasma synoviae*) were identified in these samples (data not shown). These results, coupled with those of other studies that have already reported the possibility that a live attenuate AMPV-A vaccine can undergo reversion to virulence under field conditions[12] provides support for a similar scenario for AMPV-B, as previously suggested by other researchers[,17,18].

To verify the hypothesis of a reversion to virulence and exclude the role of other contributing factors (e.g., managerial factors, other infections, etc.), it will be necessary to isolate the strains under investigation and evaluate their actual pathogenicity through experimental infections. In this study, field strains of AMPV have been identified starting from the fourth week of life, even if younger animals have also been experimentally demonstrated to be susceptible to infection. This evidence is in agreement with the results reported by Cavanagh et al.,[7] which present the interference of field and vaccine strains of IBV as a hurdle to the attachment and replication of AMPV in the upper respiratory tract. In the absence of other pathogens, the evidence of AMPV both in and not in association with SHS confirms that AMPV can be considered as a primary pathogen for broilers in some circumstances.

Conclusion

The present work has allowed the demonstration of the wide presence of AMPV, particularly subtype B, in the broiler and turkey industry in Northern Italy, where the vast majority of the Italian poultry production is located. It was not possible to obtain the sequence of the only AMPV-A identified in this study in order to define whether it is of field or vaccine origin. Thus, the issue regarding the presence of subtype A field strains in Italy remains unresolved. The evidence of clinical outbreaks in vaccinated turkey flocks highlights the challenges related to the control of this infection, which can largely be ascribed to the antigenic evolution of Italian AMPV strains. In addition, the circulation of vaccine strains during clinical outbreaks further support the hypothesis that AMPV-B vaccines can revert to virulence and spread in the environment.

References

1. Gough R.E., Jones R.C. (2008). Avian Metapneumoviruses. In Diseases of Poultry, Eds. Saif Y.M., Fadly A.M., Glisson J.R., McDougald L.R., Nolan L.K., Swayne D.E., 12th ed., 100-110. Ames, Blackwell Publishing.

2. Fabris G., D'Aprile P.N. (1990). Rinotracheite infettiva del tacchino: osservazioni sul campo ed indagini di laboratorio. Zootecnica International, 6: 36-40.

3. Catelli, E., Cecchinato M., Delogu M., De Matteo P., Ortali G., Franciosi C., De Marco M.A., Naylor, C.J. (2004). Avian Pneumovirus infection in turkey and broiler farms in Italy: a virological, molecular and serological field survey. Ital. J. Anim. Sci., 3: 287-292.

4. Cecchinato M., Catelli E., Lupini C., Ricchizzi E., Clubbe J., Battilani. M., Naylor C.J. (2010). Avian metapneumovirus (AMPV) attachment protein involvement in probable virus evolution concurrent with mass live vaccine introduction. Vet Microbiol, 146: 24-34.

5. Catelli, E. (2006) Dati epidemiologici sulle infezioni da Pneumovirus Aviare in Italia. Giornata di studio INTERVET su "Malattie respiratorie e problemi di produzione", Bologna, 7 giugno 2006, 16-23.

6. Listorti V., Lupini C., Cecchinato M., Pesente P., Rossi G., Naylor C. J., Catelli E., (2012). Rapid differentiation of vaccine and field strains of Avian Metapneumovirus of subtype B by restriction enzyme digestion of PCR products. In: Proceedings of VII International Symposium on Avian Corona and Pneumovirus and Complicating Pathogens, Rauischholzhausen, Germany, 18-21 June 2012, 313-317, Druckerei Schröder, Wetter, Germany, ISBN 978-3-9805907-9-8.

7. Cavanagh D., Mawditt K., Britton P., Naylor, C.J. (1999). Longitudinal field studies of infectious bronchitis virus and avian pneumovirus in broiler using type-specific Polymerase chain reactions. Avian Pathol, 28: 593-605.

8. Naylor C., Shaw K., Britton P., Cavanagh D. (1997). Appearance of type B avian pneumovirus in Great Britain. Avian Pathol, 26:327-338.

9. Cecchinato M., Lupini C., Munoz Pogoreltseva O.S., Listorti V., Mondin A., Drigo M., Catelli E. (2013). Development of a Real-Time RT-PCR assay for the simultaneous identification, quantitation and differentiation of avian metapneumovirus subtypes A and B. Avian Pathol, 42: 283-9.

10. Cecchinato M., Lupini C., Ricchizzi E., Falchieri M., Meini A., Jones R.C., Catelli E. (2012). Italian field survey reveals a high diffusion of avian metapneumovirus subtype B in layers and weaknesses in the vaccination strategy applied. Avian Dis, 56: 720-4.

11. Cecchinato M., Catelli E., Savage C.E., De Matteo P., Faenzi M., Naylor C.J. (2003). Evidenza di pneumovirus aviare sottotipo A in corso di un focolaio di TRT in tacchini da carne in Italia. Atti del XLII Convegno della Società Italiana di Patologia Aviare, Forlì 2-3 ottobre 2003. Large Animal Review, 9: 121-122.

12. Catelli E., Cecchinato M., Savage C.E., Jones R.C., Naylor C.J. (2006). Demonstration of loss of attenuation and extended field persistence of a live avian metapneumovirus vaccine. Vaccine, 24: 6476-6482.

13. Lupini C., Cecchinato M., Ricchizzi E., Naylor C.J., Catelli E. (2011). A turkey rhinotracheitis outbreak caused by the environmental spread of a vaccine-derived avian metapneumovirus. Avian Pathol, 40, 525-530.

14. Pattison M. (1998). TRT in the field: field situation and control. European perspective of TRT. In: Proceedings & Technical Supplements of the Roche Avian Pneumovirus Workshop, 43-49.

15. Banet-Noach C., Simanov L., Laham-Karam N., Perk S., Bacharach E. (2009). Longitudinal survey of avian metapneumoviruses in poultry in Israel: infiltration of field strains into vaccinated flocks. Avian Dis, 53:184-189.

16. Catelli E., Lupini C., Cecchinato M., Ricchizzi E., Brown P., Naylor, C.J. (2010). Field avian metapneumovirus evolution avoiding vaccine induced immunity. Vaccine, 28: 916-921.

17. Ongor H., Karahan M., Kalin R., Bulut H., Cetinkaya B. (2010). Detection of avian metapneumovirus subtypes in turkeys using RT-PCR. Vet Rec, 166: 363-366.

18. Chacón J.L., Mizuma M., Vejarano M.P., Toquín D., Eterradossi N., Patnayak D.P., Goyal S.M., Ferreira A.J. (2011). Avian metapneumovirus subtypes circulating in Brazilian vaccinated and non-vaccinated chicken and turkey farms. Avian Dis, 55: 82-89.

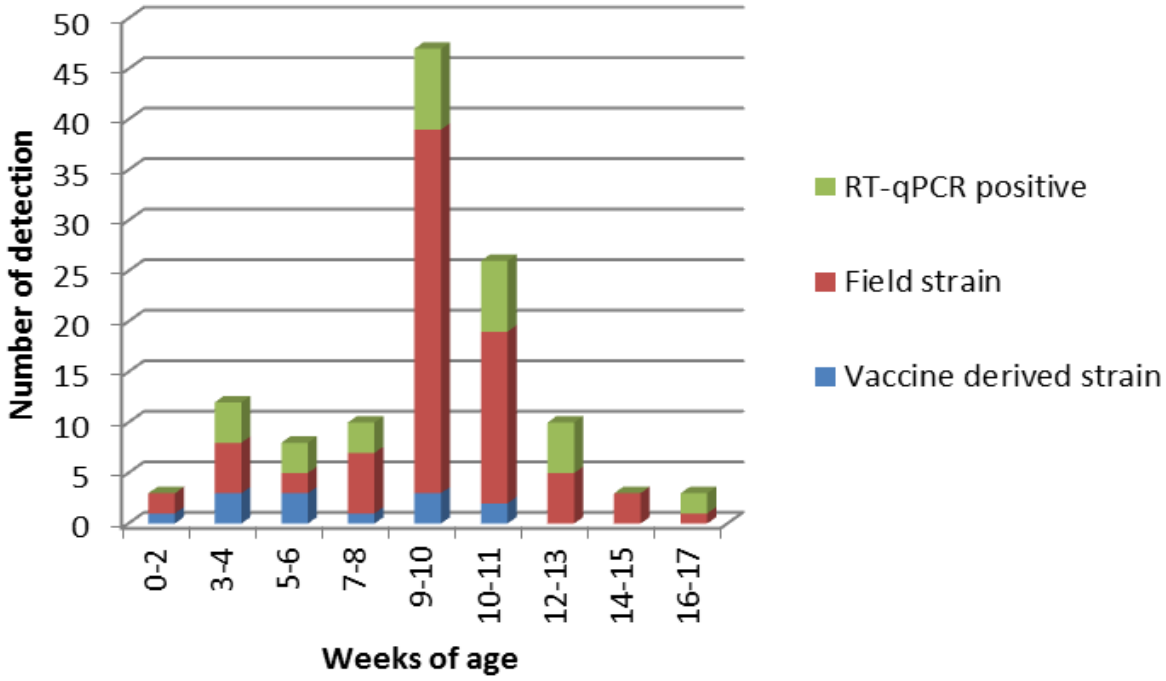

Figure 1. Distribution of AMPV-B positivity at different weeks of age in turkey farms.

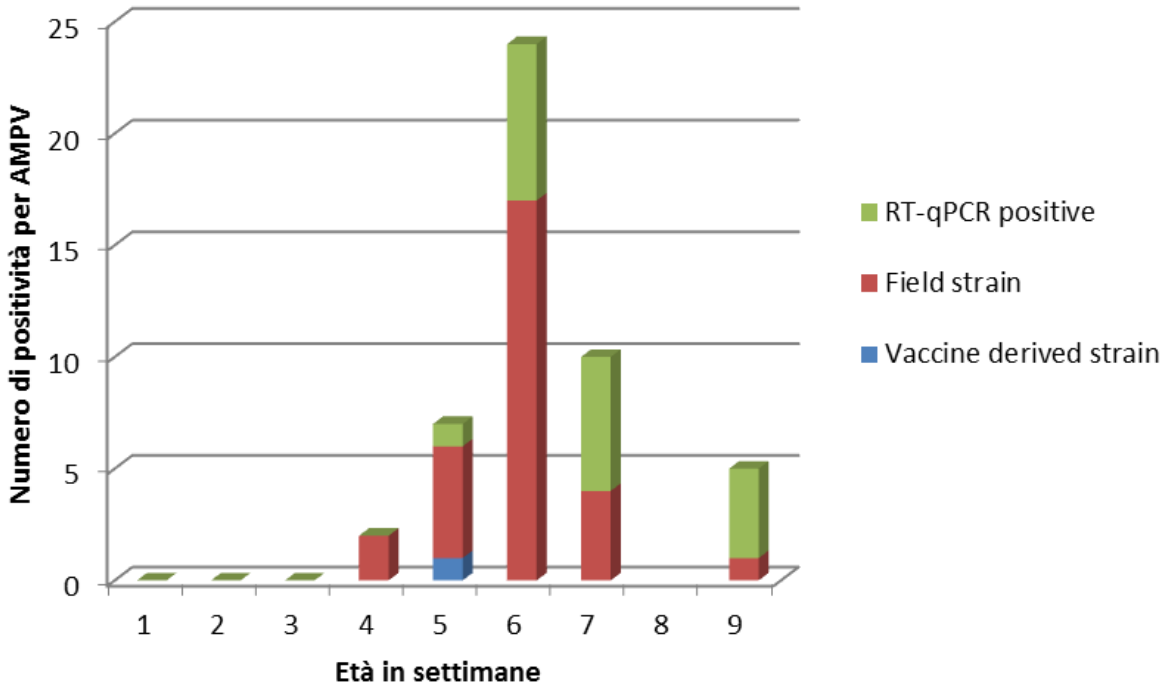

Figure 2. Distribution of AMPV-B positivity at different weeks of age in broiler farms.

AVIAN DISEASES 58:216–222, 2014

A Sensitive, Reproducible, and Economic Real-Time Reverse Transcription PCR Detecting Avian Metapneumovirus Subtypes A and B

G. Franzo,[AD] M. Drigo,[A] C. Lupini,[B] E. Catelli,[B] A. Laconi,[B] V. Listorti,[B] M. Bonci,[B] C. J. Naylor,[C] M. Martini,[A] and M. Cecchinato[A]

[A]Department of Animal Medicine, Production and Health, University of Padua, Viale dell'Università, 16, 35020, Legnaro (PD), Italy
[B]Department of Veterinary Medical Sciences, University of Bologna, Via Tolara di Sopra, 50, 40064 Ozzano Emilia (BO), Italy
[C]Department of Infection Biology, Institute of Infection and Global Health, Faculty of Health and Life Sciences, University of Liverpool, Leahurst Campus, Neston, Cheshire, CH64 7TE, United Kingdom

Received 1 October 2013; Accepted 3 January 2014; Published ahead of print 6 January 2014

SUMMARY. Use of real-time PCR is increasing in the diagnosis of infectious disease due to its sensitivity, specificity, and speed of detection. These characteristics make it particularly suited for the diagnosis of viral infections, like avian metapneumovirus (AMPV), for which effective control benefits from continuously updated knowledge of the epidemiological situation. Other real-time reverse transcription (RT)-PCRs have been published based on highly specific fluorescent dye–labeled probes, but they have high initial cost, complex validation, and a marked susceptibility to the genetic variability of their target sequence. With this in mind, we developed and validated a SYBR Green I–based quantitative RT-PCR for the detection of the two most prevalent AMPV subtypes (i.e., subtypes A and B). The assay demonstrated an analytical sensitivity comparable with that of a previously published real-time RT-PCR and the ability to detect RNA equivalent to approximately 0.5 infectious doses for both A and B subtypes. The high efficiency and linearity between viral titer and crossing point displayed for both subtypes make it suited for viral quantification. Optimization of reaction conditions and the implementation of melting curve analysis guaranteed the high specificity of the assay. The stable melting temperature difference between the two subtypes indicated the possibility of subtyping through melting temperature analysis. These characteristics make our assay a sensitive, specific, and rapid tool, enabling contemporaneous detection, quantification, and discrimination of AMPV subtype A and B.

RESUMEN. Un método de transcripción reversa y PCR en tiempo real para la detección de Metapneumovirus aviares subtipos A y B sensible, reproducible y económico.

El uso de PCR en tiempo real es cada vez mayor en el diagnóstico de enfermedad infecciosas, debido a su sensibilidad, especificidad y rapidez de detección. Estas características la hacen especialmente adecuada para el diagnóstico de las infecciones virales, como metapneumovirus aviares (aMPV), en donde se facilita el control mediante la constante actualización de la situación epidemiológica. Otras metodologías de transcripción reversa y PCR en tiempo real han sido publicadas con base en sondas altamente específicas marcadas con colorante fluorescentes, pero tienen alto costo inicial, su validación es compleja, y muestran una marcada susceptibilidad de acuerdo con la variabilidad genética de su secuencia blanco. Con esto en mente, se ha desarrollado y validado un método de RT-PCR cuantitativo basado en SYBR Green I para la detección de los dos subtipos aMPV más prevalentes (por ejemplo, los subtipos A y B). El ensayo demostró una sensibilidad analítica comparable con la de un método de RT-PCR en tiempo real publicado previamente y con capacidad de detectar ARN equivalente a aproximadamente 0.5 dosis infecciosas para ambos subtipos A y B. La alta eficiencia y linealidad entre la titulación viral y el punto de cruce para ambos subtipos muestra que el método es adecuado para la cuantificación viral. La optimización de las condiciones de reacción y la aplicación de análisis de la curva de fusión garantiza la alta especificidad del ensayo. La diferencia de temperatura de fusión estable entre los dos subtipos indicó la posibilidad de subtipificación a través del análisis de la temperatura de fusión. Estas características hacen de este ensayo una herramienta sensible, específica y rápida, lo que permite la detección simultánea, la cuantificación y la discriminación de metapneumovirus subtipos A y B.

Key words: avian metapneumovirus, diagnosis, real-time qRT-PCR, validation, subtyping, quantification

Abbreviations: AMPV = avian metapneumovirus; Cp = crossing point; CV = coefficient of variation; ELISA = enzyme-linked immunosorbent assay; GLM = General Linear Model; LoD = limit of detection; qRT-PCR = quantitative reverse transcription PCR; TCID = tissue culture infections dose; Tm = melting temperature

Avian metapneumovirus (AMPV) causes an upper respiratory tract infection in turkeys and in some other avian species, including chickens, guinea fowls (15), pheasants (11), and ducks (32). It is associated with serious economic losses in unprotected birds, particularly turkeys, especially when secondary pathogens are involved (17). Based on genome sequence differences, four subtypes of AMPV (A, B, C, and D) have been identified (5,24,28). AMPV subtypes A and B are the most widespread subtypes, and they have an almost worldwide distribution (17). Subtype C has been detected in the United States (29), France (32), Korea (23), and more recently in China (34), while subtype D has only been detected once, in France (5). Currently, the etiological diagnosis of AMPV infection is carried out by either virus isolation, detection of virus-specific antibodies using enzyme-linked immunosorbent assay (ELISA), or genome detection. Serological assays have been widely applied to this purpose; however, their results are invariably postponed due to the latency before the development of humoral immune response, and they are potentially affected by antigenic differences among the four subtypes (18). Virus isolation, although it permits detection of viral infection, is often tricky, requiring the presence of viable virus, a well-equipped laboratory, and experienced personnel. In addition, the isolated virus has to be further characterized by ELISA, indirect immunofluorescence, or molecular

[D]Corresponding author: E-mail: giovanni.franzo@studenti.unipd.it

Table 1. Sequences of primers used for real-time qRT-PCR.

Name	Sequences (5′-3′)	Position in the genome	Polarity
SH-F	TAGTTTTGATCTTCCTTGTTGC	5499–5520[A]	Sense
SH-R	GTAGTTGTGCTCAGCTCTGATA	5607–5586[A]	Antisense
SHa-R	GTAGTTGTGCTCGCTCCTGATA	5603–5582[B]	Antisense

[A]Based on the sequence of AMPV subtype B, strain VCO3/60616 (GenBank accession no. AB548428.1).
[B]Based on the sequence of AMPV subtype A, strain IT/Ty/A/259-01/03 (GenBank accession no. JF424833.1).

techniques. The degree of success depends on the strain of virus, the type and timeliness of sample collection, and the storage and handling of specimens (1). The relative insensitivity and technical requirements in virus culture and the retrospective nature of serological results have prompted researchers to develop molecular methods. Several conventional PCR assays have been developed that have demonstrated a good sensitivity and specificity (4,14,31). These tests target different regions of the AMPV genome or mRNA, with the aim of detecting and in some cases differentiating AMPV subtypes. Real-time technology has enabled PCR testing to be more rapid and more sensitive, while reducing nonspecific amplification. (8,26). Real-time reverse-transcription (RT)-PCR protocols have also been developed to detect AMPV subtypes A (16) and C (2,33) and to detect and differentiate between subtypes A and B (15,22) or any of the four AMPV subtypes (18). In all cases, different probes were used but until now no protocols have been designed using SYBR Green I chemistry.

In the present study a SYBR Green I real-time qRT-PCR coupled with melting curve analysis was developed, allowing detection, quantification, and differentiation of subtype A and B AMPVs.

MATERIALS AND METHODS

Primer design. Three oligonucleotides were used. Forward SH-F and reverse SH-R were obtained from a previous study (15), and to improve the detection of subtype A, a specific reverse primer SHa-R was designed based on a conservative region using Primer3 (21) (Table 1). Each primer was evaluated *in silico* for secondary structures and self- and hetero-dimers using OligoAnalyzer 3.1 (http://eu.idtdna.com/analyzer/applications/oligoanalyzer/). The number and position of mismatches with target sequences were also considered. The specificity of primers selected was investigated using Primer-BLAST (35).

AMPV isolates and RNA extraction. Two previously characterized viruses, IT/Ty/A/259-01/03 (subtype A) (25) and IT/Ty/B/531/10 (subtype B) (15), were cultured and titrated in Vero cell and tracheal organ culture, respectively. Titers were calculated using the Reed Muench method. A titer of $10^{5.9}$ tissue culture infections dose (TCID)50/ml and $10^{4.6}$ TCID50/ml were obtained for IT/Ty/A/259-01/03 and IT/Ty/B/531/10, respectively. RNA was extracted from 200 µl of cell culture using the High Pure RNA Isolation Kit (Roche Diagnostic, Marnes La Coquette, France) according to the manufacturer's instructions. Cell culture and RNA extracts were stored at −80 C until processing.

SYBR Green I qRT-PCR optimization. For assay optimization, 10-fold RNA dilutions for both AMPV A and B were used to perform a first evaluation of the assay's sensitivity, repeatability, and efficiency with a range of primer concentrations, annealing temperatures, RT conditions, and extension step lengths. Particular care was taken to minimize the contribution of nonspecific fluorescent amplification to total signal. qRT-PCR was performed on a LightCycler®nano system (Roche Diagnostic) using the SuperScript® III Platinum® SYBR® Green One-Step qRT-PCR Kit (Life Technologies, Carlsbad, CA). Data were analyzed using LightCycler®nano Software V1.1 (Roche Diagnostic).

Analytic validation. Analytic performances of the qRT-PCR alone and of the entire assay (i.e., RNA extraction plus qRT-PCR) were evaluated. The qRT-PCR limit of detection (LoD), efficiency, error, and coefficient of determination were evaluated using serial 10-fold viral RNA dilution ranging from $10^{5.9}$ to $10^{-2.1}$ TCID50/ml for strain IT/Ty/A/259-01/03 and from $10^{4.6}$ to $10^{-2.4}$ TCID50/ml for strain IT/Ty/B/531/10. RNA was extracted from undiluted cell culture and then serially diluted in sterile nanopure water. The lowest dilution detected as positive was then tested nine times by two operators. LoD was defined as the lowest viral amount that can be detected in at least 50% of replicates. Full process (i.e., extraction phase plus qRT-PCR) LoD, error, efficiency, and coefficient of determination were also evaluated using a similar approach. To this end a 10-fold dilution of titrated viruses was performed in uninoculated cell culture.

Repeatability. Repeatability of the assay was evaluated by selecting three RNA dilutions for each subtype, corresponding to high, medium and low dilution (i.e., one log higher than LoD). Three dilution series were produced, and each replicate of the selected dilution was tested by two operators (six replicates) during three different experiments performed at weekly intervals (18 replicates). Based on previous results, dilutions 10^{-1}, 10^{-3}, and 10^{-5} were selected for strain IT/Ty/A/259-01/03, while undiluted samples and dilution 10^{-2} and 10^{-4} were chosen for strain IT/Ty/B/531/10. The assay repeatability and the factors involved as source of variability were tested using a repeated measure GLM (7), using the following approach: the sphericity of variance-covariance matrix of the dependent variable was checked using Mauchly's test for each within subject effect. When this criterion was not met, the lower-bound epsilon was used to correct the degree of freedom. When possible *post hoc* tests based on Bonferroni correction were performed. In particular, the effect of dilution, replicate, operator, and week of experiment were considered. Additionally coefficient of variation (CV) was calculated for different experiments (replicate, operator, experiment, and dilution).

Diagnostic validation and genotype differentiation. Differentiation between AMPV type A and B based on melting temperature (Tm) curve analysis was tested using three type A and 20 type B strains (Table 2). The significance of Tm difference between subtypes and Tm variation among different runs and operators (i.e., different reaction mixes) was also evaluated using the melting curves obtained during repeatability validation. The existence of a distinguishable within-run-Tm difference between subtype A and B was evaluated on four replicates of each subtype tested in nine runs. To verify the absence of any overlap between melting peaks, these samples were also used to calculate the pairwise Tm difference by considering every combination of samples, both within and between runs, belonging to different subtypes.

Analytical specificity. Specificity was tested over a panel of several avian pathogens: *Mycoplasma synoviae*, *M. gallisepticum*, infectious bronchitis virus (serotypes QX, 793B, D274, Q1), Newcastle disease virus, classical and very virulent infectious bursal disease virus, and AMPV type C.

Statistical analysis. Statistical analysis were performed using SPSS v20.0 (30) and statistical significance was set to $P = 0.05$.

RESULTS

Assay optimization. The combination of reagents, concentrations, and cycling protocol maximizing qRT-PCR performances, while at the same time avoiding nonspecific amplification, was defined. Real-time qRT-PCR was carried out using the SuperScript III Platinum One-Step qRT-PCR Kit on a LightCycler Nano

Table 2. AMPV-A and AMPV-B strains used in the study and results of their qRT-PCR (Cp) and melting peaks (Tm) analyses.

Strain	Year of detection	Species	Results[A]			
			Cp		Tm	
AMPV-B						
IT/Ty/53/09	2009	Turkey	20.01	20.74	81.72	81.72
IT/Ck/1348-01/07[B]	2007	Chicken	18.24	19.36	81.71	81.85
IT/Ty/B/129-08/04[C]	2004	Turkey	19.99	19.69	81.83	81.81
IT/Ty/B/132-08/04[C]	2004	Turkey	22.20	22.19	81.57	81.63
IT/Ty/205-16/04[B]	2004	Turkey	17.96	15.56	81.75	81.74
IT/Ty/B/Vr240/87	1987	Turkey	28.47	28.16	81.24	81.24
IT/Ty/B/259-16/04[B]	2004	Turkey	19.04	19.34	82.42	82.39
IT/Ty/B/531/10	2010	Turkey	23.58	23.30	81.55	81.47
IT/Ty/B/532/10	2010	Turkey	19.72	17.84	82.33	82.04
IT/Ty/B/1077-02/09	2009	Turkey	14.21	14.12	81.72	81.70
IT/Ty/129-18/04[B]	2004	Turkey	21.94	21.99	80.62	80.70
IT/Ty/B/Vr240/87	1987	Turkey	25.64	25.62	81.87	81.89
IT/Ty/B/2a/01[D]	2001	Turkey	18.44	17.97	81.78	81.78
IT/Ck/B/33a/02[D]	2002	Chicken	19.61	19.03	82.16	82.13
IT/Ck/B/34a/02[D]	2002	Chicken	18.98	17.14	82.13	82.09
IT/Ty/B/531/10	2010	Turkey	26.03	26.57	81.05	81.17
Vaccine VCO3	1986	Turkey	21.45	20.85	81.79	81.82
Italy/16-91	1991	Turkey	30.76	30.81	81.73	81.68
IT/GuineaFowl/1818/12	2011	Guinea fowl	32.41	32.51	81.63	81.61
IT/Ck/B/2077/11	2011	Layer	26.03	26.13	81.39	81.35
AMPV-A						
IT/Ty/A/309/04[E]	2004	Turkey	34.78	35.45	82.85	82.83
IT/Ty/A/259-01/03[F]	2003	Turkey	16.61	16.56	83.04	83.03
IT/Ck/13574/13	2013	Chicken	35.96	35.16	82.82	83.06
AMPV C	2003	Turkey	42.16	40.29	78.92	78.92

[A]Samples tested in two replicates.
[B]Cecchinato *et al.* (13).
[C]Catelli *et al.* (12).
[D]Catelli *et al.* (9).
[E]Lupini *et al.* (25).
[F]Catelli *et al.* (10).

(Roche) with a final volume of 10 µl. Two microliters of extracted RNA was added to a standard reaction mix composed of 1X SYBR Green Reaction Mix, 0.2 µl of SuperScript III RT/Platinum Taq Mix, and 0.4 µM concentration of each primer.

The thermal profile was 50 C for 3 min (reverse transcription), 95 C for 5 min (initial PCR activation step), 45 cycles of 95 C for 15 sec (denaturation), and 60 C for 30 sec (annealing/extension). Fluorescence data were collected at the end of extension period. After the last extension, melting curve analysis was performed by progressively increasing the temperature (ramp rate = 0.1 C/sec) from 40 C to 90 C and continuously monitoring the fluorescence data.

Limit of detection. The detection range for RNA dilution was from $10^{4.6}$ to $10^{-0.4}$ TCID50/ml for IT/Ty/B/531/10 and from $10^{5.9}$ to $10^{-0.1}$ TCID50/ml for IT/Ty/A/259-01/03. Each maximum dilution of template was tested 18 times and positive results were obtained in 9 (50%) and 15 (83.33%) replicates for IT/Ty/B/531/10 and IT/Ty/A/259-01/03, respectively. LoD for cell culture dilution was $10^{0.6}$ TCID50/ml for IT/Ty/B/531/10 and $10^{0.9}$ TCID50/ml for IT/Ty/A/259-01/03 (Table 3).

Efficiency, error, and coefficient of determination. For both subtypes, qRT-PCR demonstrated a good efficiency in the range between undiluted RNA and the dilution preceding the LoD. Including the lowest detectable RNA concentration in the standard curves revealed a slight detrimental effect, particularly on efficiency and error value (Table 3). Viral dilution displayed inferior performance in terms of efficiency and linearity (Table 3; Fig. 1).

Repeatability. Our assay demonstrated overall a consistent repeatability. With the obvious exception of the dilution, the repeated measure GLM analysis revealed no significant differences among replicates, operator, and week of experiment with the following exceptions: a significant effect of operator ($P < 0.001$) for

Table 3. Summary of qRT-PCR assay analytical performances. Performances were estimated both including and excluding the higher dilution detected (LoD).

Strain	Template	Detection range (TCID50/ml)	LoD excluded			LoD included		
			Efficiency	Error	R^2	Efficiency	Error	R^2
IT/Ty/A/259-01/03	RNA	$10^{5.9}$ to $10^{-0.1}$	1.88	0.201	0.999	1.869	0.574	0.994
	Virus	$10^{5.9}$ to $10^{0.9}$	1.694	0.429	0.9944	1.705	0.42	0.9962
IT/Ty/B/531/10	RNA	$10^{4.6}$ to $10^{-0.4}$	1.919	0.103	0.9995	1.853	0.463	0.9938
	Virus	$10^{4.6}$ to $10^{0.6}$	1.878	0.241	0.9964	1.823	0.328	0.9968

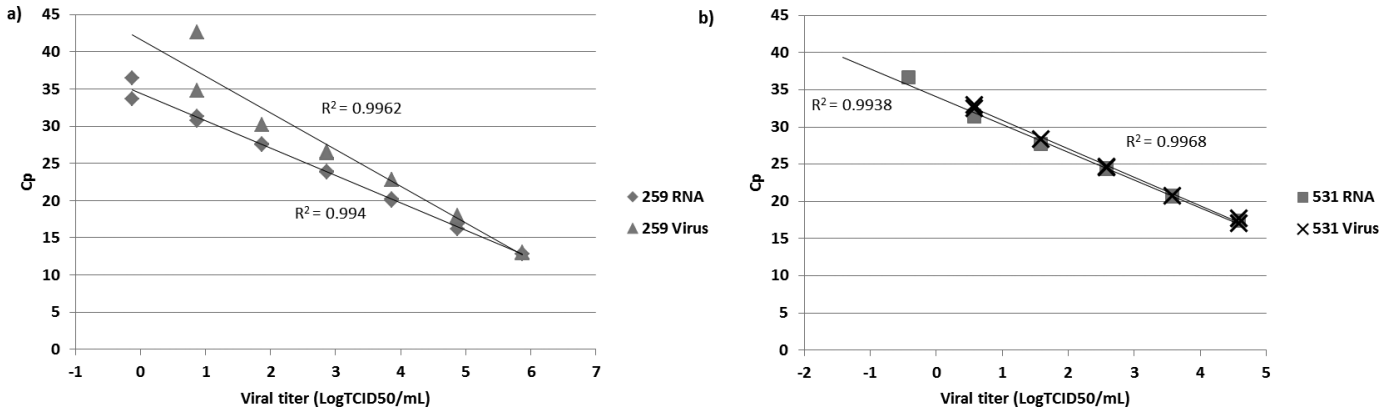

Fig. 1. Correlation between Cp and viral titer or the corresponding RNA dilution is displayed for strain IT/Ty/A/259-01/03 (a) and IT/Ty/B/531/10 (b). Viral titers and RNA dilutions are expressed in $\log_{10}$ (TCID50/ml).

IT/Ty/B/531/10 and of week of experiment ($P = 0.035$) for IT/Ty/A/259-01/03. However the CV was constantly low both within and among days of experiment (Table 4).

Diagnostic validation and genotype differentiation. Melting temperatures were constant and reproducible among different replicates, operators (i.e., different master mixes), and runs (Fig. 2). *T*-tests confirmed the presence of a significant ($P > 0.001$) difference in the Tm of the two subtypes (mean Tm = 82.72 and 81.94 for subtypes A and B, respectively). However, a significant difference in mean Tm between different runs was also detected both on total data and within subtype. To guarantee that interrun variability did not affect subtype discrimination, several replicates of both subtypes were tested in nine different runs, demonstrating a clear difference between each subtype's Tm, independent of the absolute value of Tm (Fig. 2). Additionally pairwise differences among replicates of the same and different runs demonstrated an invariably positive difference between IT/Ty/A/259-01/03 and IT/Ty/B/531/10 (Fig. 3). Similar results were obtained by comparing Tm of several strains belonging to subtype A and B (Table 2).

Specificity. No increase of fluorescence was detected for either the no template control or other avian pathogens. The only exception was AMPV C, for which a late increase in fluorescence signal was detected (crossing point [Cp] = 42.16). However the different Tm (Tm = 78.92) distinguished this subtype from AMPVs A and B.

DISCUSSION

For the first time a qRT-PCR aimed to detect, quantify, and differentiate AMPV subtype A and B was developed using the SYBR Green I chemistry.

Several multiplex real-time RT-PCR protocols were developed for the detection and the identification of AMPV different subtypes (15,18,22). However, these assays are all based on fluorophore-

Table 4. Results of the repeatability evaluation of qRT-PCR. Mean Cp, standard deviation (SD), and CV are reported, for both subtypes, according to the following factors: dilution, operator, and week of experiment.

			Week 1		Week 2		Week 3		Total	
Type	Dilution	Operator	CP (mean ± SD)	CV	CP (mean ± SD)	CV	CP (mean ± SD)	CV	CP (mean ± SD)	CV
259	−1	1	16.63 ± 0.19[A]	0.012	16.76 ± 0.04[A]	0.003	16.32 ± 0.08[A]	0.005	16.57 ± 0.22[B]	0.013
		2	16.59 ± 0.12[A]	0.007	16.38 ± 0.12[A]	0.007	16.35 ± 0.14[A]	0.009	16.44 ± 0.16[B]	0.010
		Total	16.61 ± 0.14[C]	0.009	16.57 ± 0.22[C]	0.013	16.33 ± 0.1[C]	0.006	16.51 ± 0.2[D]	0.012
	−3	1	23.82 ± 0.32[A]	0.013	23.99 ± 0.19[A]	0.008	23.65 ± 0.1[A]	0.004	23.82 ± 0.24[B]	0.010
		2	24.1 ± 0.13[A]	0.005	23.91 ± 0.22[A]	0.009	23.52 ± 0.2[A]	0.009	23.84 ± 0.31[B]	0.013
		Total	23.96 ± 0.27[C]	0.011	23.95 ± 0.19[C]	0.008	23.58 ± 0.16[C]	0.007	23.83 ± 0.27[D]	0.011
	−5	1	30.52 ± 0.48[A]	0.016	31.46 ± 0.12[A]	0.004	31.56 ± 0.44[A]	0.014	31.18 ± 0.6[B]	0.019
		2	30.61 ± 0.45[A]	0.015	30.81 ± 0.16[A]	0.005	30.84 ± 1[A]	0.032	30.75 ± 0.56[B]	0.018
		Total	30.56 ± 0.42[C]	0.014	31.13 ± 0.38[C]	0.012	31.2 ± 0.79[C]	0.025	30.96 ± 0.6[D]	0.020
531	Tq	1	17.56 ± 0.08[A]	0.004	17.91 ± 0.45[A]	0.025	17.99 ± 0.26[A]	0.014	17.82 ± 0.33[B]	0.018
		2	16.77 ± 0.04[A]	0.003	16.62 ± 0.19[A]	0.011	17.22 ± 0.12[A]	0.007	16.87 ± 0.29[B]	0.017
		Total	17.16 ± 0.43[C]	0.025	17.26 ± 0.77[C]	0.045	17.61 ± 0.46[C]	0.026	17.34 ± 0.57[D]	0.033
	−2	1	25.25 ± 0.31[A]	0.012	24.94 ± 0.16[A]	0.006	25.35 ± 0.3[A]	0.012	25.18 ± 0.3[B]	0.012
		2	23.71 ± 0.16[A]	0.007	23.35 ± 0.3[A]	0.013	23.87 ± 0.1[A]	0.004	23.64 ± 0.29[B]	0.012
		Total	24.48 ± 0.87[C]	0.036	24.14 ± 0.9[C]	0.037	24.61 ± 0.83[C]	0.034	24.41 ± 0.84[D]	0.034
	−4	1	31.3 ± 0.32[A]	0.010	32.01 ± 0.56[A]	0.018	30.93 ± 0.47[A]	0.015	31.41 ± 0.62[B]	0.020
		2	31.23 ± 0.53[A]	0.017	31.36 ± 0.98[A]	0.031	31.84 ± 0.46[A]	0.014	31.47 ± 0.66[B]	0.021
		Total	31.26 ± 0.39[C]	0.013	31.68 ± 0.8[C]	0.025	31.38 ± 0.65[C]	0.021	31.44 ± 0.62[D]	0.020

[A]Standard deviation based on three replicates.
[B]Standard deviation based on nine replicates.
[C]Standard deviation based on six replicates.
[D]Standard deviation based on 18 replicates.

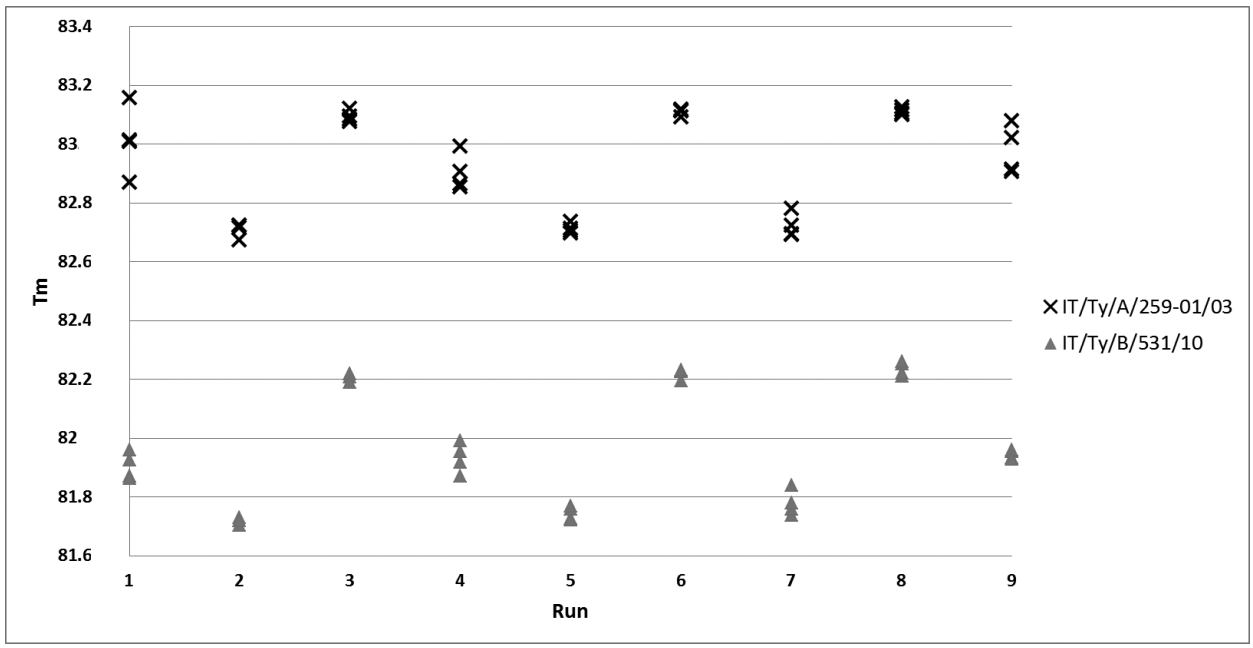

Fig. 2. Tm of four replicates of both subtype A (strain IT/Ty/A/259-01/03) and B (strain IT/Ty/B/531/10) performed in nine different runs.

labeled oligonucleotide probes (i.e., hydrolysis probes and molecular beacon probes). These chemistries, while providing great specificity and ease of genotyping (6,18,19,22), are particularly susceptible to the presence of mismatches within their target sequences (3). This problem is particularly relevant for RNA viruses (20), whose high substitution rates increase the likelihood of this phenomenon, with potential detrimental effects on analytical and diagnostic sensitivity. Having these considerations in mind and aiming to minimize diagnostic costs, we developed and validated a SYBR-Green I qRT-PCR for AMPV diagnosis.

The assay demonstrates a high analytical sensitivity for both subtypes A and B. LoD results are comparable to those of other RT-PCRs (15,16,18). In comparison with our previously published method, which was based on two primers (15), the introduction of a subtype A–specific primer in the SYBR green protocol allowed obtaining a lower LoD for this subtype.

The good efficiency and linearity, coupled with the high repeatability of our assay, provide a good basis for its use for viral quantification, making it suited for both diagnostic and research purposes. The effects of the operator or the week of experiment in the repeatability of the test are not surprising and do not nullify the assay validity. They can be easily ascribed to the different operator experience and to the time-dependent RNA degradation. In addition, the low standard deviation and the low coefficient of variation, demonstrated that the statistically significant differences observed were mainly due to the extremely high within-operator and within-run repeatability, which also overemphasized the minor differences between these factors. The loss of efficiency and the increase of the error of the standard curves, which is observed at dilutions corresponding to LoDs, suggest that Cp matching to viral titer lower than $10^{0.9}$ TCID/ml (subtype A) and $10^{0.6}$ TCID/ml (subtype B), are not suited for an accurate virus quantification, although still valuable to detect the infection. As we described in our previous study (15), RNA extraction represents a critical phase that can severely affect diagnostic performances of the assays. Our study revealed a loss of sensitivity (i.e., LoD) and higher Cp values in comparisons of viral dilutions and the corresponding RNA dilutions, demonstrating a relevant loss of RNA. Besides, the decrease in assay performances of our method suggests that RNA extraction efficiency is affected by nucleic acid concentration. Due to the broad availability of different commercial RNA extraction kits, a proper validation should be performed by each laboratory to account for

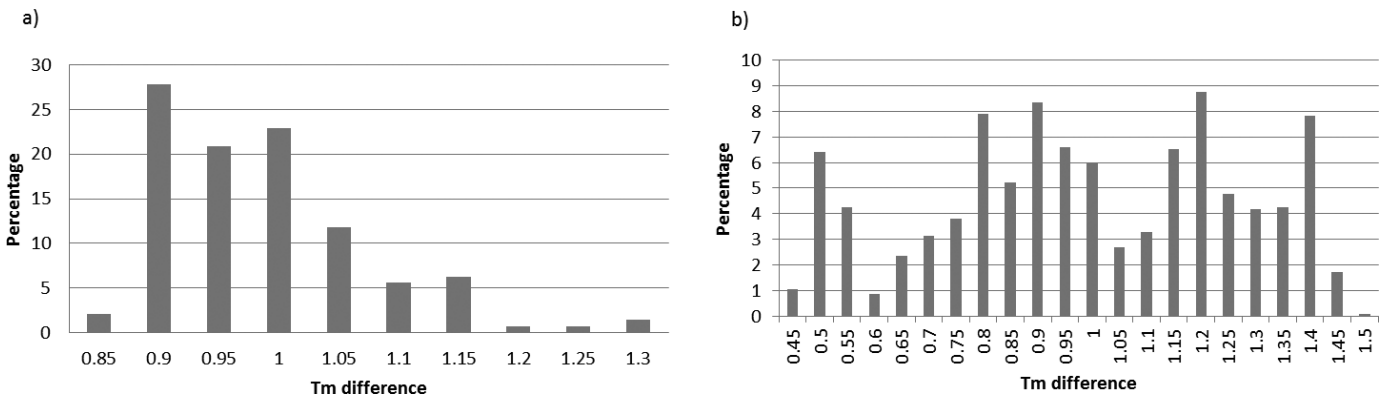

Fig. 3. Distribution of all pairwise Tm differences between replicates of subtype A (strain IT/Ty/A/259-01/03) and B (strain IT/Ty/B/531/10). The percentage of the Tm differences between subtypes is reported for both within runs (a) and among runs (b).

this factor. The optimization of primer design, reaction chemistries, and thermal protocol avoided the presence of nonspecific amplifications, ensuring a high specificity of our test in challenges with several avian pathogens. The only exception was represented by AMPV C, for which a late increase in fluorescence signal was detected. However, the implementation of Tm analysis allowed an easy recognition of this nonspecific PCR product. Although a limited specific interaction between primers and AMPV C genome cannot be excluded, the late increase of fluorescence, coupled with a high predicted number of mismatches, makes the nonspecific nature of the amplification highly probable. Tm analysis allowed discriminating between A and B subtypes due to their statistically significant Tm difference. However a significant Tm difference between runs was also observed, although some variations in Tm are not surprising and can be due to factors like dye and salt concentration and temperature ramp rate (27). This phenomenon can complicate subtyping based only on absolute Tm values. Starting from this consideration, further investigations were performed. A pairwise comparison between samples belonging to the same and/or different runs, clearly demonstrated the absence of any overlapping. In addition, a comparison of Tm over several runs revealed a consistent within-run difference between subtypes A and B, independent of the absolute values. However, given the clear within-run Tm difference, we recommend including a positive control of both subtypes to further increase the ease of subtyping.

An issue that might hinder adequate epidemiological investigations, which is often neglected in the scientific literature, is the cost of diagnostic assays. Our method addresses this issue. Our method detects and discriminates different subtypes, avoiding the costs associated with the design, purchase, and validation of updated specific probes. The implementation of our real-time qRT-PCR in the context of a high-throughput laboratory could allow substantial financial saving (approximately 25%) with respect to an assay based on two subtype-specific probes, like those used in other real-time RT-PCR protocols specific for AMPV subtypes A and B (15,16,18,22). Additionally, when an assay's use is limited, the adoption of a SYBR Green I–based method reduces the risk of losses of sensitivity due to time-dependent probe degradation.

In summary, a sensitive and reproducible real-time qRT-PCR protocol specific for AMPV subtypes A and B was developed, whose low cost represents a strong advantage when financial limitations are a major issue.

REFERENCES

1. Anonymous. Chapter 2.3.15. Turkey rhinotracheits. In: OIE terrestrial manual. Available from: http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/2.03.15_TURKEY_RHINO.pdf 2009.

2. Abbassi, H., K. Rossow, A. F. Zeigler, and K. S. Faaberg. Molecular diagnosis of avian metapneumovirus (aMPV). In: Proc. 56th Western Poultry Disease Conference, Las Vegas, NV. pp. 28–30. March 27–29, 2007.

3. Balka, G., A. Hornyák, A. Bálint, Z. Benyeda, and M. Rusvai. Development of a one-step real-time quantitative PCR assay based on primer-probe energy transfer for the detection of porcine reproductive and respiratory syndrome virus. J. Virol. Methods 158:41–45. 2009.

4. Bäyon-Auboyer, M. H., V. Jestin, D. Toquin, M. Cherbonnel, and N. Eterradossi. Comparison of F-, G- and N-based RT-PCR protocols with conventional virological procedures for the detection and typing of turkey rhinotracheitis virus. Arch. Virol. 144:1091–1109. 1999.

5. Bayon-Auboyer, M., C. Arnauld, D. Toquin, and N. Eterradossi. Nucleotide sequences of the F, L and G protein genes of two non-A/non-B avian pneumoviruses (APV) reveal a novel APV subgroup. J. Gen. Virol. 81:2723–2733. 2000.

6. Belák, S. Molecular diagnosis of viral diseases, present trends and future aspects. A view from the OIE Collaborating Centre for the Application of Polymerase Chain Reaction Methods for Diagnosis of Viral Diseases in Veterinary Medicine. Vaccine 25:5444–5452. 2007.

7. Burns, M. J., G. J. Nixon, C. A. Foy, and N. Harris. Standardisation of data from real-time quantitative PCR methods—evaluation of outliers and comparison of calibration curves. BMC Biotechnol. 5:31. 2005.

8. Bustin, S. A., V. Benes, T. Nolan, and M. W. Pfaffl. Quantitative real-time RT-PCR—a perspective. J. Mol. Endocrinol. 34:597–601. 2005.

9. Catelli, E., M. Cecchinato, M. Delogu, P. de Matteo, G. Ortali, C. Franciosi, M. A. de Marco, and C. J. Naylor. Avian pneumovirus infection in turkey and broiler farms in Italy: a virological, molecular and serological field survey. Ital. J. Anim. Sci. 3:287–292. 2004.

10. Catelli, E., M. Cecchinato, C. E. Savage, R. C. Jones, and C. J. Naylor. Demonstration of loss of attenuation and extended field persistence of a live avian metapneumovirus vaccine. Vaccine 24:6476–6482. 2006.

11. Catelli, E., M. A. De Marco, M. Delogu, C. Terregino, and V. Guberti. Serological evidence of avian pneumovirus infection in reared and free-living pheasants. Vet. Rec. 149:56–58. 2001.

12. Catelli, E., C. Lupini, M. Cecchinato, E. Ricchizzi, P. Brown, and C. J. Naylor. Field avian metapneumovirus evolution avoiding vaccine induced immunity. Vaccine 28:916–921. 2010.

13. Cecchinato, M., E. Catelli, C. Lupini, E. Ricchizzi, J. Clubbe, M. Battilani, and C. J. Naylor. Avian metapneumovirus (AMPV) attachment protein involvement in probable virus evolution concurrent with mass live vaccine introduction. Vet. Microbiol. 146:24–34. 2010.

14. Cecchinato, M., E. Catelli, C. E. Savage, R. C. Jones, and C. J. Naylor. Design, validation, and absolute sensitivity of a novel test for the molecular detection of avian pneumovirus. J. Vet. Diagn. Invest. 16:582–585. 2004.

15. Cecchinato, M., C. Lupini, O. S. Munoz Pogoreltseva, V. Listorti, A. Mondin, M. Drigo, and E. Catelli. Development of a real-time RT-PCR assay for the simultaneous identification, quantitation and differentiation of avian metapneumovirus subtypes A and B. Avian Pathol. 42:283–289. 2013.

16. Ferreira, H. L., F. R. Spilki, M. M. A. B. dos Santos, R. S. de Almeida, and C. W. Arns. Comparative evaluation of conventional RT-PCR and real-time RT-PCR (RRT-PCR) for detection of avian metapneumovirus subtype A. Ciencia Rural 39:1445–1451. 2009.

17. Gough, R. E., and R. C. Jones. Avian metapneumovirus. In: Diseases of poultry, 12th ed. Y. M. Saif, A. M. Fadly, J. R. Glisson, L. R. McDougald, N. L. Nolan, and D. E. Swayne. Wiley-Blackwell, Ames, IA. pp. 100–110. 2008.

18. Guionie, O., D. Toquin, E. Sellal, S. Bouley, F. Zwingelstein, C. Allée, S. Bougeard, S. Lemière, and N. Eterradossi. Laboratory evaluation of a quantitative real-time reverse transcription PCR assay for the detection and identification of the four subgroups of avian metapneumovirus. J. Virol. Methods 139:150–158. 2007.

19. Hoffmann, B., M. Beer, S. M. Reid, P. Mertens, C. A. L. Oura, P. A. van Rijn, M. J. Slomka, J. Banks, I. H. Brown, D. J. Alexander, and D. P. King. A review of RT-PCR technologies used in veterinary virology and disease control: sensitive and specific diagnosis of five livestock diseases notifiable to the World Organisation for Animal Health. Vet. Microbiol. 139:1–23. 2009.

20. Holmes, E. C. The evolution and emergence of RNA viruses (Oxford series in ecology and evolution). Oxford University Press, New York. 2009.

21. Koressaar, T., and M. Remm. Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289–1291. 2007.

22. Kwon, J. S., H. J. Lee, S. H. Jeong, J. Y. Park, Y. H. Hong, Y. J. Lee, H. S. Youn, D. W. Lee, S. H. Do, S. Y. Park, I. S. Choi, J. B. Lee, and C. S. Song. Isolation and characterization of avian metapneumovirus from chickens in Korea. J. Vet. Sci. 11:59–66. 2010.

23. Lee, E. H., M. S. Song, J. Y. Shin, Y. M. Lee, C. J. Kim, Y. S. Lee, H. Kim, and Y. K. Choi. Genetic characterization of avian metapneumovirus subtype C isolated from pheasants in a live bird market. Virus Res. 128:18–25. 2007.

24. Li, J., J. K. A. Cook, T. D. K. Brown, K. Shaw, and D. Cavanagh. Detection of turkey rhinotracheitis virus in turkeys using the polymerase chain reaction. Avian Pathol. 22:771–783. 1993.

25. Lupini, C., M. Cecchinato, E. Ricchizzi, C. J. Naylor, and E. Catelli. A turkey rhinotracheitis outbreak caused by the environmental spread of a vaccine-derived avian metapneumovirus. Avian Pathol. 40:525–530. 2011.

26. Mackay, I. M. Real-time PCR in microbiology: from diagnosis to characterization. Caister Academic Press, Norfolk, England. 2007.

27. Ririe, K. M., R. P. Rasmussen, and C. T. Wittwer. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Anal. Biochem. 245:154–160. 1997.

28. Seal, B. S. Matrix protein gene nucleotide and predicted amino acid sequence demonstrate that the first US avian pneumovirus isolate is distinct from European strains. Virus Res. 58:45–52. 1998.

29. Senne, D. A., R. K. Edson, J. C. Pederson, and B. Panigrahy. Avian pneumovirus update. In: Proc. American Veterinary Medical Association, 134th Annual Congress, Reno, Nevada. p. 190. July 19–24, 1997.

30. SPSS Statistics for Windows, version 20.0. IBM Corp., Armonk, NY. 2012.

31. Toquin, D., M. H. Bayon-Auboyer, N. Eterradossi, and V. Jestin. Isolation of a pneumovirus from a Muscovy duck. Vet. Rec. 145:680. 1999.

32. Toquin, D., O. Guionie, V. Jestin, F. Zwingelstein, C. Allee, and N. Eterradossi. European and American subgroup C isolates of avian metapneumovirus belong to different genetic lineages. Virus Genes 32:97–103. 2006.

33. Velayudhan, B. T., B. McComb, R. S. Bennett, V. C. Lopes, D. Shaw, D. A. Halvorson, and K. V. Nagaraja. Emergence of a virulent type C avian metapneumovirus in turkeys in Minnesota. Avian Dis. 49:520–526. 2005.

34. Wei, L., S. Zhu, X. Yan, J. Wang, C. Zhang, S. Liu, R. She, F. Hu, R. Quan, and J. Liu. Avian metapneumovirus subgroup C infection in chickens, China. Emerg. Infect. Dis. 19:1092–1094. 2013.

35. Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, and T. Madden. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13:134. 2012.

ACKNOWLEDGMENT

The work here described was supported by University of Padua grant (ex 60%, 2011).

Virus Research 194 (2014) 159–166

Contents lists available at ScienceDirect

Virus Research

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

Observation of high recombination occurrence of Porcine Reproductive and Respiratory Syndrome Virus in field condition

Giovanni Franzo[a], Mattia Cecchinato[a], Marco Martini[a], Letizia Ceglie[b], Alessandra Gigli[b], Michele Drigo[a,*]

[a] Department of Animal Medicine, Production and Health (MAPS), Viale dell'Università 16, 35020 Legnaro (PD), Italy
[b] Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università 10, 35020 Legnaro (PD), Italy

ARTICLE INFO

Article history:
Available online 21 August 2014

Keywords:
PRRSV
Italy
Recombination
ORF5
ORF7

ABSTRACT

Recombination in *Porcine Reproductive and Respiratory Syndrome Virus* (PRRSV) is a well-documented phenomenon. A high recombination frequency has been reported in experimental conditions both *in vitro* and *in vivo*, and its role in driving viral evolution has been postulated by several authors. However field evidences are rare, mainly obtained from large-scale sampling and typically represented by single sequences rather than by groups of circulating "recombinant progenies". The present work was aimed to investigate the gray area between experimental studies and large-scale epidemiological investigations. The study was performed on ORF5, ORF7 and concatenated sequences obtained in our laboratory or available in GenBank collected between 2009 and 2012 in northern Italy. Six independent recombinant strains out of 66 concatenated sequences (~9%) were found, demonstrating a high recombination frequency respect to previous field studies but comparable to *in vitro* experiments. *In silico* analysis let speculate that this new strain displayed physicochemical features diverse enough to potentially alter its immunological properties. Taken altogether, the results of our study support previous experimental evidences that depict PRRSV to be extremely prone to recombination. The limited temporal and geographical spread of recombinant strains however states in favor of a limited fitness of the recombinant progeny compared to parental strains and the marginal role of this phenomenon in PRRSV evolution.

1. Introduction

Porcine Reproductive and Respiratory Syndrome (PRRS) was first recognized quite contemporaneously in the U.S. and in Europe between the end of 1980s and the early 1990s. Since then PRRS has emerged as the most prevalent disease of swine in the world, causing remarkable economic losses (Neumann et al., 2005; Nieuwenhuis et al., 2012). The agent of the disease, *Porcine Reproductive and Respiratory Syndrome Virus* (PRRSV), classified in the order *Nidovirales*, family *Arteriviridae*, genus *Arterivirus*, is an enveloped, single-stranded positive-sense RNA virus. The viral genome is approximately 15 kb in length and contains nine open reading frames (ORFs) (Firth et al., 2011; Meulenberg, 2000). Two main genotypes, Type I (European-like) and Type II (North American-like) have been identified sharing 50–70% nucleotides and 50–80% amino acids (Shi et al., 2010a). A great and progressively increasing (Mateu et al., 2006; Pesch et al., 2005) genetic variability has been observed: mean nucleotide diversity within European and American genotypes has been estimated to be about 15% and 12.5%, respectively (Cho and Dee, 2006; Shi et al., 2010a,b). Genetic distance, calculated on ORF5, has reached a maximum of about 30% in Type I and 21% in Type II (Murtaugh et al., 2010). RNA virus evolution is assumed to result primarily from RNA polymerase infidelity. Indeed the PRRSV nucleotide substitution rate has been estimated to vary between 4.7×10^{-2} and 1.55×10^{-3} (Murtaugh et al., 2010; Yoon et al., 2012). Although the role of recombination in evolution of RNA viruses is still debated (Simon-Loriere and Holmes, 2011), several authors assert that recombination is an important mechanism of genetic diversity generation in PRRSV (Liu et al., 2011; Mengeling, 2002; Murtaugh et al., 2010), playing a potential role in conditioning virulence, antigenic escape and diagnostic failure. Several studies have demonstrated recombination in both *in vitro* (van Vugt et al., 2001; Yuan et al., 1999) and *in vivo*, in experimental (Liu et al., 2011) and field conditions (Fang et al., 2007; Forsberg et al., 2002; Li et al., 2009; Shi et al., 2010a; Stadejek

* Corresponding author. Tel.: +39 049 827 2968; fax: +39 049 827 2973.
E-mail addresses: giovanni.franzo@studenti.unipd.it (G. Franzo), mattia.cecchinato@unipd.it (M. Cecchinato), marco.martini@unipd.it (M. Martini), lceglie@izsvenezie.it (L. Ceglie), agigli@izsvenezie.it (A. Gigli), michele.drigo@unipd.it (M. Drigo).

http://dx.doi.org/10.1016/j.virusres.2014.08.005
0168-1702/

et al., 2008). In the latter case, results where typically obtained comparing sequences obtained from large-scale (*i.e.* country level) sampling. The aim of this study was to investigate recombination on a smaller scale in terms of geographic distance and time window (Forsberg et al., 2002).

2. Materials and methods

2.1. Samples

The samples used in this study were drawn from the Istituto Zooprofilattico Sperimentale delle Venezie's historical archive, a regional public veterinary laboratory collecting passive field samples brought by practitioners for diagnostic purposes. All of the 163 samples (serum and lung), coming from 52 pig farms among 12 provinces in northeastern Italy (enclosing a geographic area of about 28,000 km^2), found positive at routine RT-PCR to PRRSV between 2010 and 2012 and stored at −80 °C, were analyzed. RNA had been extracted from 200 μl of serum or 200 μl of lung homogenate using the High Pure viral RNA kit and High Pure RNA tissue kit, respectively (Roche Diagnostics, Monza, Italy). Each sample had been routinely tested using a classical two step RT-PCR targeting a genomic fragment within the ORF7 region and allowing the differentiation between the Type I and Type II strains through electrophoresis on acrylamide gels (Persia et al., 2001).

2.2. Sequencing

ORF5 and ORF7 of each sample were amplified using a one-step RT-PCR as described by Oleksiewicz et al. (1998). Briefly, ORF5 sequence was amplified using the primer ORF5F (5′ CAA TGA GGT GGG CIA CAA CC 3′) and ORF5R (5′ TAT GTI ATG CTA AAG GCT AGC AC 3′) while ORF7 was amplified using the primer pair ORF7F (5′ GCC CCT GCC CAI CAC G 3′) and ORF7R (5′ TCG CCC TAA TTG AAT AGG TGA 3′), obtaining an amplicon of 719 bp and 637 bp respectively.

Amplification and band specificity were visualized using a SYBR safe stained 2% agarose gel, after electrophoresis. Amplicons were sequenced with the same primers, in both senses, using the BygDye terminator v.3.1 Cycle Sequencing Kit (Applied Biosystem®, Monza, Italy). Sequences were obtained using ABI PRISM® 3100 Genetic Analyzer (Applied Biosystem®, Monza, Italy). Chromatograms were evaluated by FinchTV (http://www.geospiza.com) and consensus sequences were reconstructed using CromasPro (CromasPro Version 1.5). When both ORFs were available, concatenated sequences were constructed using Mesquite (Maddison and Maddison, 2011).

2.3. Sequence analysis

Sequences obtained plus those (*i.e.* 11 ORF7 and 64 ORF5) derived from Pesente et al. (2006) were aligned by Guidance (using PRANK as alignment method) (Penn et al., 2010) and score evaluated. For clarification purposes, all ORF7 sequences published by Pesente et al. (2006) were renamed with the accession number assigned to ORF5. JModelTest 2.1.2 (Darriba et al., 2012) was used to select the model of evolution according to Akaike Information Criterion (AIC). Phylogenetic trees based on ORF5 and ORF7 were reconstructed applying the Maximum Likelihood method implemented in PhyML 3.0 (Guindon et al., 2010) assuming the GTR + Γ4 + I nucleotide substitution model. Phylogenetic tree reliability was evaluated using a fast nonparametric version of the aLRT (Shimodaira–Hasegawa [SH]-aLRT), which was developed and implemented in the PhyML 3.0 (Anisimova et al., 2011).

ORF5, ORF7 and concatenated sequence alignments were tested for evidence of recombination using RDP3 (Martin et al., 2010). In order to obtain a conservative estimate, a recombination event was accepted only when detected by two or more methods implemented in the program with a *p*-value lower than 5×10^{-5}. A collection of partitions without recombination was obtained dividing the original alignment at the recombination breakpoint. Phylogenetic trees were reconstructed for each partition using RAXML (Silvestro and Michalak, 2012) and used to calculate per site log likelihoods for each alignment partition. Statistical significance of topological incongruence between segments separated by recombination breakpoints were assessed through SH, KH, ELW and AU tests implemented in CONSEL (Shimodaira and Hasegawa, 2002). A p-value < 0.05 was assumed to indicate statistical significance.

A discrete states phylogeographic reconstruction of PRRSV strains migration pattern was performed using BEAST 1.7.5 (Drummond et al., 2012) as described by Lemey et al. (2009). The 12 provinces where the samples had been collected were considered to be discrete states. An asymmetric substitution model, coupled with the Bayesian Stochastic Search Variable Selection (BSSVS), was implemented. Non-recombinant, concatenated ORF5–ORF7 sequences, for which sampling data was known, were analyzed for this purpose. Bayesian Factor (BF) was calculated in order to define well supported diffusion rates using SPREAD (Bielejec et al., 2011). Rates yielding a BF > 10 were considered adequately supported (Kass and Raftery, 1995). The same software was used to generate the KML file compatible with Google Earth displaying migration history.

2.4. In silico structural analysis

Structural consequences of recombination on GP5 were considered for a recombinant cluster that demonstrated circulation over time in a farm. Nucleotide and amino acid *p*-distance of recombinant strains from their parents were calculated using MEGA5 (Tamura et al., 2011). Hydrophobicity profile was calculated using ProtScale (Wilkins et al., 1999) assuming the Kyte & Doolittle scale. Secondary structure and transmembrane topology of GP5 were predicted using Psipred (http://bioinf.cs.ucl.ac.uk/psipred/). N-linked glycosylation sites were estimated using NetNGlyc 1.0 Server (Gupta et al., 2004). The possible role of recombination in generating strains with different immunological properties was evaluated through *in silico* prediction of T- and B-epitopes. Linear B-cell epitopes were predicted using the BepiPred 1.0 Server (Larsen et al., 2006).

For cytotoxic T lymphocytes epitopes, NetCTLpan 1.1 Server (Stranzl et al., 2010) a pan-specific major histocompatibility complex class I epitope predictor, integrating prediction of proteasomal cleavage, antigen transport efficiency and MHC-I binding affinity, was used. All swine MHC-I alleles deposited in the program database were selected to predict 8-,9-,10-,11-mer peptides. MHC-II binders were predicted using NetMHCII 2.2 server (Nielsen and Lund, 2009), searching 15-mer peptides that bound the collection of human MHC-II (loci DR and DQ). Considering that the stronger the binding, the more likely the peptide to become T-cell epitopes (Gustiananda, 2011), highly stringent cut-offs were applied. Peptides were accepted as possible epitopes when their rank score was <1% and IC50nM < 500 (including high and intermediate affinity binder of SLA-I). To limit the presence of false positive results using HLA-II based software, only strong binder (IC50nM < 50) predicted by NetMHCII were accepted, according to Díaz et al. (2009) and Gustiananda (2011).

3. Results

3.1. Recombination analysis

A total of 131 ORF5 and 111 ORF7 were obtained, including those achieved from Pesente et al. (2006). The accession numbers

Table 1
Strains detected as recombinants with RDP3. When identified parental viruses are also reported. Recombination breakpoints are defined by nucleotide position assuming the beginning of ORF5 as position 1.

Event	Major parent	Minor parent	Recombinant	Recombination breakpoint
1	440/32	736/5	2522	606
2	2287/26	162/31	1054, 84	606
3	Unknown	958	1040/39	606
4	AY39981	AY74007	AY743937	210–606
5	162/31	84	893/16, 162/32, 893/17, 177/22	123–474
6	Unknown	547/27	92	606

of sequences obtained in our study are provided in Supplementary Data 1. All of the strains were collected from a restricted geographic area of <30,000 km^2, with a maximum distance between farms of 270 km. For 66 samples both ORF5 and ORF7 were available allowing the construction of a third database, based on the concatenation (*i.e.* the joining of two character strings end-to-end) of the respective ORF5 and ORF7 sequences of each strain. All the strains belonged to the Type I subtype I, according to Stadejek's classification (Stadejek et al., 2008, 2013). Phylogenetic analysis revealed some incongruences between trees obtained from ORF5 and ORF7 (Fig. 1).

The recombination scan on concatenated sequences coupled with topology comparison revealed 6 statistically significant recombination events (Table 1 and Supplementary Data 2). Four recombinants (*i.e.* strains 2522, 1040/39, 84 and 92) displayed a single breakpoint between the end of ORF5 and the beginning of ORF7. Unfortunately, a more precise localization could not be performed because the segment spanning these ORFs was not sequenced. Besides, only for strain 2522 both parental viruses were clearly identifiable. Only donors of the ORF7 segment were clearly detected for both 1040/39 and 92 while no closely related sequences were identified in ORF5. However, the analysis of the wider ORF5 database displays a certain relatedness of 1040/39 with strains 390/45. Another recombination event between ORF5 and ORF7 was detected (strain AY743937), although only one parent could be evidently identified (AY739981). Analysis refinement using the ORF5 dataset confirmed the presence of recombination breakpoints in position 210 and 606 and identifies AY739981 and AY74007 as major and minor parents. The ORF7 tree reconstruction revealed a close relationship between AY743937 and AY739995. The last recombinant strain was identified to be the result of two recombination events. Strains 1054 and 84 were classified as recombinants between 2287/26 (ORF5) and 162/31(ORF7). At the same time strain 84 and the closely related 1054 were predicted as donors in a recombination event within ORF5 (segment 123–474): strains 893/17, 177/22, 162/32 and 893/16 were detected as recombinants between 84 (minor parent) and 162/31 (major parent). Remarkably, both parental and recombinants strains were collected from the same farm at different time periods (see Supplementary Data 1). Particularly, recombinant strains were found twice about 3 months apart while their co-circulation with 162/31. On the contrary, the minor parent (strain 84), was recorded only two years later. Predicted recombination events involved both relatively neighbor (<10 km) and distant (more than 100 km) farms. Remarkably, the same strain (*i.e.* 162/31) was implied in two different recombination events (*i.e.* events 2 and 4) (Table 1) that took place in the same geographic area. The parental strains involved in these events were sampled from the same farm (*i.e.* 162/31 and 84) or in nearby farms (*i.e.* origin farms of 162/31 and 2287/26 were about 10 km apart), supporting the short-range spread of PRRSV strains. At the same time, discrete state phylogeographic analysis evidenced the presence of 12 well-supported migration rates among provinces of northeastern Italy and an extensive PRRSV circulation over time (Supplementary Data 3).

Table 2
Nucleotide and amino acid *p*-distance in the segment internal and external to recombination breakpoints. Distance is calculated as mean distance between recombinants' group and parental viruses.

Region	Strain	84	162/31	Recombinants
(a) Nucleotide p-distance				
AA 41–158	84	0	0.191	0.031
	162/31	0.191	0	0.18
	Recombinants	0.031	0.18	0
AA 159–40	84	0	0.194	0.194
	162/31	0.194	0	0.004
	Recombinants	0.194	0.004	0
(b) Amino acid p-distance				
AA 41–158	84	0	0.184	0.056
	162/31	0.184	0	0.153
	Recombinants	0.056	0.153	0
AA 159–40	84	0	0.253	0.253
	162/31	0.253	0	0
	Recombinants	0.253	0	0

3.2. Structural analysis

GP5 of recombinant strains 893/17, 177/22, 162/32 and 893/16, demonstrating a prolonged circulation, was further analyzed to explore the presence of structural differences from parental strains which may affect viral fitness. Strains 84 and 162/31, assumed as representative of minor and major parents, respectively, were also included in the analysis. The calculation of the *p*-distance demonstrated a relevant amino acidic difference between the two parents: 18.4% in the internal region to recombination breakpoints (AA 41–158) and about 25.3% in the external. Comparable results were obtained considering nucleotides (Table 2). Hydrophobicity profile and secondary structure were also affected and recombinants displayed some different secondary structures from both major and minor parents (Supplementary Data 4). However, the transmembrane regions' prediction revealed no differences among strains even though recombination spanned part of the first ectodomain, the three transmembrane segments and the first part of the main endodomain. Some differences were observed in glycosylation pattern: the strain 162/31, 893/17, 893/16 and 162/32 were predicted to be glycosylated in position 36, 46 and 53. A similar pattern was displayed by strain 84 with the only exception that N-linked glycosylation was present in position 37 instead of 36. Certain dissimilarity was displayed by strain 177/22 which lost the glycosylation in position 46. Consequences of recombination on GP5 antigenicity were also estimated *in silico.* For computational easiness and clarification, considering the high percentage of identity purposes (*p*-distance = 0.005 and 0.004 at amino acid and nucleotide level, respectively), 893/17 was assumed as representative of all recombinant strains. Linear B-cell epitopes, although slightly different in extension, were substantially unchanged in position: a region around the first glycosylation site and two sequences within the C-terminal endodomain were identified as possible epitopes in all strains (Fig. 2). The first one, which claimed to be a neutralizing epitope (AA 37–44) (Mateu and Diaz, 2008; Plagemann, 2004), spanned the region where the initial

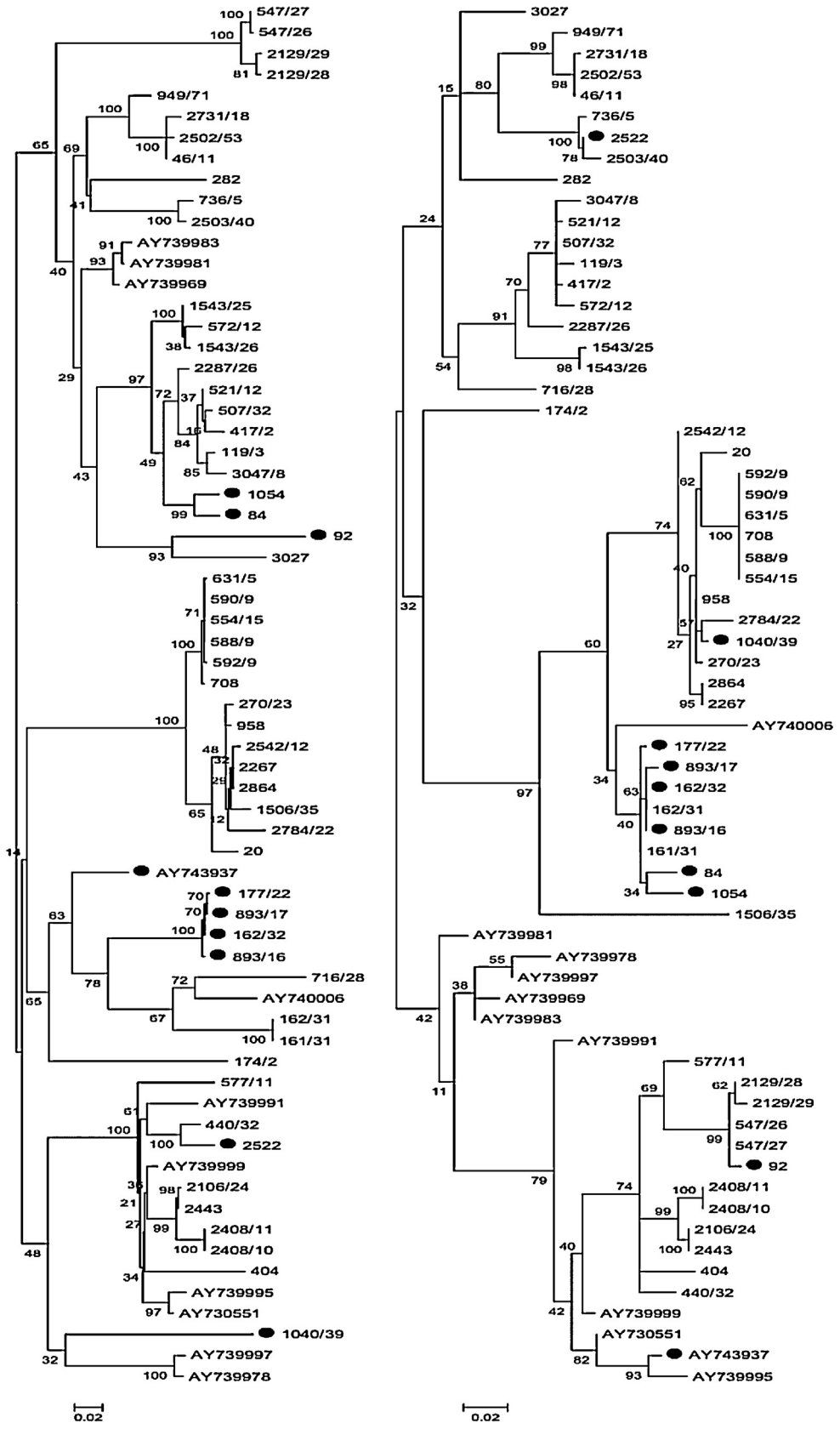

Fig. 1. Phylogenetic trees of ORF5 (on the left) and ORF7 (on the right) reconstructed using RAxML. Only strains for which both ORF5 and ORF7 were available were included in this picture. Recombinants are marked with black circles to emphasize the different topology between the two ORFs.

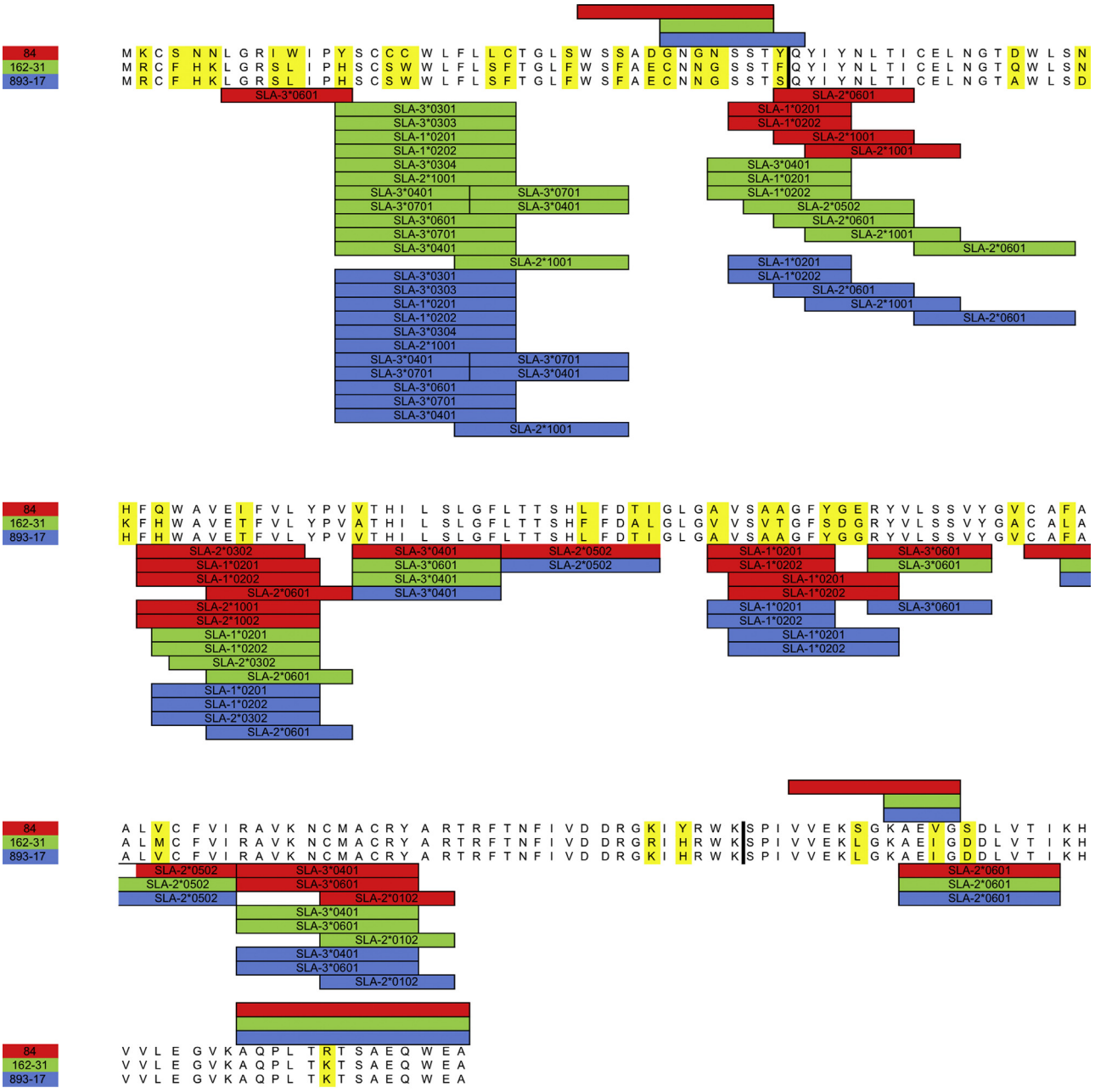

Fig. 2. SLA-I alleles are displayed under the respective peptides of GP5 predicted as ligands while linear B-cell epitopes are represented above the corresponding sequence. In both cases a color code is used to associate major (green), minor (red) parents and recombinant strain (blue) to the respective estimated epitopes. Recombination breakpoints are represented as solid black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

recombination breakpoint was estimated. One amino acidic difference between parental viruses was identifiable at amino acid 37 (G-N), with the recombinants sharing the major parent sequence. AA 41 was different in both parental and recombinant strains with the latter sharing the same residue. The analysis of T-cell epitope evidenced that, comparing major, minor and recombinant strains, approximately the same SLA-I alleles recognized the common region of GP5. Two main differences were estimated between major and minor parents in predicted ligands of SLA-I (region 7–14 and 84–93) (Fig. 2). Accordingly, the recombinant epitopic profile can be described as a combination of parental patterns. A similar picture was obtained with regard to HLA-II. Parents differed in region 61–69, 113–121 and 153–162. Recombinants resulted to be approximately a combination of parental strains. However, epitopes in position AA 41–49 and AA 61–69 were lost while one peculiar epitope (AA 103–111) was predicted (Fig. 3).

4. Discussion

RNA viruses represent both a fascinating opportunity and a challenge in the study of evolutionary processes (Holmes, 2009). Their populations usually harbor abundant genetic variability, due in large part to the combination of high mutation rate and large population size. Although recombination had been thought to be rare in these viruses, recent studies disavowed this theory, demonstrating not only that it is a frequent phenomenon in some virus families, especially retroviruses and positive-sense single stranded RNA viruses, but also that it can sometimes have a major impact on their emergence, evolution, and epidemiology (Simon-Loriere and Holmes, 2011). A similar picture can be drawn for PRRSV, a highly variable RNA virus, recently emerged as agent of devastating impact on pig production. Although the remarkable genetic distance is mainly attributable to the high substitution rate (Yoon

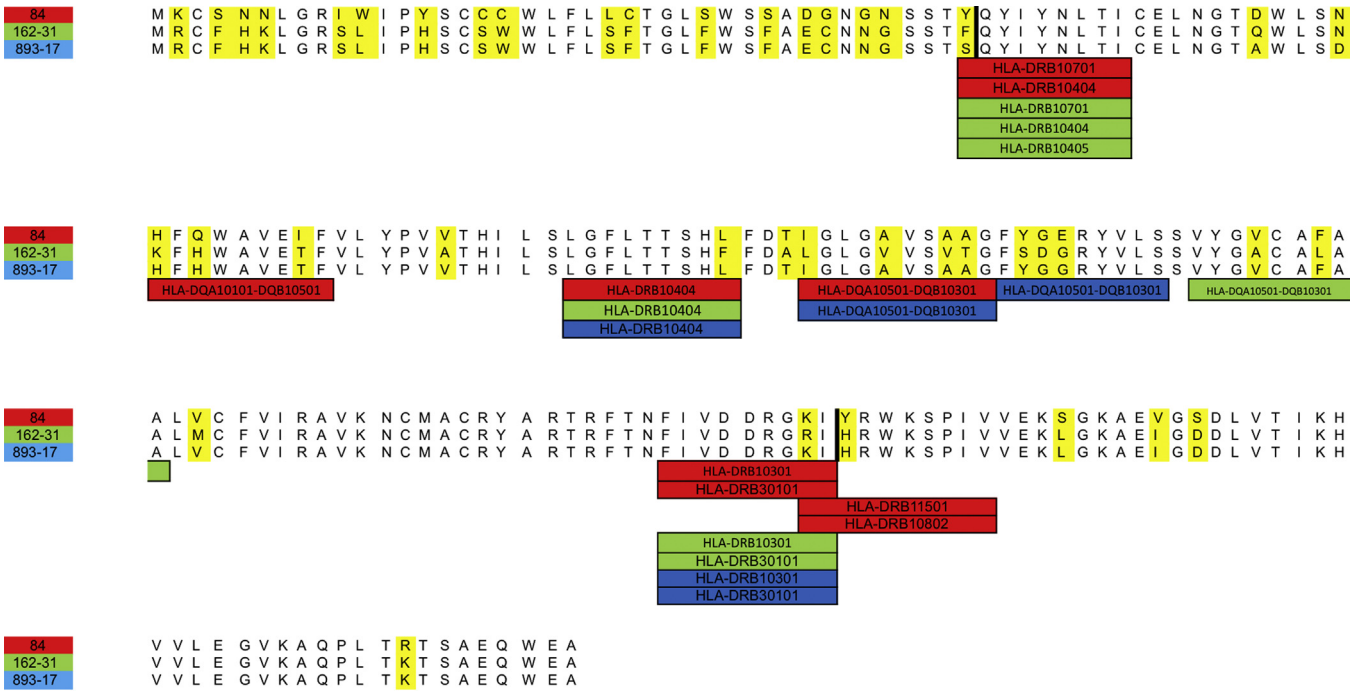

Fig. 3. HLA-II alleles are displayed under the respective peptides of GP5 predicted as ligands. A color code is used to associate major (green), minor (red) parents and recombinant strain (blue) to the respective estimated epitopes. Recombination breakpoints are represented as solid black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

et al., 2012) and to the abrupt increase in population size, consequent to the advent of high density confinement management practice in Europe and North America coupled with increasing pig transportation (Murtaugh et al., 2010), the role of recombination is nowadays increasingly studied. A high recombination frequency was demonstrated both *in vitro* (van Vugt et al., 2001; Yuan et al., 1999) and in experimentally infected pigs (Liu et al., 2011). Nevertheless, the frequency of naturally occurring recombination was believed to be much lower. In fact, most of the recombination events identified so far in the field were only represented by single sequences rather than by circulating "recombinant progenies"(Fang et al., 2007; Forsberg et al., 2002; Li et al., 2009; Murtaugh et al., 2001; Shi et al., 2010a; Stadejek et al., 2008; Yoshii et al., 2008; Yuan et al., 1999). However the vast majority of field evidences were based on sequences obtained from strains collected at different time points and from distant locations. An underestimate of the recombination rate in the field due to sampling limitations (Shi et al., 2010a) is plausible. Our study attempted to deal with this inadequacy, performing an intensive sampling based on a narrow spatial and temporal scale. Six recombination events were detected in concatenated ORF5–ORF7, affecting ORF5 and the region between ORF5 and ORF7. According to previous studies (Shi et al., 2010a), no recombination events were displayed within ORF7. Six independent recombinant strains out of 66 were found. This proportion of recombinant strains was comparable with those reported *in vitro* by Yuan and van Vugt for PRRSV (van Vugt et al., 2001; Yuan et al., 1999) and approximate the range reported for some *Coronavirus* (Lai, 1996). Several hypotheses can be proposed to explain the high frequency of recombination events observed. Firstly, it should be accounted that sequences considered, originating from a small area, are relatively similar (mean *p*-distance: ORF5 = 0.135; ORF7 = 0.098; concatenated = 0.122). van Vugt et al. (2001) demonstrated that recombination occurs preferentially between genome regions with a high percentage of similarity, which is compatible with the transcription strategy of Arterivirus (Pasternak et al., 2006). Liu et al. (2011) demonstrated that PRRSV is really prone to recombine in experimentally co-infected pigs. The occurrence of transmission of different strains between farms therefore plays a key role in the likelihood of recombination (Forsberg et al., 2002). Analysis of geographic relation between parental strains revealed that viruses originating from both distant and neighboring farms participate to this phenomenon. Accordingly, phylogeographic analysis revealed the presence of 12 well-supported migration rates among provinces of northeastern Italy. These data support the field evidence that multiple strains co-infection is relatively frequent in Italian pig farms. High farm density, intense animal movements and ineffective biosecurity measures are the most probable cause of this situation. A further hypothesis is that, collecting numerous samples on a narrow spatial and temporal scale makes it more likely to find recombinants with low fitness, whose survival and spreading among farms is unlikely, and which otherwise would not have been collected performing a less intensive sampling. The evidence that quite all recombinants were sampled only once and did not demonstrate any persistence or geographical spread supports this hypothesis. A partial exception was represented by strains 893/17, 177/22, 162/32 and 893/16. These recombinants were collected from the same farm approximately 3 months apart demonstrating a certain fitness. *In silico* analysis revealed a noteworthy amino acid difference from parental strains resulting in different hydrophobicity and secondary structure. Recombinants secondary structure was not a simple combination of the parental ones but presented some peculiar characters, probably as a consequence of interaction between distant amino acidic regions or minor AA substitutions occurred after recombination took place. Both B-cell and T-cell epitopes are constrained by sequence specificity, and mutations within epitopes can result in immune escape. Obviously, mutations within an epitope can directly affect antibody–antigen interactions or epitope-MHC and T-cell receptor interactions. Additionally, mutations outside of the epitope can inhibit antibody binding through conformational changes, or inhibit proper cleavage and processing of T-cell epitopes (Korber et al., 2006). For these reasons, the role

of recombination in generating strains with potentially different immunological features was investigated. Two linear B-cell epitopes were predicted within major endodomain and one within the major ectodomain, in substantial agreement with antigenic region identified *in vitro* by other authors (Vanhee et al., 2011). The first two are unlikely involved in viral neutralization due to their localization (Vanhee et al., 2011). On the contrary, the epitope identified within the GP5 main ectodomain is usually regarded as one of the major neutralization epitopes (Mateu and Diaz, 2008). Interestingly the beginning recombination breakpoint was predicted within this region. All the predicted antigenic regions displayed some amino acidic differences between parental viruses affecting also the sequences of recombinants strains (Fig. 2). It has been demonstrated that different PRRSV isolates differ in susceptibility to neutralization (Martínez-Lobo et al., 2011). Even if a clear correlation with protein sequence was not identified for genotype I, Kim et al. demonstrated that amino acid substitution in specific positions of ectodomain can affect neutralization in genotype II strains (Kim et al., 2013). As a consequence, it is possible to suppose that the recombination events discovered could have influenced viral susceptibility to humoral immunity through changes in epitope sequence, possibly in association with minor conformational changes that might have modified epitope accessibility or immunogenicity (Martínez-Lobo et al., 2011).

Glycosylation pattern showed some differences among parental and recombinant strains. It is well established that glycosylation plays a major role in protecting against humoral response. It has been reported that removal of N-glycosylation site renders the virus more susceptible to neutralization and elicits a significantly greater neutralizing antibodies response (Ansari et al., 2006; Darwich et al., 2010; Vu et al., 2011). However an actual correlation between glycosylation number and neutralization phenotype has not been found yet (Martínez-Lobo et al., 2011). Loss of glycosylation in position 46, predicted in strain 177/22, is usually regarded as deleterious, being it strongly required for both assembly and infectivity in LV (Dokland, 2010). However, although not frequently, field strains without this glycosylation site have been reported (Mateu et al., 2006; Stadejek et al., 2006). Besides, Balka reported a case of reversion to virulence of a PRRSV vaccine strain in which the new virulent strain was characterized by the loss of glycosylation in position 46 associated with glycosylation in position 37 (while vaccine has N-35) (Balka et al., 2008). It is also possible that apparently minor changes in glycosylation profile could significantly affect strain virulence. Also, considering potential T-cell epitopes, as expected, recombinant strains presented a combination of different parental epitopes, besides the one additional MHC-II ligand that was predicted. Despite our conservative settings, *in silico* prediction of epitopes cannot be considered an accurate tool, even if used with some success also in veterinary medicine (Díaz et al., 2009). However it should be stressed that the main purpose of this analysis was to verify if recombination between parental strains may generate an amino acidic difference relevant enough to affect their original physicochemical and antigenic properties. The results provide strong evidence that through recombination a new strain was generated which displayed characters different enough to potentially alter the development of acquired immunity or decrease the effectiveness of recall response against parental strains. Nevertheless the definition of actual B and T-cell epitopes of these strains, their effect on cross-protection or the effects of recombination on virulence are beyond the scope of this study.

5. Conclusion

This study, conducted in a restricted geographical area (28,000 km^2) in a limited time frame, evidenced a frequent occurrence of PRRSV recombination in pig farms, compatible with that reported *in vitro* and *in vivo* in experimental conditions. These results provide further confirmation that PRRSV, as other *Nidovirales* family members, is really prone to undergo recombination events, probably due to their peculiar replicative strategy (Pasternak et al., 2006; Simon-Loriere and Holmes, 2011). On the other hand, the lower frequency of recombinant strains detected on a broader scale in other studies, suggests that recombinants usually have low fitness and rarely gain an evolutionary advantage over their parents. In the present study the prolonged circulation of recombinant viruses, displaying a combination of parental physicochemical and antigenic profiles, was demonstrated in a farm. Their temporal survival and spatial spread remained limited, suggesting a marginal role of recombination in driving PRRSV evolution. Anyhow, recent evidences demonstrated the emergence of virulent Type II strains through recombination (Chen et al., 2013; Shi et al., 2013). Considering the frequency of recombination and its ability to generate strains of unexpected behavior (Li et al., 2009; Liu et al., 2011), further efforts should be deserved to study the consequences of this phenomenon on various aspects as infectivity, virulence, immunogenicity and diagnosability providing more extensive knowledge on the evolutionary driving forces of PRRSV.

Acknowledgements

The work here described was supported by the University of Padua grant, year 2013 (60A08-2170/13).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.virusres.2014.08.005.

References

Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C., Gascuel, O., 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst. Biol. 60, 685–699.

Ansari, I.H., Kwon, B., Osorio, F.A., Pattnaik, A.K., 2006. Influence of N-linked glycosylation of porcine reproductive and respiratory syndrome virus GP5 on virus infectivity, antigenicity, and ability to induce neutralizing antibodies. J. Virol. 80, 3994.

Balka, G., Hornyák, A., Bálint, A., Kiss, I., Kecskeméti, S., Bakonyi, T., Rusvai, M., 2008. Genetic diversity of porcine reproductive and respiratory syndrome virus strains circulating in Hungarian swine herds. Vet. Microbiol. 127, 128.

Bielejec, F., Rambaut, A., Suchard, M.A., Lemey, P., 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics 27, 2910–2912.

Chen, N., Yu, X., Wang, L., Wu, J., Zhou, Z., Ni, J., Li, X., Zhai, X., Tian, K., 2013. Two natural recombinant highly pathogenic porcine reproductive and respiratory syndrome viruses with different pathogenicities. Virus Genes 46, 473–478.

Cho, J.G., Dee, S.A., 2006. Porcine reproductive and respiratory syndrome virus. Theriogenology 66, 655–662.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.

Darwich, L., Díaz, I., Mateu, E., 2010. Certainties, doubts and hypotheses in porcine reproductive and respiratory syndrome virus immunobiology. Virus Res. 154, 123–132.

Díaz, I., Pujols, J., Ganges, L., Gimeno, M., Darwich, L., Domingo, M., Mateu, E., 2009. In silico prediction and ex vivo evaluation of potential T-cell epitopes in glycoproteins 4 and 5 and nucleocapsid protein of genotype-I (European) of porcine reproductive and respiratory syndrome virus. Vaccine 27, 5603.

Dokland, T., 2010. The structural biology of PRRSV. Virus Res. 154, 86.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969.

Fang, Y., Schneider, P., Zhang, W.P., Faaberg, K.S., Nelson, E.A., Rowland, R.R.R., 2007. Diversity and evolution of a newly emerged North American Type 1 porcine arterivirus: analysis of isolates collected between 1999 and 2004. Arch. Virol. 152, 1009–1017.

Firth, A.E., Zevenhoven-Dobbe, J.C., Wills, N.M., Go, Y.Y., Balasuriya, U.B., Atkins, J.F., Snijder, E.J., Posthuma, C.C., 2011. Discovery of a small arterivirus gene that overlaps the GP5 coding sequence and is important for virus production. J. Gen. Virol. 92, 1097–1106.

Forsberg, R., Storgaard, T., Nielsen, H.S., Oleksiewicz, M.B., Cordioli, P., Sala, G., Hein, J., Bøtner, A., 2002. The genetic diversity of European type PRRSV is similar to

that of the North American type but is geographically skewed within Europe. Virology 299, 38–47.
Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.
Gupta, R., Jung, E., Brunak, S., 2004. Prediction of N-glycosylation sites in human proteins, Available from: http://www.cbs.dtu.dk/services/NetNGlyc/
Gustiananda, M., 2011. Immunoinformatics analysis of H5N1 proteome for designing an epitope-derived vaccine and predicting the prevalence of pre-existing cellular-mediated immunity toward bird flu virus in Indonesian population. Immunome Res. 7, 1.
Holmes, E.C., 2009. The Evolution and Emergence of RNA Viruses. Oxford University Press, USA.
Kass, E.R., Raftery, E.A., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795.
Kim, W.-I., Kim, J.-J., Cha, S.-H., Wu, W.-H., Cooper, V., Evans, R., Choi, E.-J., Yoon, K.-J., 2013. Significance of genetic variation of PRRSV ORF5 in virus neutralization and molecular determinants corresponding to cross neutralization among PRRS viruses. Vet. Microbiol. 162, 10–22.
Korber, B., LaBute, M., Yusim, K., 2006. Immunoinformatics comes of age. PLoS Comput. Biol. 2, 0484.
Lai, M.M.C., 1996. Recombination in large RNA viruses: coronaviruses. Semin. Virol. 7, 381.
Larsen, J.E., Lund, O., Nielsen, M., 2006. Improved method for predicting linear B-cell epitopes. Immunome Res. 2, 2.
Lemey, P., Rambaut, A., Drummond, A.J., Suchard, M.A., 2009. Bayesian phylogeography finds its roots. PLoS Comput. Biol. 5, e1000520.
Li, B., Fang, L., Xu, Z., Liu, S., Gao, J., Jiang, Y., Chen, H., Xiao, S., 2009. Recombination in vaccine and circulating strains of porcine reproductive and respiratory syndrome viruses. Emerg. Infect. Dis. 15, 2032–2035.
Liu, D., Zhou, R., Zhang, J., Zhou, L., Jiang, Q., Guo, X., Ge, X., Yang, H., 2011. Recombination analyses between two strains of porcine reproductive and respiratory syndrome virus in vivo. Virus Res. 155, 473–486.
Maddison, W., Maddison, P.D.R., 2011. Mesquite: a modular system for evolutionary analysis.
Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefeuvre, P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26, 2462–2463.
Martínez-Lobo, F.J., Díez-Fuertes, F., Simarro, I., Castro, J.M., Prieto, C., 2011. Porcine reproductive and respiratory syndrome virus isolates differ in their susceptibility to neutralization. Vaccine 29, 6928–6940.
Mateu, E., Diaz, I., 2008. The challenge of PRRS immunology. Vet. J. 177, 345–351.
Mateu, E., Díaz, I., Darwich, L., Casal, J., Martín, M., Pujols, J., 2006. Evolution of ORF5 of Spanish porcine reproductive and respiratory syndrome virus strains from 1991 to 2005. Virus Res. 115, 198–206.
Mengeling, W.L., 2002. The potential role of genetic recombination in the evolution of new strains of porcine reproductive and respiratory syndrome virus (PRRSV). J. Swine Health Prod. 10, 273–275.
Meulenberg, J.J., 2000. PRRSV, the virus. Vet. Res. 31, 11–21.
Murtaugh, M.P., Stadejek, T., Abrahante, J.E., Lam, T.T., Leung, F.C., 2010. The ever-expanding diversity of porcine reproductive and respiratory syndrome virus. Virus Res. 154, 18–30.
Murtaugh, M.P., Yuan, S., Faaberg, K.S., 2001. Appearance of novel PRRSV isolates by recombination in the natural environment. Adv. Exp. Med. Biol. 494, 31–36.
Neumann, E.J., Kliebenstein, J.B., Johnson, C.D., Mabry, J.W., Bush, E.J., Seitzinger, A.H., Green, A.L., Zimmerman, J.J., 2005. Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States. J. Am. Vet. Med. Assoc. 227, 385–392.
Nielsen, M., Lund, O., 2009. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. BMC Bioinformatics 10, 296, http://dx.doi.org/10.1186/1471-2105-10-296.
Nieuwenhuis, N., Duinhof, T.F., van Nes, A., 2012. Economic analysis of outbreaks of porcine reproductive and respiratory syndrome virus in nine sow herds. Vet. Rec. 170, 225.
Oleksiewicz, M.B., Bøtner, A., Madsen, K.G., Storgaard, T., 1998. Sensitive detection and typing of porcine reproductive and respiratory syndrome virus by RT-PCR amplification of whole viral genes. Vet. Microbiol. 64, 7–22.
Pasternak, A.O., Spaan, W.J.M., Snijder, E.J., 2006. Nidovirus transcription: how to make sense. . .? J. Gen. Virol. 87, 1403–1421.
Penn, O., Privman, E., Ashkenazy, H., Landan, G., Graur, D., Pupko, T., 2010. GUIDANCE: a web server for assessing alignment confidence scores. Nucleic Acids Res. 38, W23–W28.
Persia, D., Pacciarini, M., Cordioli, P.S., 2001. Evaluation of three RT-PCR assays for the detection of porcine and respiratory syndrome virus (PRRSV) in diagnostic samples. In: Proceedings of the 10th International Symposium of Veterinary Laboratory Diagnosticians and OIE Seminar on Biotechnologies, pp. 440–441.
Pesch, S., Meyer, C., Ohlinger, V.F., 2005. New insights into the genetic diversity of European porcine reproductive and respiratory syndrome virus (PRRSV). Vet. Microbiol. 107, 31–48.
Pesente, P., Rebonato, V., Sandri, G., Giovanardi, D., Ruffoni, L.S., Torriani, S., 2006. Phylogenetic analysis of ORF5 and ORF7 sequences of porcine reproductive and respiratory syndrome virus (PRRSV) from PRRS-positive Italian farms: a showcase for PRRSV epidemiology and its consequences on farm management. Vet. Microbiol. 114, 214–224.
Plagemann, P., 2004. GP5 ectodomain epitope of porcine reproductive and respiratory syndrome virus, strain Lelystad virus. Virus Res. 102, 225–230.
Shi, M., Holmes, E.C., Brar, M.S., Leung, F.C., 2013. Recombination is associated with an outbreak of novel highly pathogenic porcine reproductive and respiratory syndrome viruses in china. J. Virol. 87, 10904–10907.
Shi, M., Lam, T.T., Hon, C.C., Hui, R.K., Faaberg, K.S., Wennblom, T., Murtaugh, M.P., Stadejek, T., Leung, F.C., 2010a. Molecular epidemiology of PRRSV: a phylogenetic perspective. Virus Res. 154, 7–17.
Shi, M., Lam, T.T., Hon, C.C., Murtaugh, M.P., Davies, P.R., Hui, R.K., Li, J., Wong, L.T., Yip, C.-W., Jiang, J.-W., Leung, F.C., 2010b. Phylogeny-based evolutionary, demographical, and geographical dissection of North American type 2 porcine reproductive and respiratory syndrome viruses. J. Virol. 84, 8700–8711.
Shimodaira, H., Hasegawa, M., 2002. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246.
Silvestro, D., Michalak, I., 2012. RaxmlGUI: a graphical front-end for RAxML. Organ. Divers. Evol. 12, 335.
Simon-Loriere, E., Holmes, E.C., 2011. Why do RNA viruses recombine? Nat. Rev. Microbiol. 9, 617–626.
Stadejek, T., Oleksiewicz, M.B., Potapchuk, D., Podgórska, K., 2006. Porcine reproductive and respiratory syndrome virus strains of exceptional diversity in eastern Europe support the definition of new genetic subtypes. J. Gen. Virol. 87, 1835–1841.
Stadejek, T., Oleksiewicz, M.B., Scherbakov, A.V., Timina, A.M., Krabbe, J.S., Chabros, K., Potapchuk, D., 2008. Definition of subtypes in the European genotype of porcine reproductive and respiratory syndrome virus: nucleocapsid characteristics and geographical distribution in Europe. Arch. Virol. 153, 1479–1488.
Stadejek, T., Stankevicius, A., Murtaugh, M.P., Oleksiewicz, M.B., 2013. Molecular evolution of PRRSV in Europe: current state of play. Vet. Microbiol. 165, 21–28.
Stranzl, T., Larsen, M.V., Lundegaard, C., Nielsen, M., 2010. NetCTLpan: pan-specific MHC class I pathway epitope predictions. Immunogenetics 62, 357.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
van Vugt, J., Storgaard, T., Oleksiewicz, M.B., Bøtner, A., 2001. High frequency RNA recombination in porcine reproductive and respiratory syndrome virus occurs preferentially between parental sequences with high similarity. J. Gen. Virol. 82, 2615–2620.
Vanhee, M., Van Breedam, W., Costers, S., Geldhof, M., Noppe, Y., Nauwynck, H., 2011. Characterization of antigenic regions in the porcine reproductive and respiratory syndrome virus by the use of peptide-specific serum antibodies. Vaccine 29, 4794–4804.
Vu, H.L.X., Kwon, B., Yoon, K., Laegreid, J., Pattnaik, W.W., Osorio, A.K.F.A., 2011. Immune evasion of porcine reproductive and respiratory syndrome virus through glycan shielding involves both glycoprotein 5 as well as glycoprotein 3. J. Virol. 85, 5555–5564.
Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D., Hochstrasser, D.F., 1999. Protein identification and analysis tools in the ExPASy server. Methods Mol. Biol. (Clifton, N.J.) 112, 531.
Yoon, S.H., Kim, H., Park, B., Kim, H., 2012. Tracing the genetic history of porcine reproductive and respiratory syndrome viruses derived from the complete ORF 5-7 sequences: a Bayesian coalescent approach. Arch. Virol. 157, 2143–2151.
Yoshii, M., Okinaga, T., Miyazaki, A., Kato, K., Ikeda, H., Tsunemitsu, H., 2008. Genetic polymorphism of the nsp2 gene in North American type-porcine reproductive and respiratory syndrome virus. Arch. Virol. 153, 1323–1334.
Yuan, S.S., Nelsen, C.J., Murtaugh, M.P., Schmitt, B.J., Faaberg, K.S., 1999. Recombination between North American strains of porcine reproductive and respiratory syndrome virus. Virus Res. 61, 87–98.

Phylodynamic analysis of Porcine reproductive and respiratory syndrome virus (PRRSV) in Italy: action of selective pressures and interactions between different clades.

Submitted to Infection, Genetics and Evolution

Franzo Giovanni[a*], Dotto Giorgia[a], Cecchinato Mattia[a], Pasotto Daniela[a], Martini Marco[a] and Drigo Michele[a].

[a]*Department of Animal Medicine, Production and Health (MAPS), Viale dell'Università 16, 35020 Legnaro (PD), Italy;*

**Corresponding author. Tel: +390498272968; fax: +390498272973;*

email address: giovanni.franzo@studenti.unipd.it

Keyword

PRRSV ORF5 phylodynamic clades interaction selection

Abstract

Porcine reproductive and respiratory syndrome (PRRS) is the most relevant and challenging infectious disease to affect swine breeding. Despite this, several aspects of the virus' evolution and virus-host interaction are still poorly understood and largely based on knowledge obtained through *in vitro* or *in vivo* experimental infections. Due to peculiar experimental conditions, our understanding is often contradictory and difficult to infer with respect to actual field conditions. Our phylodynamic study, based on ORF5 sequences of 141 samples collected in Italy from 1993 to 2012, explores different aspects of PRRSV epidemiology, evolution, and virus-host interaction. Two major clades, belonging to Type 1 subtype 1, were demonstrated to co-circulate while harboring a relevant intra- and inter-clade genetic diversity. Most Recent Common Ancestor (MRCA), evolution rates, and population dynamics were estimated using a serial coalescent-based approach, and different demographic histories were reconstructed for the two clades. Analysis of selective pressure revealed that sites subjected to diversifying selection were mainly located in the region of glycoprotein 5 (GP5) exposed to the host environment. Similarly, the vast majority of strains were highly glycosylated, confirming the proposed protective role of the

glycan shield against the humoral immune response. Overall, our study reports both interactions among the viral populations as well as between virus and host, and their relevance in shaping viral evolution: different population dynamics over time seem to reflect a competition between clades. Some evidence argues in favor of the role of immune pressure in affecting GP5 evolution, including frequent changes in the region exposed to the host immune response, and preserving glycosylation profiles that can hamper humoral immunity.

1. Introduction

Porcine reproductive and respiratory virus (PRRSV) is a member of the order *Nidovirales,* family *Arteriviridae,* genus *Arterivirus.* Two genotypes, sharing about 60% of nucleotide identity, have been identified: Type 1, reported mainly in Europe, and Type 2, comprised mainly of strains from North America and Asia (Lunney et al., 2010). However, several exceptions to this geographical localization have nowadays been reported (Shi et al., 2010a). Since its discovery PRRSV has emerged as one of the most devastating swine diseases, prompting intensive research and diagnostic activity. However, scientific papers dealing with PRRSV epidemiological trends are sporadic, limited to restricted time windows or constituted by reviews of different studies that often lack a proper leitmotif. This picture is particularly severe in Europe (and so for PRRSV Type I) due to the lack of an organized system for sequencing and sharing data among scientists. As a consequence, current knowledge about PRRSV Type 1 is based on a limited number of available sequences (Nguyen et al., 2014, Stadejek et al., 2013) compared to PRRSV Type 2 (Shi et al., 2013). Nevertheless, the progressive accumulation of studies (Balka et al., 2008, Drigo et al 2014, Forsberg et al., 2002, Indik et al., 2005, Mateu et al., 2003, Pesch et al., 2005, Stadejek et al., 2002, 2006, 2008, 2013) demonstrates the remarkable heterogeneity of European strains and of Italian ones in particular (Fosberg 2002, Drigo et al., 2014). The PRRSV genome is approximately 15 kb long and consists of 10 open reading frames (ORFs). Approximately three-quarters of the genome is occupied by ORF1a and ORF1b, encoding for 14 nonstructural proteins, while the terminal part consists of eight partially overlapping ORFs (ORF2a, ORF2b, ORF3, ORF4, ORF5, ORF5a, ORF6 and ORF7). These ORFs encode for the structural proteins constituting the envelope, matrix, and nucleocapsid. Glycoprotein 5 (GP5), encoded by ORF5, is the most widely studied PRRSV

protein due to its high genetic diversity and its role in virulence, virus-host interaction and immunity (Kimman et al., 2009).

Transmembrane region prediction using bioinformatics software revealed the presence of a major ectodomain of approximately 30 aa (aa 32-65) (preceded by the signal peptide, aa 1-31), followed by three transmembrane regions (aa 66-84, 89-104 and 112-131) and a major endodomain (aa 132-201) (Franzo et al., 2014, Kimman et al., 2009). GP5 forms a heterodimer through a disulfide link with the M protein; together they constitute the major component of the viral envelope. The complex is involved, along with other membrane proteins, in viral attachment and internalization through interaction with heparin sulfate and sialoadhesin, both of which are located on target cells (Van Breedam et al., 2010). As expected, GP5 has been reported to be one of the main targets of humoral and, to a lesser extent, cellular immunity (Darwich et al., 2010). Three epitopes involving humoral immunity have been reported, one of which is located in the major ectodomain and has been proposed to be the major neutralizing epitope (Ostrowski et al., 2002, Plagemann et al., 2002). In close proximity to this region, up to three and four glycosylation sites have been reported for Type 1 and 2, respectively. As described for other viruses, glycosylation seems to play an important role in the production of infectious virions (Wei et al., 2012; Wissink et al., 2004). It has been reported that a loss of glycosylation in GP5 could enhance immunogenicity and susceptibility to antibody-mediated neutralization (Ansari et al., 2006; Darwich et al., 2010; Faaberg et al., 2006). Recently, another highly conserved protein (i.e. GP5a) codified within ORF5 has been identified and seems to play an important role in the virus infectivity (Firth et al., 2011, Robinson et al., 2013a). All these factors could significantly affect the virus-host interaction and outcomes of PRRSV infection, although conclusive knowledge is still lacking. Although experimental results are available, experimental conditions vary and our knowledge of ORF5 evolutionary driving forces and their effect on virus-host interaction is often contradictory. Moreover, few studies have considered this phenomenon from an epidemiological prospective. Our paper, which examines the evolution based on field data, aims to provide more information on the dynamics of PRRSV epidemiology and evolution in Italy. In particular, the demographic history of different PRRSV clades has been reconstructed and the selective pressures that shape GP5 evolution have been estimated allowing to infer the evolutionary driving forces that influence the PRRSV relationship with the host as well as its epidemiological dynamics.

2. Material and methods

2.1 Dataset

The sequences of ORF5 obtained from samples collected in Italy between 1993 and 2012, of which collection data was available, were obtained from Fosberg al., (2002), Pesente et al., (2006) and Franzo et al., (2014) (Table S1). To account for their coding nature, the sequences of ORF5 were first aligned at the amino acid level; next, the nucleotide sequences were superimposed using the MAFFT algorithm implemented in TranslatorX (Abascal et al., 2010). Using RDP3, a recombination analysis was performed implementing poorly conservative parameters to increase the probability of excluding all recombinants. A sequence was considered recombinant if detected by at least one of the following methods: RDP, GENECONV, Chimaera, MaxChi, BootScan, SiScan, and 3Seq. Statistical significance was set to p-value =0.01 using the Bonferroni correction to account for multiple comparisons. Substitution saturation for each codon position was evaluated using DAMBE (Xia, 2013), with the Xia test accompanied by a visual inspection of the plot of the observed number of transitions and transversions against the corrected genetic distance.

2.2 Phylogeny, typing and subtyping

All of the Italian ORF5 sequences, plus Type 1 sequences published by Stadejek (Stadejek et al., 2013), were included in the phylogenetic analysis. To summarize, substitution models were selected based on the results of the Bayesian Information Criterion (BIC) calculated with jModelTest 2.1.1. (Darriba et al., 2012). Phylogenetic trees were reconstructed using the Maximum Likelihood (ML) method implemented in PhyML 3.0 (Guindon et al., 2010). A combination of nearest neighbor interchange (NNI) and sub-tree pruning and regrafting (SPR) were selected as the tree rearrangement strategy. To evaluate the robustness of the monophyly of the taxa subsets, a fast non-parametric version of the aLRT (Shimodaira–Hasegawa [SH]-aLRT), developed and implemented in the PhyML 3.0 (Anisimova et al., 2011), was used.

2.3 Population dynamics

A serial coalescent-based approach was applied on the non-recombinant Italian ORF5 sequences to measure the population dynamics and evolution of PRRSV in Italy. Phylogeny, Most Recent Common Ancestor (MRCA) and substitution rate (substitutions/site/year)

were jointly estimated using the Bayesian framework implemented in BEAST 1.7.5. (Drummond et al., 2012). As described in the previous paragraph, substitution models were selected using jModelTest 2.1.1 (Darriba et al., 2012). A relaxed molecular clock (Drummond et al., 2006), calibrated using the sample's collection dates, was chosen over the other molecular clock models based on Bayesian Factor (BF) calculation (Baele et al., 2012). A logBF>5 was considered as strong evidence in favor of the model. Demographic history was estimated using the time-aware Bayesian skyride (Minin et al., 2008) implemented in BEAST. Results were reported by plotting the relative genetic diversity (Ne*t) against time. "Ne" refers to the effective population size, representative of the number of individuals that contribute offspring to the descendent generation (Ho and Shapiro, 2011), while "t" refers to the generation time between infections (Frost and Volz, 2010). The relative genetic diversity is actually directly proportional to the rate of transmission if panmixis is assumed and the population considered does not have a significant subdivision. To limit possible bias due to this factor, all strains that were not part of the main Italian clades and/or displayed little or no time persistence were categorized as imported strains and were therefore excluded from further analysis.
Data was obtained performing a Markov chain Monte Carlo run of 100 million generations, during which parameters and trees were sampled every 2000 generations. Run results were accepted only if Estimate Sample Size exceeded 200 and convergence and mixing, evaluated by visually inspecting the trace plot of run results using Tracer 1.5, were adequate. Parameter estimation was summarized in terms of mean and 95% Highest Posterior Density (HPD) after the exclusion of a burn-in equal to 20% of the run length. The same analysis was executed independently on the two main Italian clades identified with the phylogenetic analysis.

2.4 Selective pressure

The action of selective pressure was evaluated on Italian sequences through the dN/dS method. A dN/dS larger, equal to, or lower than 1 is suggestive of a diversifying, neutral and purifying selection, respectively.
Pervasive and episodic site-by-site diversifying and purifying selection were evaluated on GP5 and GP5a (encoded within ORF5) (Firth et al., 2011) using the SLAC, FEL, FUBAR and MEME methods implemented in HyPhy 2.2 (Pond, 2005). A significance value was set to p-value < 0.05 for FEL and MEME analyses, while a less conservative p-value (p-value < 0.1)

was accepted for the SLAC method. FUBAR results, obtained through a Bayesian approach, were considered significant when posterior probability was greater than 0.9. The results of different methods were compared; only sites detected as significant by more than 50% of the methods were considered under non-neutral selective pressure. Biochemical properties driving substitutions at a given site were investigated using PRIME (http://www.datamonkey.org/help/citations.php), assuming the Conant-Stadler amino-acid property-specific distance measures. A Significance level was set to p-value < 0.05. The presence of lineages under selective episodic diversifying selection was assessed using the Branch-site REL implemented in HyPhy 2.2 (Pond, 2005).

2.5 Glycosylation

The Glycosylation profile was predicted *in silico* using NetNGlyc 1.0 Server (Gupta et al., 2004) and PROSITE (de Castro et al., 2006). To investigate its variation over time, an ancestral state reconstruction was performed only on Italian strains using BEAST (Drummond et al., 2012). Glycosylation history at each site, treated as a discrete character, was estimated together with ORF5 genealogy and genetic parameters using the same settings described above. The trees were summarized assuming a 20% burn-in and setting Maximum clade credibility as the target tree.

3. Results

3.1 Recombination and phylogenetic analysis

A total of 141 Italian sequences were included in this study. Nine Italian strains (162/32_29/01/10, 177/22_29/01/10, 2129/28_20/09/10, 2129/29_20/09/10, 547/26_16/03/10, 547/27_16/03/10, 893/16_22/04/10, 893/17_22/04/10 and AY743937_15/06/03) were detected as recombinants (as described in an our previous work (Franzo et al., 2014), and were excluded from further analysis. The phylogenetic analyses demonstrated that all Italian sequences belong to Type 1 subtype 1 (Stadejek et al., 2008), and were included in two clades (Clade A and Clade B) (Figure 1). The first clade consists of sequences collected over the entire time period considered, while Clade B includes sequences from 2002 (Figure 1, Figure S1 and Table S1). The mean genetic distance within Clade A was 12.4% and 8.8% for Clade B, while the mean between the clades was 15.3%. In addition, a few strains collected in 1993 (2), 2003-2004 (4) and 2010

(2) were not included among the main clades identified. Nevertheless, they were limited in number and apparently dead-ends. Earlier sequences AY035926_19/01/93 and AY035927_27/02/93 were closely related (i.e. percentage of identity 98%) to the Lelystad virus.

3.2 Population dynamics and rates of evolutionary changes

The serial coalescent approach based on Bayesian inference allowed estimates of the PRRSV demographic history in Italy. Skyride plots depicting an estimation of demographic history from MCRA to 2012 displayed a variable trend. An increase in incidence was reconstructed from MCRA, i.e 1979.6 (95%HPD 1973.9 - 1984.6) until approximately 1995, after which the transmission rate remained substantially constant until about 2004, when a progressive decline was estimated (Figure 2a). The mean rate of nucleotide substitution, estimated under relaxed molecular clock assumption, was $6{,}12x10^{-3}$(95%HPD $4{,}71x10^{-3}$ - $7{,}80x10^{-3}$) substitution/site/year. Independent population history reconstruction revealed a different scenario for the two clades (Figure 2b). The MRCA of Clade A was predicted in 1973.9 (95%HPD 1964.4 - 1982.6) and the population displayed a rapid increase until the mid 1990's, when relative genetic diversity started to decline until about 2003. A second decline phase began after 2005-06 and continued until the end of the sampling period. Clade B originated approximately in 1996.6 (95%HPD 1993 - 1999.6) and, conversely to Clade A, grew slowly until about 2004 when Ne became substantially constant. A second limited increase was estimated in the last couple years, when the relative genetic diversity of Clade B exceeded that of Clade A. The mean nucleotide substitution rates were $5.08x10^{-3}$ (95%HPD $3.35x10^{-3}$ - $6.94x10^{-3}$) and $7.43x10^{-3}$ (95%HPD $5.29x10^{-3}$ - $9.84x10^{-3}$) for Clade A and Clade B, respectively.

3.3 Selective pressure

Global ω (i.e. ratio between non-synonymous and synonymous substitution rates) of GP5 was 0.307 (95%CI = 0,287 - 0,328), which is consistent with an evolutionary process shaped mainly by purifying selection.

Site-by-site selective pressure analysis was executed. In ORF5 the majority of sites were under negative selection: 15 sites were predicted under diversifying selection (Table 1) while 117 were under purifying selective pressure (Table S2). Both significant and non-significant sites whose dN/dS was greater than one were located in the region

corresponding to the signal peptide, the main and the second ectodomain (Figure 3a). Two sites within the major endodomain (aa 173 and 199) displayed traces, although not statistically significant, of positive selection. On the other hand, part of the main ectodomain (aa 43-55), transmembrane regions and all of the other major endodomains were under negative selection. Similarly, PRIME analysis revealed that aa, whose biochemical properties were significantly affected by selective pressure, were located within (or relatively close to) the signal peptide, major ectodomain and second ectodomain (Table S3). In short, substitutions that were radical with respect to the considered properties were accelerated relative to substitutions that were conservative in 5 sites according to volume, 3 sites by Iso-electric point and hydropathy, 2 sites by polarity and 1 site for chemical composition (Table S3). Site-branch REL revealed no lineage subjected to episodic diversifying selection at 0.05 significance level. GP5a was mainly under neutral or purifying selection with only aa 13 being detected as positively selected using the consensus approach previously described (Table S4 and Figure 3b).

3.4 Comparison with the vaccine strain

No particular differences were found in terms of the amino acidic composition between Clade A and Clade B compared to vaccine strains in the ectodomain regions (data not shown). Amino acid distances between groups were calculated between the two clades and the vaccines commercially available (i.e. Porsilis®, Hypra® and Amervac®) by using a sliding windows approach. Window of 20 aa was moved one aa at a time and p-distance was calculated (Figure S2).

Clade B showed a lower percentage of aa identity with both vaccine strains in the peptide signal region. This difference, as expected, was also characteristic of the ectodomain and the transmembrane region of the GP5a protein (Figure S3).

3.5 Glycosylation

Glycosylation site analysis revealed that most of the strains presented glycosylation in positions 37, 46 and 53. However, several exceptions were noted. Out of the 132 sequences, 3 (2.27%) were glycosylated in position 35, 7 (5.30%) in position 36, and 114 (86.6%) in position 37. 114 GP5 were predicted to be glycosylated in position 46 (86.4%). Out of the 18 sequences where glycosylation 46 was absent, 3 were glycosylated only in position 53 while the others were glycosylated in positions 37 and 53. Only two strains

lacked glycosylation in position 53 (i.e. AY739962_15/03/04 and AY739958_15/06/03), being glycosylated only in aa 37/46 and 35/46, respectively. However, 11 sequences, predicted as glycosylated using Prosite, displayed a low score for the last glycosylation site when evaluated using NetNGlyc 1.0 Server (potential < 0.5). Similar glycosylation profiles were predicted in the two clades, even if only Clade A included strains with glycosylation in positions 35 or 36, and only one glycosylation site (Table 2). Bayesian ancestral state reconstruction revealed that sequences glycosylated on asparagine 35 included only terminal branches scattered throughout the tree. A similar picture could be drawn for site 53. On the other hand, sequences glycosylated in position 36 formed a cluster including sequences from 15/09/2002 (AY740007_15/09/02) to 02/05/2011 (716/28_02/05/11) (Data not shown). Remarkably, sequences AY740008_15/07/02 and 161/31_29/01/10, part of this cluster, presented no glycosylation in this area, or glycosylation in position 37. Loss of glycosylation on aa 46 was mainly represented by a single sequence or pair of sequences randomly distributed throughout the phylogenetic tree. However, a four-sequence clade, including strains collected between 15/04/2002 and 15/02/2004, was also identified (Data not shown).

4. Discussion

Porcine reproductive and respiratory syndrome has nowadays become one the most relevant infectious diseases of swine breeding. Despite all the efforts of the last two decades, many aspects of its pathogenesis and evolution remain largely unknown. A phylodynamic approach has rarely been applied to PRRSV (Nguyen et al., 2014, Shi et al., 2010b, Yoon et al., 2012). In particular, this approach has never been performed for PRRSV on a small geographic area, although studies on other pathogens have demonstrated the effectiveness of this approach in understanding the demographic dynamics of the viral population and the forces that are involved (Alfonso-Morales et al., 2013, Murcia et al., 2011, Pybus et al., 2003). Traditional epidemiological data is sparse in Italy and generally obtained from serological studies, and is therefore biased by the widespread use of vaccination (De Mateo-Aznar, 2008). Nevertheless, the prevalence of farms infected by PRRSV has been estimated at around 90%, even though frequently in absence of overt clinical signs (Candotti et al., 2006). The economical and practical constraints have made impossible to perform extensive longitudinal studies to understand the actual dynamics of this infection using traditional epidemiologic tools. The Italian situation represents an

interesting model because PRRSV strains have been collected and sequenced since the early ’90's and intensive swine breeding is mainly concentrated in the restricted area of Northern Italy. Based on these features, it was possible to perform a coalescent-based phylodynamic analysis, using sequences serially collected from a substantially homogenous population. Using this approach it was possible to avoid the limitations of previous studies. Although the Italian strains belong to Type 1 subtype 1, our study demonstrated that they harbored a significant genetic diversity. The phylogenetic analysis identified two clades comprising 95% of the strains sampled, potentially originating from different events of introduction at different moments in time. The presence of several “out of clade” strains without relevant progeny suggests that even if multiple introductions of strains from different countries manifested, only a few strains gained predominance and gave rise to the two major clades co-circulating until recent years. The global population dynamic reconstruction demonstrated that, after the first introduction of PRRSV in Italy, there was a rapid growth in the viral population that nearly decupled in 10 years (Figure 2a), followed by a period (i.e. between 1997 and 2008) characterized by a substantially constant population size. The progressive acquisition of population immunity through natural infection or vaccination (attenuated live vaccine became available in Italy in 2000) likely played a role in limiting transmission rate. Contrary to what was reported for other pathogens (Alfonso-Morales et al., 2013, Murcia et al., 2011), limited contraction of incidence after the vaccine introduction, if present, attests to the inadequate efficacy of this control measure in limiting infection and the spread of PRRSV (Kimman et al., 2009, Charerntantanakul, 2012). The last time window displayed a trend directed to a decrease in relative genetic diversity, with a relevant step in 2010. Considering the absence of any innovation with regard to vaccination, this tendency is probably due to the crescent implementation of biosecurity measures and improved management. Other authors described a similar trend, shifted in the past of about 10 years (Yoon et al., 2012). However, in that paper, sequences obtained from a larger scale were selected, including both Type 1 and 2. As a consequence, different population dynamics could reflect sequence selection, different evolution of the two genotypes or different characters of swine industry. The peculiarity of Italian swine breeding, where pigs are raised until about 160 kg, creates a dense population of animals whose management is usually problematic and characterized by a declining PRRSV-specific immunity. This background determines conditions that can facilitate persistent circulation of PRRSV. The analysis of the viral population dynamic for

each clade revealed an interesting interaction between the two clades. In particular, a major decline in the relative genetic diversity of Clade A reflects the introduction and sharp expansion of Clade B, suggesting a competition between the two clades. The actual causes of this phenomenon are unknown. Nevertheless, it is likely that varied susceptibility to host response or virulence plays a major role in modeling viral dynamic.

The method used by PRRSV to counteract the action of immune response by establishing prolonged infection and circumventing existing acquired immunity is still a matter of discussion (Murtaugh and Genzow, 2011). However, anti-PRRSV immunity was demonstrated to be completely effective in clearing infection (Murtaugh and Genzow, 2011, Murtaugh et al., 2010). Despite that, several studies have reported the limited cross-protection between heterologous strains, both vaccine and wild type. It is acknowledged that viral heterogeneity plays a major role in determining inter-host transmission success and viral fitness within a host population. Accordingly, evasion from immunity should be one of the driving forces of PRRSV evolution. Our study demonstrates that ORF5 is mainly under purifying selection, probably for structural and functional constraints. The only sites detected under significant diversifying selection were located in the predicted major and minor ectodomains, which are the only regions exposed to the actions of antibodies. One epitope (aa 33-44) has been characterized in the major ectodomain in several studies (Vanhee et al., 2011), even if its role in virus neutralization is still under debate. Kim et al., (2013) demonstrated that aa substitution in specific regions (aa 32-34; 38-39; 57-61) can severely affect the Ab mediated viral neutralization of PRRSV Type 2, and it is well-known that changes in the amino-acidic sequence and physic-chemical composition can hamper the recognition of epitopic sites (Korber et al., 2006). According to our study seven sites within the major ectodomain have been detected under diversifying selection (Table 1 and Table S3). Remarkably, sites between aa 43-55 are under neutral or purifying selection (Figure 3a), probably because of structural constraints imposed by the presence of the binding site with protein M (Cys 50) (Dokland, 2010). Interestingly, sites at the intersection between the second transmembrane region and the second ectodomain were detected under positive selective pressure, with a tendency to radical changes in hydropathy (aa 101-104), polarity, and iso-electric point (aa106). Changes in these properties could alter the portion of GP5 exposed to the outer side of the envelope, affecting its exposure to immune system action as proposed by Pesh et al. (Pesch et al., 2005). Two other epitopes (aa 165-176 and 189-201) were demonstrated through *in vitro* studies (Vanhee et al.,

2011) within the main endodomain. Only slight and non-significant evidence of positive selection was identified in this region. Although potentially suited to induce a specific B-cell response, the localization of these sites render them inaccessible and therefore ineffective in hindering viral infectivity. On the other hand, region 117-133, reported by Diaz et al.(Díaz et al., 2009) to be the target of cell-mediated immunity, displayed no sign of diversifying pressure. Biological constraints could contrast the action of immune pressure, rendering this region relatively conserved (Pesch et al., 2005) and possibly justifying the proposed role of cell-mediated immunity in heterologous protection (Díaz et al., 2006, Mateu and Diaz, 2008). However, further studies would be necessitated to clarify the CMI role, the epitopes involved and their susceptibility to PRRSV heterogeneity. Similarly, the true role of vaccination in affecting viral evolution is still unknown. In fact, Clade B displayed a lower decrease in incidence than Clade A after the introduction of vaccination, a trend that could reflect a lower rate of protection induced by the vaccine against this clade. However, no particular differences were present between the two clades in the ectodomain with respect to the vaccine strain, suggesting that other proteins play a greater role in PRRSV protection (Vanhee et al., 2011).

Remarkably, the major differences between the two clades when compared to the vaccines were concentrated in the signal peptide, a region that has been reported to be a virulence determinant (Allende et al., 2000). This observation suggests a different virulence of the two clades, with the less virulent Clade A (more similar to the attenuate vaccine strain) being progressively outclassed by Clade B. Moreover, the presence of several positively selected sites predicted in the signal peptide presents conflicting evidence. Even if the presence of a decoy epitope in this region has been reported, a recent study has demonstrated that, in the vast majority of cases, the signal peptide is cleaved and does not become part of the mature protein (Thaa et al., 2013). Nevertheless, there is an overlapping between ORF5 and ORF5a, located 5 nucleotides downstream from the GP5 starting codon. A fascinating hypothesis is that the high number of non-synonymous mutations is due to the superimposition of the sequence of the highly conserved GP5a in a different reading frame (Robinson et al., 2013a). Although the presence of ORF5a could restrain and shape ORF5 sequence evolution, GP5a can be targeted by the host immune response, and as a result can affect viral fitness. Remarkably, contrary to Clade A, Clade B strains displayed several differences with the vaccine strain in aa composition in the ectodomain and

transmembrane regions of the GP5a. This evidence could contribute to differential cross protection. However, evidence of the non-neutralizing role of the antibodies directed against GP5a (Robinson et al., 2013b), coupled with the purifying nature of the selection acting on this protein imposes caution when proposing the immune pressure on GP5a as one of the determinants of viral fitness.

The N-linked glycans of the GP5 ectodomain may be critical for the proper folding and functioning of the protein (Ansari et al., 2006). For PRRSV it has been reported that glycans act as a shield that masks epitopes from humoral immunity (Ansari et al., 2006, Darwich et al.,2010, Faaberg et al., 2006, 2010,Vu et al., 2011). Several glycosylation profiles were identified in our study, with the most prevalent being glycosylation in positions 37, 46, and 53. Exceptions were represented mainly by a terminal branch, suggesting poor fitness of these mutants. However, a long-lasting clade presenting glycosylation in position 36 was identified, suggesting a certain interchangeability of these sites. Contrary from what was previously reported in literature (Wissink et al., 2004), the presence of several strains collected at different points in time without N-46 confirms that this glycosylation is not essential for viral infectivity, a phenomenon recently reported for Type 2 (Wei et al., 2012). However, the high prevalence of this glycosylation site suggests that it could influence susceptibility to pig immune response (Ansari et al., 2006, Vu et al., 2011) and may consequently have repercussions on pigs' performance (Badaoui et al., 2013). Accordingly, the extremely limited number of strains lacking glycosylation in positions 35 to 37 and 53 could be due to an earlier and higher induction of neutralizing antibodies by strains devoid of these glycosylation sites, as reported by Faaberg et al., (2006). The selection of strains inducing weaker neutralizing antibodies has presumably driven the trend reported in our study (Mateu et al., 2006).

5. Conclusions

In summary, our study reconstructed PRRSV population dynamics in Italy from its probable emergence till nowadays, demonstrating that viral population, after an initial expansion, is currently slowly decreasing. However, different clades, likely in competition against one another, have been displaying different trends. Some evidence is in favor of the role of immune pressure in affecting GP5 evolution, favoring frequent changes in regions exposed to the host immune response and preserving the glycosylation profiles that can hamper humoral immunity.

Further studies, based on the entire genome sequencing, may be necessary to highlight the role of immunity induced by other proteins in conditioning the viral population on an epidemiological scale. Such data and approach could provide remarkable information for the evaluation of the effectiveness of strategies and tools used for the control of this disease.

6. Acknowledgements

The work here described was funded by a grant from the University of Padua (ex 60%, 2013).

7. References

Abascal, F., Zardoya, R., Telford, M.J., 2010. TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 38, W7-W13.

Alfonso-Morales, A., Martinez-Perez, O., Dolz, R., Valle, R., Perera, C.L., Bertran, K., Frias, M.T., Majo, N., Ganges, L., Perez, L.J., 2013. Spatiotemporal Phylogenetic Analysis and Molecular Characterisation of Infectious Bursal Disease Viruses Based on the VP2 Hyper-Variable Region. PLoS One 8, e65999.

Allende, R., Kutish, G.F., Laegreid, W., Lu, Z., Lewis, T.L., Rock, D.L., Friesen, J., Galeota, J.A., Doster, A.R., Osorio, F.A., 2000. Mutations in the genome of porcine reproductive and respiratory syndrome virus responsible for the attenuation phenotype. Arch. Virol. 145, 1149-1161.

Anisimova, M., Gil, M., Dufayard, J., Dessimoz, C., Gascuel, O., 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst. Biol. 60, 685-699.

Ansari, I.H., Kwon, B., Osorio, F.A., Pattnaik, A.K., 2006. Influence of N-linked glycosylation of porcine reproductive and respiratory syndrome virus GP5 on virus infectivity, antigenicity, and ability to induce neutralizing antibodies. J. Virol. 80, 3994.

Badaoui, B., Grande, R., Calza, S., Cecere, M., Luini, M., Stella, A., Botti, S., 2013. Impact of genetic variation and geographic distribution of porcine reproductive and respiratory syndrome virus on infectivity and pig growth. BMC Vet. Res. 9, 58-6148-9-58.

Baele, G.F., Lemey, P.F., Bedford, T.F., Rambaut, A., Suchard, M.A., FAU, S.M., Alekseyenko, A.V., 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Molecular biology and evolution JID – 8501455.

Balka G, Hornyák A, Bálint A, Kiss I, Kecskeméti S, Bakonyi T, Rusvai M. 2008. Genetic diversity of porcine reproductive and respiratory syndrome virus strains circulating in Hungarian swine herds. Vet Microbiol. 127(1-2):128-35.

Candotti P., Rota-Nodari S., Leotti G., Longo S., Joisel F. 2006. Relationship between piglet PRRS seroprevalence at the end of nursery, and reproduction disorders due to PRRS in sows. In: 19th IPVS Congress, Copenhagen, Denmark, 2006. Volume 1: 54.

Charerntantanakul, W., 2012. Porcine reproductive and respiratory syndrome virus vaccines: Immunogenicity, efficacy, and safety aspects. World J. Virol. 1, 23-30.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: More models, new heuristics and parallel computing. Nature Methods 9, 772.

Darwich, L., Díaz, I., Mateu, E., 2010. Certainties, doubts and hypotheses in porcine reproductive and respiratory syndrome virus immunobiology. Virus Res. 154, 123-132.

de Castro, E., Sigrist, C.J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gasteiger, E., Bairoch, A., Hulo, N., 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res. 34, W362-5.

De Mateo Aznar, Maria. (2008). Porcine reproductive and respiratory syndrome virus: evaluation of the viral circulation and antibody response in swine herds in Veneto and Friuli-Venezia Giulia. PhD Thesis. http://paduaresearch.cab.unipd.it/946/

Díaz, I., Darwich, L., Pappaterra, G., Pujols, J., Mateu, E., 2006. Different European-type vaccines against porcine reproductive and respiratory syndrome virus have different immunological properties and confer different protection to pigs. Virology 351, 249.

Díaz, I., Pujols, J., Ganges, L., Gimeno, M., Darwich, L., Domingo, M., Mateu, E., 2009. In silico prediction and ex vivo evaluation of potential T-cell epitopes in glycoproteins 4 and 5 and nucleocapsid protein of genotype-I (European) of porcine reproductive and respiratory syndrome virus. Vaccine 27, 5603.

Dokland, T., 2010. The structural biology of PRRSV. Virus Res. 154, 86.

Drigo M, Franzo G, Gigli A, Martini M, Mondin A, Gracieux P, Ceglie L. 2014. The impact of porcine reproductive and respiratory syndrome virus genetic heterogeneity on molecular assay performances. J Virol Methods. 202:79-86.

Drummond, A.J., Ho, S.Y., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969.

Faaberg, K.S., Hocker, J.D., Erdman, M.M., Harris, D.L.H., Nelson, E.A., Torremorell, M., Plagemann, P.G.W., 2006. Neutralizing antibody responses of pigs infected with natural GP5 N-glycan mutants of porcine reproductive and respiratory syndrome virus. Viral Immunol. 19, 294.

Firth, A.E., Zevenhoven-Dobbe, J.C., Wills, N.M., Go, Y.Y., Balasuriya, U.B., Atkins, J.F., Snijder, E.J., Posthuma, C.C., 2011. Discovery of a small arterivirus gene that overlaps the GP5 coding sequence and is important for virus production. J. Gen. Virol. 92, 1097-1106.

Forsberg, R., Storgaard, T., Nielsen, H.S., Oleksiewicz, M.B., Cordioli, P., Sala, G., Hein, J., Bøtner, A., 2002. The genetic diversity of European type PRRSV is similar to that of the North American type but is geographically skewed within Europe. Virology 299, 38-47.

Franzo G, Cecchinato M, Martini M, Ceglie L, Gigli A, Drigo M. 2014. Observation of high recombination occurrence of porcine reproductive and respiratory syndrome virus in

field condition. Virus Res. Aug 21. pii: S0168-1702(14)00333-5. doi: 10.1016/j.virusres.2014.08.005. [Epub ahead of print].

Frost, S.D., Volz, E.M., 2010. Viral phylodynamics and the search for an 'effective number of infections'. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365, 1879-1890.

Guindon, S., Dufayard, J., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst. Biol. 59, 307-321.

Gupta, R., Jung, E., Brunak, S., 2004. Prediction of N-glycosylation sites in human proteins.

Ho, S.Y., Shapiro, B., 2011. Skyline-plot methods for estimating demographic history from nucleotide sequences. Mol. Ecol. Resour. 11, 423-434.

Indik S, Schmoll F, Sipos W, Klein D. 2005. Genetic variability of PRRS virus in Austria: consequences for molecular diagnostics and viral quantification. Vet Microbiol.107(3-4):171-8.

Kim WI, Kim JJ, Cha SH, Wu WH, Cooper V, Evans R, Choi EJ, Yoon KJ. 2013. Significance of genetic variation of PRRSV ORF5 in virus neutralization and molecular determinants corresponding to cross neutralization among PRRS viruses. Vet Microbiol. 162(1):10-22.

Kimman, T.G., Cornelissen, L.A., Moormann, R.J., Rebel, J.M.J., Stockhofe-Zurwieden, N., 2009. Challenges for porcine reproductive and respiratory syndrome virus (PRRSV) vaccinology. Vaccine 27, 3704-3718.

Korber, B., LaBute, M., Yusim, K., 2006. Immunoinformatics comes of age. PLoS Computational Biology 2, 0484.

Lunney, J.K., Benfield, D.A., Rowland, R.R.R., 2010. Porcine reproductive and respiratory syndrome virus: An update on an emerging and re-emerging viral disease of swine. Virus Res. 154, 1-6.

Martin, D.P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26, 2462-2463.

Mateu E, Martín M, Vidal D. 2003. Genetic diversity and phylogenetic analysis of glycoprotein 5 of European-type porcine reproductive and respiratory virus strains in Spain. J Gen Virol. (Pt 3):529-34.

Mateu, E., Diaz, I., 2008. The challenge of PRRS immunology. Veterinary Journal 177, 345-351.

Mateu, E., Díaz, I., Darwich, L., Casal, J., Martín, M., Pujols, J., 2006. Evolution of ORF5 of Spanish porcine reproductive and respiratory syndrome virus strains from 1991 to 2005. Virus Res. 115, 198-206.

Minin, V.N., Bloomquist, E.W., Suchard, M.A., 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Mol. Biol. Evol. 25, 1459-1471.

Murcia, P.R., Wood, J.L., Holmes, E.C., 2011. Genome-scale evolution and phylodynamics of equine H3N8 influenza A virus. J. Virol. 85, 5312-5322.

Murtaugh, M.P., Genzow, M., 2011. Immunological solutions for treatment and prevention of porcine reproductive and respiratory syndrome (PRRS). Vaccine 29, 8192-8204.

Murtaugh, M.P., Stadejek, T., Abrahante, J.E., Lam, T.T.Y., Leung, F.C., 2010. The ever-expanding diversity of porcine reproductive and respiratory syndrome virus. Virus Res. 154, 18-30.

Nguyen VG, Kim HK, Moon HJ, Park SJ, Chung HC, Choi MK, Park BK. 2014. A Bayesian Phylogeographical Analysis of Type 1 Porcine Reproductive and Respiratory Syndrome Virus (PRRSV). Transbound Emerg Dis.;61(6):537-45.

Ostrowski M, Galeota JA, Jar AM, Platt KB, Osorio FA, Lopez OJ. 2002. Identification of neutralizing and nonneutralizing epitopes in the porcine reproductive and respiratory syndrome virus GP5 ectodomain. J Virol. 76(9):4241-50.

Pesch, S., Meyer, C., Ohlinger, V.F., 2005. New insights into the genetic diversity of European porcine reproductive and respiratory syndrome virus (PRRSV). Vet. Microbiol. 107, 31-48.

Pesente, P., Rebonato, V., Sandri, G., Giovanardi, D., Ruffoni, L.S., Torriani, S., 2006. Phylogenetic analysis of ORF5 and ORF7 sequences of porcine reproductive and respiratory syndrome virus (PRRSV) from PRRS-positive Italian farms: A showcase for PRRSV epidemiology and its consequences on farm management. Vet. Microbiol. 114, 214-224.

Plagemann PG, Rowland RR, Faaberg KS. 2002. The primary neutralization epitope of porcine respiratory and reproductive syndrome virus strain VR-2332 is located in the middle of the GP5 ectodomain. Arch Virol. 147(12):2327-47.

Pond, S.L.K., 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21, 676.

Pybus, O.G., Drummond, A.J., Nakano, T., Robertson, B.H., Rambaut, A., 2003. The epidemiology and iatrogenic transmission of hepatitis C virus in Egypt: a Bayesian coalescent approach. Mol. Biol. Evol. 20, 381-387.

Robinson, S.R., Abrahante, J.E., Johnson, C.R., Murtaugh, M.P., 2013a. Purifying selection in porcine reproductive and respiratory syndrome virus ORF5a protein influences variation in envelope glycoprotein 5 glycosylation. Infect. Genet. Evol. 20, 362-368.

Robinson, S.R., Figueiredo, M.C., Abrahante, J.E., Murtaugh, M.P., 2013b. Immune response to ORF5a protein immunization is not protective against porcine reproductive and respiratory syndrome virus infection. Vet. Microbiol. 164, 281-285.

Shi, M., Lam, T.T.Y., Hon, C., Hui, R.K., Faaberg, K.S., Wennblom, T., Murtaugh, M.P., Stadejek, T., Leung, F.C., 2010a. Molecular epidemiology of PRRSV: A phylogenetic perspective. Virus Res. 154, 7-17.

Shi, M., Lam, T.T., Hon, C., Murtaugh, M.P., Davies, P.R., Hui, R.K., Li, J., Wong, L.T., Yip, C., Jiang, J., Leung, F.C., 2010b. Phylogeny-based evolutionary, demographical, and geographical dissection of north american type 2 porcine reproductive and respiratory syndrome viruses. J. Virol. 84, 8700-8711.

Shi M, Lemey P, Singh Brar M, Suchard MA, Murtaugh MP, Carman S, D'Allaire S, Delisle B, Lambert MÈ, Gagnon CA, Ge L, Qu Y, Yoo D, Holmes EC, Chi-Ching Leung F. 2013.The spread of type 2 Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) in North America: a phylogeographic approach. Virology. Dec;447(1-2):146-54.

Stadejek T, Stankevicius A, Storgaard T, Oleksiewicz MB, Belák S, Drew TW, Pejsak Z. 2002. Identification of radically different variants of porcine reproductive and respiratory syndrome virus in Eastern Europe: towards a common ancestor for European and American viruses. J Gen Virol. 83(Pt 8):1861-73.

Stadejek T, Oleksiewicz MB, Potapchuk D, Podgórska K. 2006. Porcine reproductive and respiratory syndrome virus strains of exceptional diversity in eastern Europe support the definition of new genetic subtypes. J Gen Virol. 87(Pt 7):1835-41.

Stadejek, T., Oleksiewicz, M.B., Scherbakov, A.V., Timina, A.M., Krabbe, J.S., Chabros, K., Potapchuk, D., 2008. Definition of subtypes in the European genotype of porcine reproductive and respiratory syndrome virus: Nucleocapsid characteristics and geographical distribution in Europe. Arch. Virol. 153, 1479-1488.

Stadejek, T., Stankevicius, A., Murtaugh, M.P., Oleksiewicz, M.B., 2013. Molecular evolution of PRRSV in Europe: current state of play. Vet. Microbiol. 165, 21-28.

Thaa, B., Sinhadri, B.C., Tielesch, C., Krause, E., Veit, M., 2013. Signal peptide cleavage from GP5 of PRRSV: a minor fraction of molecules retains the decoy epitope, a presumed molecular cause for viral persistence. PLoS One 8, e65548.

Van Breedam, W., Delputte, P.L., Van Gorp, H., Misinzo, G., Vanderheijden, N., Duan, X., Nauwynck, H.J., 2010. Porcine reproductive and respiratory syndrome virus entry into the porcine macrophage. J. Gen. Virol. 91, 1659.

Vanhee, M., Van Breedam, W., Costers, S., Geldhof, M., Noppe, Y., Nauwynck, H., 2011. Characterization of antigenic regions in the porcine reproductive and respiratory syndrome virus by the use of peptide-specific serum antibodies. Vaccine 29, 4794-4804.

Vu, H.L.X., Kwon, B., Yoon, K., J. Laegreid, W.W., Pattnaik, A.K., Osorio, F.A., 2011. Immune Evasion of Porcine Reproductive and Respiratory Syndrome Virus through Glycan Shielding Involves both Glycoprotein 5 as Well as Glycoprotein 3. J. Virol. 85, 5555-5564.

Wei, Z., Lin, T., Sun, L., Li, Y., Wang, X., Gao, F., Liu, R., Chen, C., Tong, G., Yuan, S., 2012. N-linked glycosylation of GP5 of porcine reproductive and respiratory syndrome virus is critically important for virus replication in vivo. J. Virol. 86, 9941-9951.

Wissink, E.H., Kroese, M.V., Maneschijn-Bonsing, J.G., Meulenberg, J.J., van Rijn, P.A., Rijsewijk, F.A., Rottier, P.J., 2004. Significance of the oligosaccharides of the porcine reproductive and respiratory syndrome virus glycoproteins GP2a and GP5 for infectious virus production. J. Gen. Virol. 85, 3715-3723.

Xia, X., 2013. DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. Mol. Biol. Evol. 30, 1720-1728.

Yoon, S.H., Kim, H., Park, B., Kim, H., 2012. Tracing the genetic history of porcine reproductive and respiratory syndrome viruses derived from the complete ORF 5-7 sequences: a Bayesian coalescent approach. Arch. Virol. , 2143-2151.

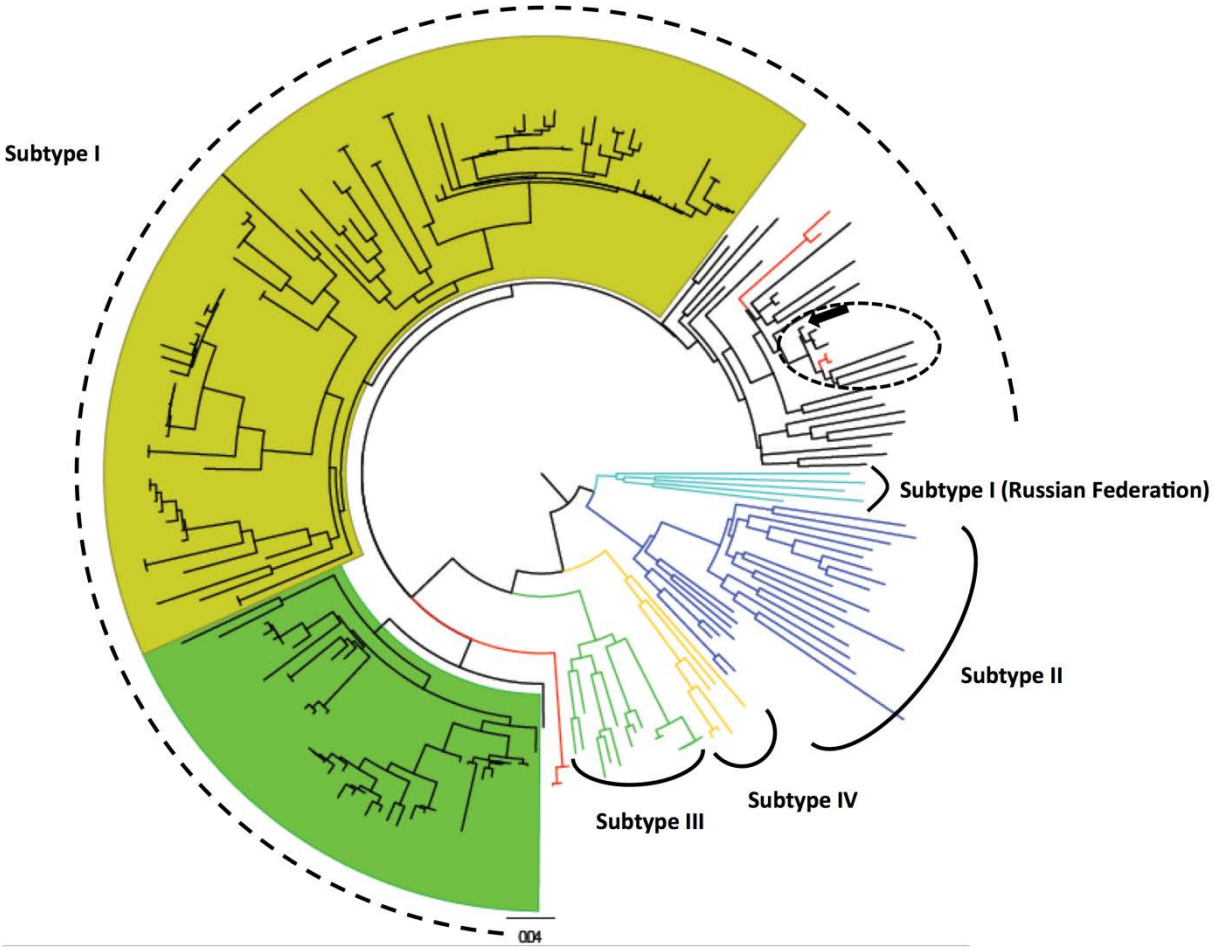

Figure 1. Maximum likelihood phylogenetic tree of ORF5 obtained from Italian strains plus some Type 1 strains obtained from literature. Clade A, B are highlighted in yellow and green, respectively. Type 1 subtype 1 (Russian Federation) is highlighted in light blue; subtype 2 in blue; subtype 3 in green and 4 in yellow (according to Stadejek classification). Italian strains that are not part of the above-mentioned clades are highlighted in red. Lelystad (black arrow)-like viruses are highlighted by a dotted circle.

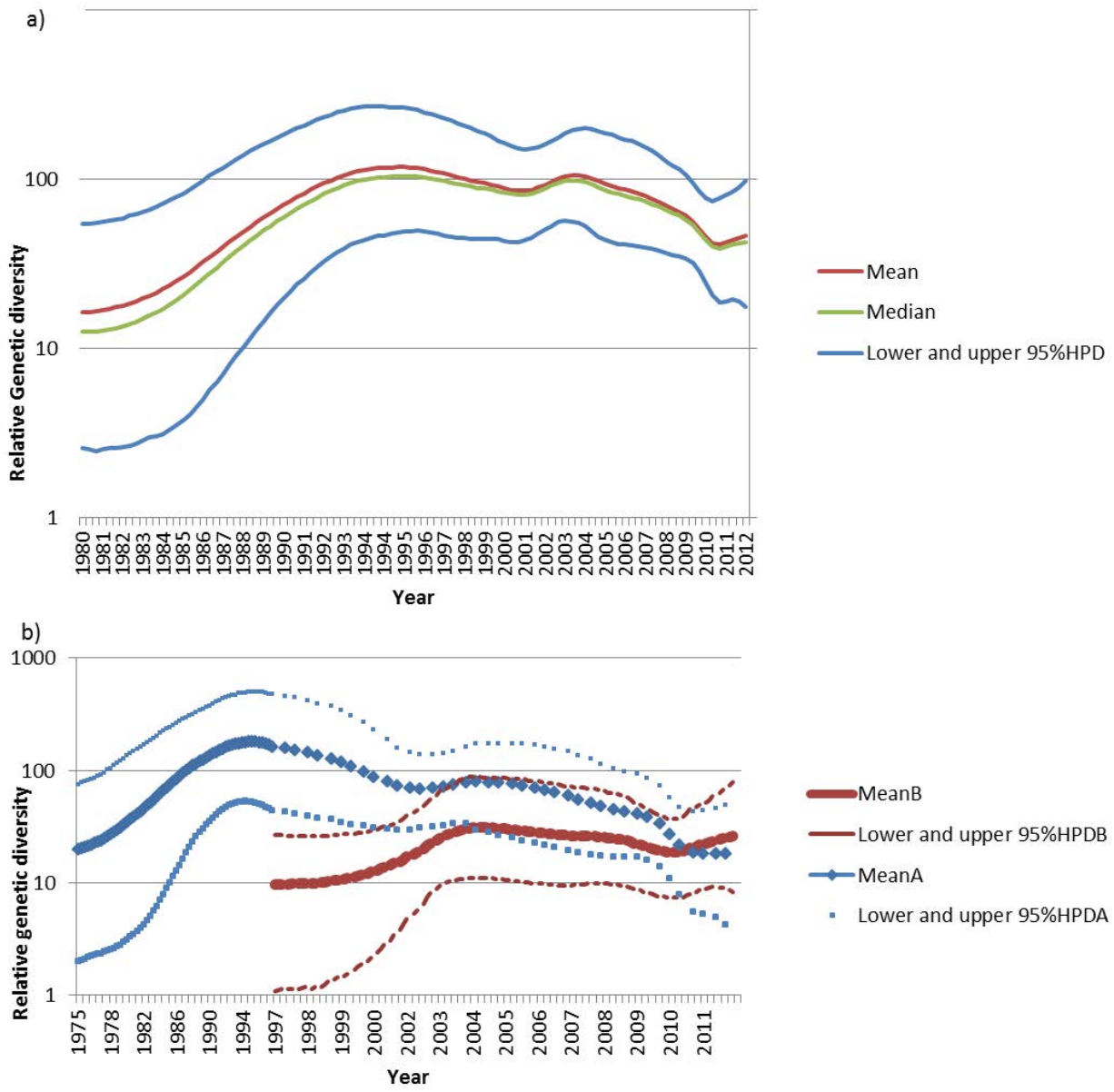

Figure 2. a) Population dynamics of PRRSV. Bayesian skyride plots showing the changing levels of relative genetic diversity through time (Ne*t) estimated using ORF5 of both clades. Mean values of Ne*t are given in red while median value is given in green. Blue lines denote the 95%HPD values. b) Population dynamics of Clade A and B. Mean and 95%HPD are reported. X axis resolution is increased after the introduction of Clade B to better display the dynamics and interaction between the two groups.

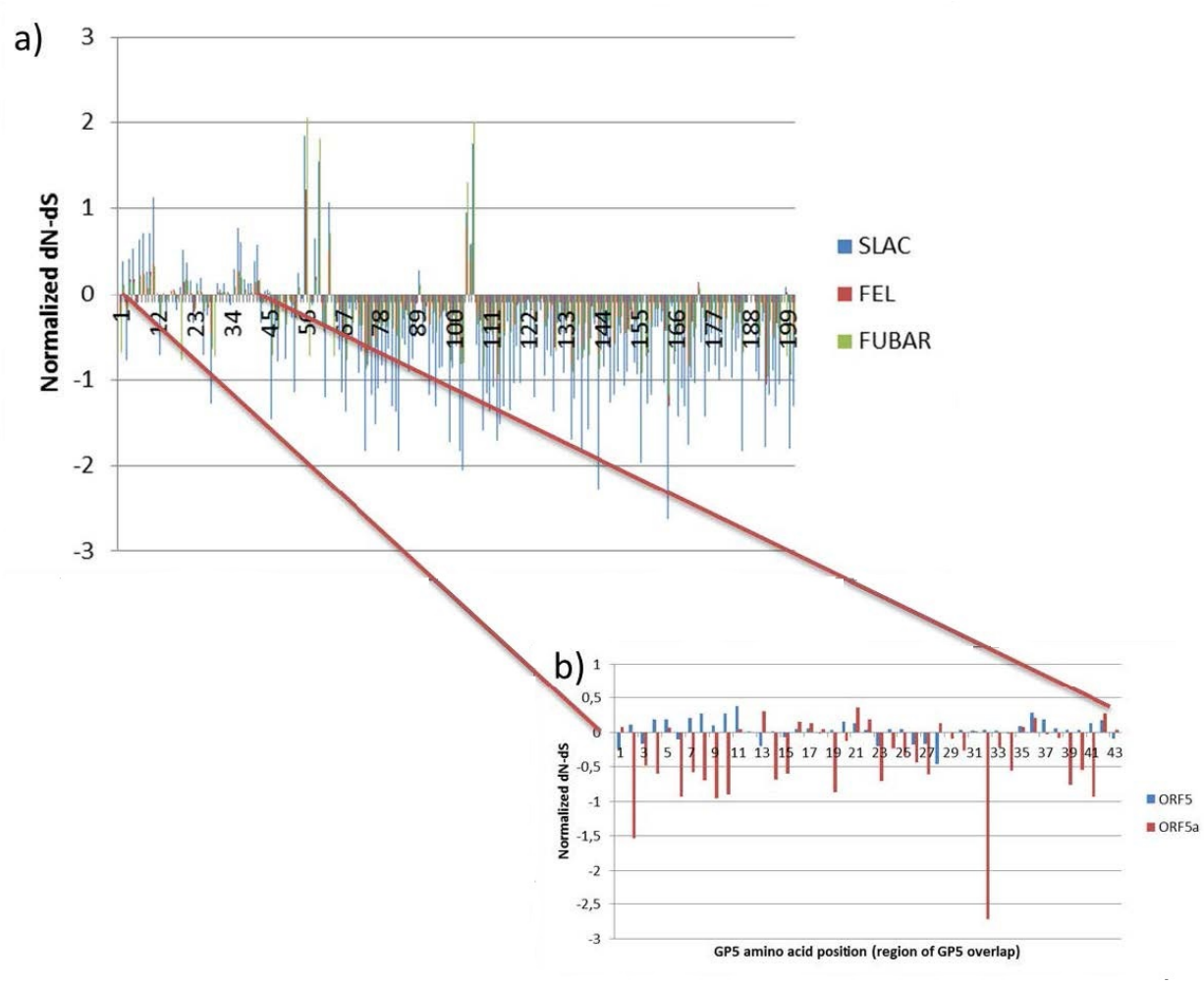

Figure 3. Diversifying selection acting on GP5 (a) and GP5a (b). Normalized dN-dS is displayed for each codon position. Only results of FEL methods are reported for GP5a for ease of representation.

Codon	SLAC dN-dS	SLAC p-value	FEL dN-dS	FEL p-value	MEME ω	MEME p-value	FUBAR dN-dS	FUBAR Post. Pr.	N° Methods
7	0.632	0.007	0.211	0.004	>100	0.007	0.226	0.990	**4**
8	0.718	0.012	0.266	0.006	>100	0.009	0.241	0.981	**4**
10	0.720	0.038	0.273	0.010	>100	0.013	0.228	0.970	**4**
11	1.131	0.000	0.378	0.000	>100	0.000	0.334	0.999	**4**
20	0.520	0.041	0.154	0.005	>100	0.001	0.177	0.985	**4**
21	0.369	0.127	0.128	0.014	>100	0.024	0.156	0.968	**3**
36	0.775	0.008	0.286	0.001	>100	0.002	0.272	0.996	**4**
37	0.608	0.034	0.184	0.007	>100	0.000	0.198	0.978	**4**
41	0.390	0.118	0.135	0.023	>100	0.005	0.153	0.947	**3**
42	0.572	0.061	0.169	0.021	>100	0.019	0.177	0.955	**4**
56	1.835	0.005	1.214	0.000	16.692	0.000	2.047	1.000	**4**
60	1.544	0.027	1.120	0.006	2.671	0.023	1.808	0.998	**4**
63	1.069	0.030	0.511	0.003	5.619	0.008	0.714	0.999	**4**
104	0.954	0.089	0.773	0.002	7.302	0.002	1.309	0.981	**4**
106	1.748	0.003	0.951	0.006	10.431	0.001	2.001	0.997	**4**

Table 1. Codon positions detected under diversifying selection with more than two methods are reported. Comparison between rates of non-synonymous and synonymous substitutions is expressed as dN-dS or ω (i.e. dN/dS). The statistical significance of positive over neutral selection is expressed in term of p-value (SLAC, FEL and MEME methods) or posterior probability (FUBAR method).

	Glycosylation site[a]					Glycosylation Number[b]		
	35	36	37	46	53	1	2	3
CladeA	0,035	0,0814	0,814	0,860	0,965	0,035	0,151	0,814
CladeB	0	0	0,949	0,821	1	0	0,231	0,769

[a] Percentage of strains of each clade presenting that potential glycosylation site
[b] Percentage of strains of each clade presenting indicted number of potential glycosylation site

Table 2. Distribution of amino acid positions in the potential N-glycosylation sites and number of potential N-glycosylation sites predicted in GP5s.

Journal of Virological Methods 201 (2014) 79–85

Contents lists available at ScienceDirect

Journal of Virological Methods

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

Validation and comparison of different end point and real time RT-PCR assays for detection and genotyping of porcine reproductive and respiratory syndrome virus

Michele Drigo[a,*], Giovanni Franzo[a], Ilaria Belfanti[b], Marco Martini[a], Alessandra Mondin[a], Letizia Ceglie[b]

[a] Department of Animal Medicine, Production and Health (MAPS), University of Padova, Viale dell'Università 16, 35020 Legnaro, PD, Italy
[b] Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università 10, 35020 Legnaro, PD, Italy

Article history:
Received 28 June 2013
Received in revised form 19 February 2014
Accepted 21 February 2014
Available online 5 March 2014

Keywords:
PRRSV
Genotyping
RT-PCR
qRT-PCR

A B S T R A C T

The accuracy and rapid diagnosis of PRRSV infection is a major prerequisite for every control and/or eradication strategy. In this study two real time RT-PCR based on different chemistry analysis (TaqMan Probes and SYBR Green) have been developed and validated before comparison to an end point two-step RT-PCR validated previously. All assays were aimed at discrimination between PRRSV genotypes. Furthermore, an exogenous internal control (IC) system had also been implemented in qRT-PCR. A rigorous analytical validation, executed on infected cell cultures and serum, demonstrated good sensitivity, specificity and repeatability. In particular RT-PCR was exceptionally sensitive and could detect a viral titre in the order of a magnitude of 1 copies/μL, 10-fold lower than other qRT-PCR described in this study. Optimal diagnostic performances have been demonstrated analyzing samples retrieved from an experimental infection, with RT-PCR again outperforming real time RT-PCR assays. All tests, showing substantial agreement between them, were able to detect early stages of viraemia (1 DPI) and some animals were classified as positive until the end of the study (76 DPI). Therefore, this supports the assays usefulness in animals with different clinical conditions and in a broad range of epidemiological scenarios. The benefits and disadvantages of different assays were also considered and discussed.

1. Introduction

Porcine reproductive and respiratory syndrome (PRRS) has emerged as one of the most prevalent diseases of swine in the world, causing severe economic losses due to both clinical manifestations as well as the costs related to its control (Lunney et al., 2010). PRRSV, a member of order *Nidovirales*, family *Arteriviridae*, genus *Arterivirus* has a positive single strand RNA genome of approximately 15 kb, carrying at least nine open reading frames (ORF) (Shi et al., 2010a). PRRSV has been classified in two genotypes: type 1 and type 2, originally discovered in Europe and North America, respectively. Nowadays both genotypes display, with different prevalence, a worldwide distribution (Shi et al., 2010a). Despite their quite simultaneous emergence and comparative clinical manifestation, they differ greatly in terms of genetic content (50–70% of nucleotide identity) (Cho and Dee, 2006; Shi et al., 2010a, 2010b). Genetic distance within the genotype was also proven to be noteworthy, most likely as a consequence of both great mutation rate and viral population size (Holmes, 2009; Murtaugh et al., 2010). Though the impact of PRRSV genetic variability on farming strategies has been proven to be crucial, its effect on biological and pathological properties of different viral strains has not been adequately elucidated. Besides considerations on vaccination efficacy, each control strategy must be founded on efficient and reliable surveillance systems (Corzo et al., 2010; Thanawongnuwech and Suradhat, 2010). Several diagnostic tools such as serology or virus isolation have been applied. The highly variable and transient immunological response that has been demonstrated does not allow discrimination between present and past infection or between active/passive immunity. On the other hand, virus isolation is labour intensive, easily affected by sample quality and storage and implies long response time; moreover it does not allow an easy discrimination between genotypes. Other techniques as immunohistochemistry and immunofluorescence are used for PRRSV antigen detection, but there is some concern about their sensitivity (Martínez et al., 2008). All these aspects raise an issue

* Corresponding author. Tel.: +39 049 827 2968; fax: +39 049 827 2973.
E-mail address: michele.drigo@unipd.it (M. Drigo).

http://dx.doi.org/10.1016/j.jviromet.2014.02.015
0166-0934/

for field veterinarians and authorities in choosing effective tools to counteract this challenging disease. Molecular biology assays regarded as sensitive, specific, rapid and relatively cheap have gained an increasing favour as diagnostic tools in human and veterinary medicine (Belák, 2007; Hoffmann et al., 2009; Mackay, 2007). Several in house and commercial RT-PCR based methods have been developed for viral infection diagnosis (Belák, 2005, 2007; Lu et al., 2013; Pol et al., 2013; Schmitt and Henderson, 2005), including PRRSV (Toplak et al., 2012). In this paper the development and validation of two real time and one end point RT-PCR protocols aimed at PRRSV detection and genotyping are described. A collection of experimental samples has been used to explore the sensitivity of the assays during different phases of viraemia.

2. Materials and methods

2.1. PRRSV isolates

One isolate of Italian type 1 genotype (Lelystad like, GenBank acc. no. AY035969) and one of type 2 (GenBank acc. no. U87392.3, reference strain-VR-2332), cultured in MARC-145 with a titre of $10^{6.2}$ copies/mL, were used as positive references for PCR validation on cell cultures. The quantification was performed as described by Bonilauri, 2007. One pig was purchased from a PRRSV-negative source farm and inoculated with 2 mL of supernatant obtained from MARC-145 infected with the same Italian PRRSV type 1 isolate (3×10^7 copies/mL) to obtain a positive serum. After titration, viral concentration in serum was estimated $10^{6.1}$ copies/ml. Negative samples were represented by mock inoculated MARC-145 and serum from a seronegative SPF animal.

2.2. Sample preparation and RNA extraction

RNA was extracted from the cell cultures or fresh refrigerated samples. Cell cultures were lysed through a freezing–thawing cycle and centrifuged for 5 min at $1800 \times g$. Two hundreds μL of supernatant were used for RNA extraction using High Pure RNA Isolation Kit (Roche Diagnostics, Monza, Italy). Blood samples were centrifuged 10 min at $1800 \times g$ and 200 μL of serum was extracted using the High Pure Viral RNA Kit (Roche Diagnostics, Monza, Italy) according to the manufacturer's instructions. Organs (lung, lymph nodes, spleen and liver) were homogenized after the addition of 10 mL of PBS for each gram of tissue. Two hundreds μL of homogenate was extracted using High Pure RNA Tissue Kit (Roche Diagnostics, Monza, Italy). Before extraction, each 200 μL aliquot was added with 5 μL of solution containing 8×10^5 copies/μL of RNA Internal Control (IC) (Hoffmann et al., 2006) immediately following the addition of a lysis buffer. All extracts were stored at −80 °C until testing.

2.3. Primer and probe selection

In order to compare the respective advantages and limits, two real time RT-PCRs, based on different chemistry analysis (i.e. TaqMan Probe and SYBR Green) were validated. Primer pairs, published by Lurchachaiwong (Lurchachaiwong et al., 2008), were used for both qRT-PCR. In addition to type 1 and type 2 specific probes (EU-1 and US respectively) reported in that study, one more probe (EU-2) was designed using Beacon Designer7 (PREMIER Biosoft International) to deal with the high genetic diversity of type 1 Italian strains (Forsberg et al., 2002; Pesch et al., 2005). A commercially available heterologous IC system (Hoffmann et al., 2006) was included both in SYBR Green and TaqMan Probe-based assays. Briefly, a segment of enhanced green fluorescent protein (EGFP) gene was amplified and inserted in a standard cloning vector. After in vitro transcription the IC was quantified and stored using a special RNA-safe buffer (RSB). Primers and probes specific for IC were selected according to Hoffmann et al. (Hoffmann et al., 2006) and implemented in the assays. Sequences of primers and probes are reported in Table 1.

2.4. Real time RT-PCR protocols

In order to obtain the highest PRRSV sensitivity and the lowest competition between IC and pathogen specific reactions, both TaqMan and SYBR Green-based multiplex real time RT-PCR were optimized by checking different PCR conditions. Particular attention was given to minimize the IC fluorescence contribution of the total signal in the SYBR Green assay. SYBR Green and Probe assays were performed on the LightCycler®480 system (Roche Diagnostics, Monza, Italy), using SuperScript® III Platinum® SYBR® Green one-step qPCR Kit and SuperScript® III Platinum® One-Step qRT-PCR kit respectively (Life Technologies, Monza, Italy). The data was analyzed using LightCycler®480 Software V1.5 (Roche Diagnostics, Monza, Italy).

2.4.1. TaqMan-based one-step qRT-PCR

One μL of extracted RNA was added to a standard reaction mix containing 0.2 μL of SuperScript® III RT/Platinum® Taq Mix, 0.3 μM of primer PRRSVf1 and PRRSVr2, 0.4 μM of primer EGBP-1-F and EGBP-2-R, 0.2 μM of PRRSV specific probes (EU-1, EU-2, US) and 0.4 μM of EGBP-Cy5. Sterile nanopure water was added to bring the final volume up to 10 μL. The cycling parameters were 50 °C for 15 min, 95 °C for 2 min, 40 cycles of 95 °C for 10 s and 60 °C for 30 s. The fluorescence signal was acquired for each cycle at the end of the extension phase.

2.4.2. SYBR Green one-step qRT-PCR

One μL of extracted RNA was added to a standard reaction mix containing 0.2 μL of SuperScript® III RT/Platinum® Taq Mix, 0.4 μM of primer PRRSVf1 and PRRSVr2, 0.1 μM of primer EGBP-1-F and EGBP-2-R. Sterile nanopure water was added to bring the final volume up to 10 μL. The cycling parameters were 50 °C for 15 min, 95 °C for 5 min, 40 cycles of 95 °C for 10 s and 60 °C for 30 s. The fluorescence signal was acquired for each cycle at the end of the extension phase. After incubation for 1 min at 40 °C, melting curves were performed gradually raising temperatures from 64 °C to 95 °C with continuous collection of fluorescence data.

2.5. End point RT-PCR protocol

End point RT-PCR, validated previously, was performed as described by Persia et al. (Persia et al., 2001). To differentiate between PRRSV types, a common forward primer and two specific reverse primers (Table 1), generating amplicons of different lengths, were used allowing a simultaneous detection of both type 1 and type 2 genotypes (181 bp and 282 bp, respectively). In order to increase the analytical sensitivity, a short PRRSV ORF 7-based specific primer had been used in reverse transcription reaction, instead of generic hexamers. According to OIE guidelines, in the original validation of this protocol, analytical performances were investigated. Limit of detection (LoD) (i.e. Lower limit of detection–see chapter 1.1.5. – OIE Terrestrial Manual 2012, p. 10), repeatability and specificity (inclusivity and exclusivity) were calculated for titrated type 1 and type 2 infected cells as well as for lung homogenates and sera. In addition, diagnostic sensitivity had been evaluated periodically through a panel of well-characterized field strains among the positive samples collected throughout the diagnostic activity.

2.6. Analytical validation

Analytical validation was performed for real time RT-PCR assays on serum and cell culture. In order to determine the LoD of

Table 1
Primer and probes implemented in the assays described in this study.

Primer/Probe	Oligonucleotides	Assay	References
RT	5′-TCGCCCTAAT-3′	RT-PCR	Persia et al., 2001
US/EU-F	5′-ATGGCCAGCCAGTCAATC-3′	RT-PCR	Persia et al., 2001
EU-R	5′-GATTGCAAGCAGAGGGAGCGTTC-3′	RT-PCR	Persia et al., 2001
US-R	5′-GGCGCACAGTATGATGCGTAG-3′	RT-PCR	Persia et al., 2001
PRRSVf1	5′-GGGGAATGGCCAGYCAGTCAA-3′	qRT-PCR	Lurchachaiwong et al., 2008
PRRSVr2	5′-GCCAGRGGAAAATGKGGCTTCTC-3′	qRT-PCR	Lurchachaiwong et al., 2008
US	5′-HEX-CTGGGTAAGATCATCGCCCAGGA-3′-IABkFQ	qRT-PCR	Lurchachaiwong et al., 2008
EU-1	5′-FAM-TTGGCCTGTCCTCCCCTAGGTTG-3′-IABKFQ	qRT-PCR	Lurchachaiwong et al., 2008
EU-2	5′-FAM-ATGATGAAATCCCAGCGCCAGCGGT-3′-IABkFQ	qRT-PCR	Present study
EGFP-1-F	5′-GACCACTACCAGCAGAACAC-3′	IC System	Hoffmann et al. (2006)
EGFP-2-R	5′-GAACTCCAGCAGGACCATG-3′	IC System	Hoffmann et al. (2006)
EGFP-Cy5	5′-Cy5-AGCACCCAGTCCGCCCTGAGCA-IBRQ-3′	IC System	Hoffmann et al. (2006)

methods, a 10-fold dilution of positive in non-infected matrices was performed. All undiluted samples, as well as each with a dilution of less than or equal to 10^{-7} were extracted three times and tested in duplicate by two operators using both TaqMan Probe and SYBR Green-based qRT-PCR. Real time RT-PCRs sensitivity was defined as the lowest concentration detected at least in 50% of the replicates for each dilution. *Intra*-laboratory repeatability was estimated choosing three dilutions representing high, medium and low viral concentration and having them tested with both methods by two operators in two different days. The selected RNA dilutions were: 10^{-1}, 10^{-3}, 10^{-5} for type 1 infected cells and undiluted RNA, 10^{-2}, 10^{-4} for both type 2 infected cells and serum. In each run a no template control, constituted by the negative matrix, was also included. Specificity was tested over a panel of several swine pathogens (haemolytic *Escherichia coli*, *Streptococcus suis*, *Clamydia suis*, *Suid herpesvirus 1*, *Swine influenza virus subtypes H1N1 and H1N2*, *Porcine circovirus type 2*, *Porcine parvovirus* and *Bovine viral diarrhoea virus*).

2.7. Experimental infection

To explore the diagnostic validity in vivo an experimental infection was performed. Seventeen 4-weeks old gilts were purchased from a PRRSV-negative source farm and housed in an isolation unit. After ten days of acclimatization, prior to PRRSV inoculation, serum samples were collected from all pigs via venipuncture of the jugular vein using a Vacutainer system. Twelve gilts were inoculated with both 1 mL of infected serum ($10^{6.1}$ copies/mL) in each nostril and 2 mL intramuscularly. Five animals, kept as control, were housed in a separate room of the isolation unit and then inoculated with saline alone. Serum from each gilt was collected on day 1, 10, 28, 42 and 76 using the same protocol. On day 10 and 28, animal 8 and 3 (randomly chosen) were humanely euthanized and samples from the lungs, lymph nodes, spleen and liver were collected. All samples were stored at −80 °C until processing. RNA extraction was performed, as described previously, and all samples were tested using end point RT-PCR and both real time RT-PCR after validation. This study was approved by the General Direction of animal health and veterinary drug of the *Ministero del Lavoro, della Salute e del Farmaco Veterinario* (30/12/2009 protocol number 11615).

2.8. Data analysis

Diagnostic test performances were evaluated considering several parameters. Aggregated efficiency (E) of both extraction and real time RT-PCR was evaluated through the formula $E = 10^{(-1/\text{slope})} - 1$. Slope was obtained through the calculation of a linear regression between crossing points (C_p) corresponding to 10-fold dilution of each matrix. The goodness of regression line fit in explaining the relationship between dilution and C_p was expressed in terms of R^2. The presence of significant differences between dilution curves and the factors involved were tested using a repeated measure GLM (Burns et al., 2005), using the following approach: the sphericity of variance–covariance matrix of the dependent variable was checked using Mauchly's test for each within subject effect. When this criteria was not met the lower-bound epsilon was used to correct the degree of freedom. Post hoc test based on Bonferroni correction was performed. *Intra*-laboratory repeatability was evaluated calculating coefficient of variation (CV) for different levels of the experiment (replicate, operator, experiment, dilution). Melting temperatures for PRRSV type 1 and type 2 and IC amplicons were expressed as mean ± standard deviation and the statistical significance of their difference was assessed using ANOVA and post hoc test with Bonferroni correction. Diagnostic performances were evaluated comparing results of different methods in terms of Cohen's kappa coefficient. In particular performances were compared within and between different collection times. Statistical analysis was performed using the statistical software SPSS (IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.) where the statistical significance was set to $p = 0.05$.

3. Results

3.1. Primer and probes

TaqMan Probe-based real time RT-PCR allowed PRRSV genotyping and the simultaneous detection of IC using different labelled probes. The colour compensation tool implemented in LightCycler®480 Software V1.5 (Roche Diagnostics, Monza, Italy) obviates the emission spectra overlapping that could lead to the misinterpretation of data. IC was constantly detected in all serum dilution and at dilution higher than 10^{-2} in cell cultures (Table 2).

Similarly, SYBR Green-based assay allowed an easy distinction between the three qRT-PCR products based on different melting temperatures (T_m). In particular type 1, type 2 and IC were characterized by a T_m of 85.57 ± 0.39, 84.25 ± 0.34 and 88.3 ± 0.8 respectively. ANOVA analysis substantiates the significance of these differences ($p = 0.00$). IC was detected only at a dilution equal to or greater than 10^{-1} for serum and 10^{-4} for cell cultures (Table 2).

3.2. Assays validation

Both of the real time RT-PCRs displayed a remarkable analytical sensitivity being able to detect up to about 15 PRRSV copies/μL in both type 1 and 2 PRRSV infected cells and 125 copies/μL in serum. Efficiency was within the range usually considered adequate according to Pestana et al. (2010) (i.e. 90–105%) or really close to it (Table 2). The value of R^2 highlighted the consistent linearity of the relationship between the C_p value and dilution, along with the low

Table 2
Analytical performances for SYBR Green and TaqMan Probe real time RT-PCR.

Assay	Matrix	b[a]	E[b]	R^2[c]	LoD[d] (copies/μL)	IC detection[e] (copies/μL)
SYBR Green	Cell cultured type 1	−3.35	0.99	0.91	15	150
	Cell cultured type 2	−3.60	0.90	0.86	15	150
	Serum	−3.21	1.05	0.89	125	150,000
TaqMan Probe	Cell cultured type 1	−3.14	1.08	0.91	15	15,000
	Cell cultured type 2	−3.67	0.87	0.99	15	15,000
	Serum	−3.16	1.07	0.92	125	1,500,000

[a] Coefficient of regression.
[b] Efficiency.
[c] Coefficient of determination.
[d] Limit of detection.
[e] Highest PRRSV titre at which the IC is first detected.

variability within replicates at the dilution level. The GLM analysis implemented to assess the repeatability of the assays revealed a substantial equality of standard curves. In this context the effect of the operator resulted as significant ($p = 0.00$) for PRRSV type 1 infected cells using TaqMan-based method and for serum using both methods. No statistically significant difference was demonstrated for the factors "replicate", "duplicate", or their interaction support a high *intra*-run repeatability. *Inter*-run repeatability was proven consistent with a maximum CV value of 3.2% (Table 3). Merely the SYBR Green-based method revealed a poor repeatability for cell cultured PRRSV type 2 because of the difficulty of result analysis due to the presence of some aspecific melting peaks. Unfortunately, it was not possible to solve the problem without affecting the sensitivity of the assay and the detection of the IC. However it is noteworthy that from a qualitative point of view repeatability is quite absolute, with only 4 negative samples (three on SYBR Green and one on Probe-based real time RT-PCR) out of 432 tested with qRT-PCR during the whole repeatability assessment. They all corresponded to a lower dilution of type 1 infected cell.

One hundred percent (100%) of the analytical specificity was demonstrated considering the different pathogens tested and the No Template Control included in each run.

End point RT-PCR LoDs were about 4 copies/μL and 1 copy/μL for cells infected respectively with PRRSV type 1 and 2, 1 copy/μL for serum and 40 copies/μL for the lungs. The repeatability, assessed comparing qualitative results, demonstrated the absence of any discordant results between different experiments or operators (Persia et al., 2001).

3.3. Experimental infection

Starting from day 1 post infection (DPI), all sera tested positive by all assays. All animals remained viremic until 28 DPI, with the exception of only one gilt that tested negative to the SYBR test. On days 42, 80% (8 out of 10) of pigs were still positive using RT-PCR while only 40% (4 out of 10) using real time RT-PCR methods. By day 76, only 3 gilts were still positive in RT-PCR and only one in SYBR Green. The agreement between tests pairs was almost perfect until 28 DPI and substantial in the following samples, according with Landis and Koch classification (Landis and Koch, 1977) (Fig. 1). All five tissues collected both on day 10 and 28 proved to be positive with RT-PCR and Probe-based real time RT-PCR. On the other hand, two samples from gilt 3, tested with SYBR Green, were negative due to inhibition. All samples collected from control gilts remained negative through the whole period of experimentation.

4. Discussion

Diagnostic assays based on molecular biology are now widely used in diagnosis of infectious diseases in human and veterinary medicine (Balka et al., 2009; Belák, 2007; Hoffmann et al., 2009). High Sensitivity and Specificity, coupled with the rapidity of diagnostic response prompt their use as a diagnostic tool for both pathogen identification in single individuals as well as for its surveillance at the population level. Several in house methods have been published and many others are commonly used in diagnostic laboratories all over the world. A recent survey, including laboratories from Europe, Turkey and China reported that utilization of real time PCR was widespread and that many laboratories implement a range of molecular tests in routine diagnostic use. All of the partners have at least one assay in the dual-labelled probe format (such as TaqMan) while intercalating dyes were used by about 40% of the participants. Nearly 90% of the participants undertake a one-step, rather than a two-step RT-PCR. It is remarkable that only half of the labs implemented internal controls (Hoffmann et al., 2009).

This study describes the validation and implementation, in an experimental context, of two real time RT-PCR protocols based on different chemistries. It is now well recognized that real time methods display several advantages over end point PCR methods in terms of sensitivity, specificity, rapidity and reduced risk of cross contamination (Belák, 2007; Hoffmann et al., 2009; Pestana et al., 2010) Our assays show good analytical performances, displaying efficiency and low LoD, being able to detect samples diluted to 15 copies/μL. Although degenerate primers were used, the sensitivity was comparable with other published real time RT-PCR (Balka et al., 2009). Dealing with such a variable virus, characterized by a remarkable genetic distance among isolates and reported as one of the viruses with the highest mutation rate (Murtaugh et al., 2010; Shi et al., 2010a), the likelihood of mismatches between pathogen sequences and oligonucleotides used should be considered high. As a consequence, the adoption of degenerated primers, as reported by Lurchachaiwong (Lurchachaiwong et al., 2008) has been implemented in our test as well. Despite the great potentiality demonstrated by PCR assay, several steps during sample processing can affect the performances of the assay and the diagnostic response. The target loss during extraction or the presence of PCR reaction inhibitors during amplification are often an underestimated cause of low sensitivity, leading to an unpredictable false negative. Despite this evidence, only a small percentage of researchers test their nucleic acid for presence of inhibition (Hoffmann et al., 2009; Mackay, 2007). Several approaches can be implemented, choosing between endogenous and exogenous IC. The use of host genes (i.e. endogenous IC) reveals some disadvantages, in particular with regard to the difficulty of selecting a gene that is constantly expressed in different tissues and clinical conditions. To overcome this problem, in this work, an exogenous IC was spiked in each sample just before extraction, minimizing the differential of degradation between non-encapsidated RNA compared to viral nucleic acid (Mackay, 2007). The IC was demonstrated to be constantly expressed, in particular at low viral template concentrations, the most affected by inhibition or poor extraction efficiency.

Table 3
Results of the repeatability performances for real time RT-PCRs. Data are summarized in term of mean (μ) ± standard deviation (sd) and coefficient of variation (CV) considering matrix, replicate, operator and day of experiment as source of variability.

Matrix	Assay	Dilution	Operator	Day 1		Day 2		Total	
				C_p[a] (μ ± sd)	CV[b]	C_p[a] (μ ± sd)	CV[b]	C_p[a] (μ ± sd)	CV[b]
Cell cultured type 1	SYBR Green	−1	1	23.55 ± 0.30	0.013	23.89 ± 0.63	0.026	23.72 ± 0.47	0.020
			2	22.69 ± 0.08	0.004	22.79 ± 0.10	0.004	22.74 ± 0.09	0.004
			Total	23.12 ± 0.51	0.022	23.34 ± 0.72	0.030	23.23 ± 0.60	0.026
		−3	1	30.11 ± 0.20	0.007	31.47 ± 0.39	0.012	30.79 ± 0.80	0.026
			2	29.81 ± 0.39	0.013	30.32 ± 0.31	0.010	30.07 ± 0.42	0.014
			Total	29.96 ± 0.32	0.010	30.89 ± 0.70	0.022	30.43 ± 0.70	0.023
		−5	1	32.28 ± 0.76	0.023	32.42 ± 0.74	0.023	32.35 ± 0.74	0.023
			2	32.99 ± 0.24	0.007	33.39 ± 0.08	0.002	33.19 ± 0.23	0.007
			Total	32.51 ± 0.68	0.020	32.91 ± 0.53	0.016	32.71 ± 0.69	0.021
	TaqMan Probe	−1	1	26.55 ± 0.83	0.031	26.05 ± 0.35	0.014	26.30 ± 0.63	0.024
			2	25.73 ± 0.22	0.008	25.54 ± 0.24	0.009	25.64 ± 0.23	0.009
			Total	26.14 ± 0.71	0.027	25.80 ± 0.39	0.015	25.97 ± 0.57	0.022
		−3	1	33.15 ± 0.29	0.009	33.66 ± 0.58	0.017	33.41 ± 0.50	0.015
			2	32.48 ± 0.23	0.007	33.19 ± 0.48	0.014	32.84 ± 0.50	0.016
			Total	32.82 ± 0.44	0.013	33.43 ± 0.54	0.016	33.13 ± 0.56	0.017
		−5	1	35[c]	0.000	35[c]	0.000	35[c]	0.000
			2	34.88 ± 0.21	0.006	35[c]	0.000	34.94 ± 0.17	0.005
			Total	34.94 ± 0.15	0.004	35[c]	0.000	34.97 ± 0.10	0.003
Cell cultured type 2	SYBR Green	Undiluted	1	16.98 ± 0.98	0.057	19.41 ± 0.81	0.042	18.2 ± 1.56	0.086
			2	19.34 ± 0.72	0.037	17.58 ± 0.14	0.008	18.46 ± 1.07	0.058
			Total	18.16 ± 1.50	0.082	18.50 ± 1.13	0.061	18.33 ± 1.28	0.070
		−2	1	24.52 ± 0.12	0.005	28.01 ± 0.12	0.004	26.27 ± 1.92	0.073
			2	29.70 ± 0.75	0.025	25.62 ± 0.67	0.026	27.66 ± 2.32	0.084
			Total	27.11 ± 2.88	0.106	26.82 ± 1.37	0.051	26.97 ± 2.16	0.080
		−4	1	31.10 ± 0.55	0.018	35[c]	0.000	33.05 ± 2.15	0.065
			2	31.47 ± 0.58	0.018	35[c]	0.000	33.24 ± 1.99	0.060
			Total	31.32 ± 0.71	0.022	35[c]	0.000	33.16 ± 1.99	0.060
	TaqMan Probe	Undiluted	1	20.46 ± 0.58	0.028	20.96 ± 0.22	0.010	20.71 ± 0.48	0.023
			2	19.95 ± 0.34	0.017	20.97 ± 0.28	0.013	20.46 ± 0.61	0.030
			Total	20.2 ± 0.51	0.025	20.96 ± 0.22	0.010	20.58 ± 0.39	0.019
		−2	1	27.84 ± 0.16	0.006	28.60 ± 0.21	0.007	28.22 ± 0.45	0.016
			2	28.14 ± 0.60	0.021	28.95 ± 0.14	0.005	28.55 ± 0.60	0.021
			Total	27.99 ± 0.42	0.015	28.78 ± 0.25	0.008	28.39 ± 0.54	0.019
		−4	1	34.41 ± 0.57	0.016	34.65 ± 0.61	0.018	34.53 ± 0.55	0.016
			2	35[c]	0.000	35[c]	0.000	35[c]	0.000
			Total	34.71 ± 0.48	0.013	34.82 ± 0.43	0.012	34.77 ± 0.45	0.013
Serum	SYBR Green	Undiluted	1	20.07 ± 0.21	0.010	20.41 ± 0.23	0.011	20.24 ± 0.26	0.013
			2	19.74 ± 0.21	0.010	20.16 ± 0.10	0.005	19.95 ± 0.28	0.014
			Total	19.90 ± 0.26	0.013	20.29 ± 0.21	0.010	20.1 ± 0.30	0.015
		−2	1	29.03 ± 0.22	0.007	28.99 ± 0.15	0.005	29.01 ± 0.17	0.006
			2	28.33 ± 0.45	0.016	28.59 ± 0.48	0.017	28.46 ± 0.43	0.015
			Total	28.68 ± 0.50	0.017	28.79 ± 0.38	0.013	28.74 ± 0.43	0.015
		−4	1	34.83 ± 0.30	0.009	35[c]	0.000	34.92 ± 0.21	0.006
			2	34.17 ± 0.35	0.010	35[c]	0.000	34.59 ± 0.52	0.015
			Total	34.50 ± 0.46	0.013	35[c]	0.000	34.75 ± 0.42	0.012
	TaqMan Probe	Undiluted	1	25.64 ± 0.55	0.021	25.80 ± 0.80	0.031	25.72 ± 0.54	0.021
			2	25.32 ± 0.11	0.004	25.43 ± 0.11	0.004	25.38 ± 0.10	0.004
			Total	25.48 ± 0.40	0.016	25.62 ± 0.55	0.021	25.55 ± 0.41	0.016
		−2	1	29.03 ± 0.12	0.004	28.93 ± 0.76	0.026	28.98 ± 0.56	0.019
			2	28.34 ± 0.37	0.013	28.72 ± 0.45	0.016	28.53 ± 0.41	0.014
			Total	28.68 ± 0.40	0.014	28.82 ± 0.61	0.021	28.75 ± 0.64	0.022
		−4	1	35[c]	0.000	35[c]	0.000	35[c]	0.000
			2	35[c]	0.000	35[c]	0.000	35[c]	0.000
			Total	35[c]	0.000	35[c]	0.000	35[c]	0.000

[a] Crossing point.
[b] Coefficient of variation.
[c] All the C_p values greater than 35 are reported by the software to a maximum value of 35.

The presence of a full process IC easily allows detecting false negative caused by PCR inhibitors or loss of RNA during the extraction phase, increasing the reliability of the assay (Hoffmann et al., 2006).

Besides, the use of specific probes that were labelled differently or the presence of characteristic T_m, allowed an easy distinction between different amplifications. However, in particular with detecting systems based on intercalating dyes, interpretability of multiplex PCR can be hampered severely by fluorescence summation. It has been described that the calculation of the ratio between specific amplicon melting peak areas and the integral of the entire melting profile can be used to calculate a coefficient to correct fluorescence value, estimating the contribution of target amplification (Ririe et al., 1997). Nevertheless it should be taken into account that this adjustment is based on the end-point amount of DNA, which is not representative of the exponential phase. As a consequence, any estimation about the initial amount of RNA based on these assumptions would be inaccurate. Although the method described was not intended to be quantitative, it is obvious that C_p value provides itself a semi-quantitative suggestion of initial RNA concentration. To maintain this advantage, the reaction mix was

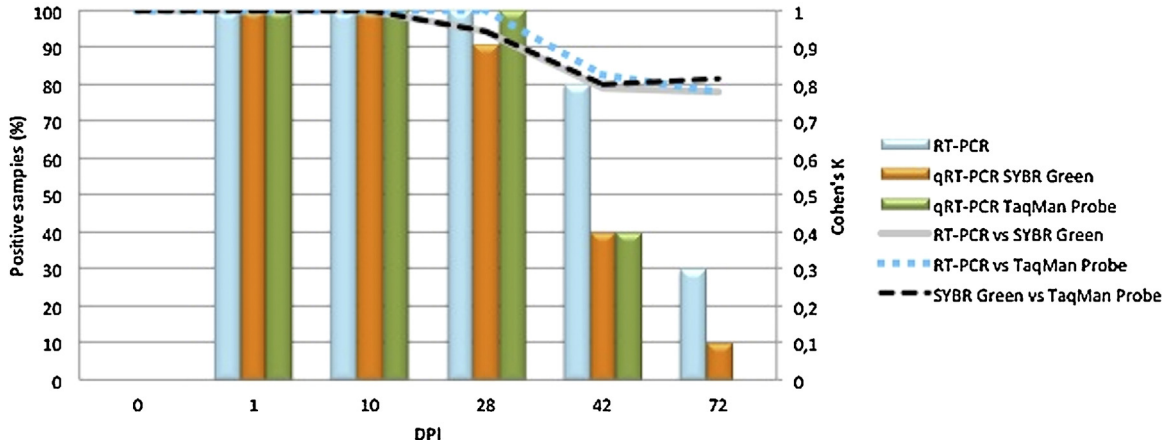

Fig. 1. Diagnostic evaluation of the three diagnostic assays during experimental infection. Lines represent the agreement between assays during the course of infection expressed in terms of Cohen's K. Histograms show the prevalence of animals tested positive.

optimized to lower IC signal contribution to total fluorescence, minimizing one of the main by-products of IC system.

It is nowadays well established that the original geographic separation between type 1 (Europe) and type 2 (North America) is no longer true (Fang et al., 2007; Shi et al., 2010a). In fact both genotypes are present worldwide, even though with different prevalence. So it is essential to set up assays able to detect and differentiate between both genotypes. This aim was easily achieved with TaqMan type specific probes. Also SYBR Green chemistry, as previously described for several pathogens (Chai et al., 2013; Franzo et al., 2014; Kong et al., 2009), allowed genotyping based on T_m. However, genetic variability leading to nucleotidic substitution within amplicon could make this analysis difficult or invalidate its effectiveness. Validation revealed a good *intra*- and *inter*-runs repeatability. Curves built over several sample dilutions were highly repeatable, only showing as significant the effect of the operator. The fact that operator skills and experience can influence the results, it is neither surprising nor nullifying to the assay response. Furthermore it should also be noted that the low *intra*-operator variability makes relatively small *inter*-operator differences significant from a statistical point of view. *Inter*-run repeatability was also high, except for cell culture infected with type 2 that displayed a high CV using SYBR Green-based method, probably due to the presence of some aspecific amplicons. On the other hand, considering the easy discrimination between aspecific and specific amplifications through melting curves analysis and given the qualitative purpose of our real time protocols, the performances of the assay can be considered adequate. According to this principle, from a qualitative point of view, repeatability can be considered almost absolute. In experimentally infected animals, all assays demonstrated good performances, two of them being able to detect positive animals up to 76 DPI. Although previous studies report a far shorter viraemia (Batista et al., 2002; Wills et al., 2003), the optimal/substantial agreement among three independent tests makes the presence of several false positive results unlikely. As a consequence, these results could be due to specific inoculated strain, housing condition or greater assays sensitivity. Although RT-PCR is usually regarded as less sensitive than real time RT-PCR (Chai et al., 2013; Mackay, 2007), in this particular study it matched or outperformed these methods. However, it should be considered that the RNA template added to the mastermix was 10-fold higher, probably justifying the difference in performance. Limiting the reaction volume represents an easy way to reduce costs, making diagnostic assays cheaper and more accessible, in particular in a high-thoroughput laboratory. Furthermore, real time methods revealed some advantages connected with the rapid diagnostic response and to the easiness of implementing an internal control system. Moreover greater specificity can be obtained due to the use of specific probes or T_m analysis coupled with the absence of any post PCR manipulation of amplificates (Hoffmann et al., 2009).

Clinical diagnostic applications and particularly the use of real time PCRs are exponentially growing also in veterinary medicine, rapidly replacing other established diagnostic methods. Rigorous assay validation is an essential prerequisite to gain a good diagnostic level and to support the confidence of its use on everyday field veterinary practice. In this study three methods were developed and validated, evaluating their benefits and disadvantages. By the way, technicians and veterinarians should be advised of the limits and opportunities when choosing a specific assay for specific diagnosis.

Financial support

This study has been supported by Ministry of Health grant (IZS VE 04/07 RC).

Acknowledgments

The authors are grateful to Dr. Bortoletto Giacomo and Dr. De Mateo Aznar Maria for their valuable work done during the conduction of the experimental infection.

References

Balka, G., Hornyák, A., Bálint, A., Benyeda, Z., Rusvai, M., 2009. Development of a one-step real-time quantitative PCR assay based on primer-probe energy transfer for the detection of porcine reproductive and respiratory syndrome virus. J. Virol. Methods 158, 41–45.

Batista, L., Dee, S.A., Rossow, K.D., Dee, J., Pijoan, C., 2002. Assessing the duration of persistence and shedding of porcine reproductive and respiratory syndrome virus in a large population of breeding-age gilts. Can. J. Vet. Res. 66, 196–200.

Belák, S., 2007. Molecular diagnosis of viral diseases, present trends and future aspects. A view from the OIE collaborating centre for the application of polymerase chain reaction methods for diagnosis of viral diseases in veterinary medicine. Vaccine 25, 5444–5452.

Belák, S., 2005. The molecular diagnosis of porcine viral diseases: a review. Acta Vet. Hung. 53, 113–124.

Bonilauri, P., 2007. Developement and validation of a real time (FRET probe) PCR for detection of PRRSV. In: Proceedings of the 5th International Symposium on Emerging and Re-emerging Pig Diseases, p. 194.

Burns, M.J., Nixon, G.J., Foy, C.A., Harris, N., 2005. Standardisation of data from real-time quantitative PCR methods – evaluation of outliers and comparison of calibration curves. BMC Biotechnol. 5, 31.

Chai, Z., Ma, W., Fu, F., Lang, Y., Wang, W., Tong, G., Liu, Q., Cai, X., Li, X., 2013. A SYBR Green-based real-time RT-PCR assay for simple and rapid detection and differentiation of highly pathogenic and classical type 2 porcine reproductive and respiratory syndrome virus circulating in China. Arch. Virol. 158, 407–415.

Cho, J.G., Dee, S.A., 2006. Porcine reproductive and respiratory syndrome virus. Theriogenology 66, 655–662.
Corzo, C.A., Mondaca, E., Wayne, S., Torremorell, M., Dee, S., Davies, P., Morrison, R.B., 2010. Control and elimination of porcine reproductive and respiratory syndrome virus. Virus Res. 154, 185–192.
Fang, Y., Schneider, P., Zhang, W.P., Faaberg, K.S., Nelson, E.A., Rowland, R.R.R., 2007. Diversity and evolution of a newly emerged North American Type 1 porcine arterivirus: analysis of isolates collected between 1999 and 2004. Arch. Virol. 152, 1009–1017.
Forsberg, R., Storgaard, T., Nielsen, H.S., Oleksiewicz, M.B., Cordioli, P., Sala, G., Hein, J., Bøtner, A., 2002. The genetic diversity of European type PRRSV is similar to that of the North American type but is geographically skewed within Europe. Virology 299, 38–47.
Franzo, G., Drigo, M., Lupini, C., Catelli, E., Laconi, A., Listorti, V., Bonci, M., Naylor, C., Martini, M., Cecchinato, M., 2014. A sensitive, reproducible and economic real-time RT-PCR detecting Avian metapneumovirus subtype A and B. Avian Dis., http://dx.doi.org/10.1637/10676-092413-Reg.1.
Hoffmann, B., Beer, M., Reid, S.M., Mertens, P., Oura, C.A.L., van Rijn, P.A., Slomka, M.J., Banks, J., Brown, I.H., Alexander, D.J., King, D.P., 2009. A review of RT-PCR technologies used in veterinary virology and disease control sensitive and specific diagnosis of five livestock diseases notifiable to the World Organisation for Animal Health. Vet. Microbiol. 139, 1–23.
Hoffmann, B., Depner, K., Schirrmeier, H., Beer, M., 2006. A universal heterologous internal control system for duplex real-time RT-PCR assays used in a detection system for pestiviruses. J. Virol. Methods 136, 200–209.
Holmes, E.C., 2009. The Evolution and Emergence of RNA Viruses. Oxford Series in Ecology and Evolution. Oxford University Press, USA.
Kong, L.L., Omar, A.R., Hair Bejo, M., Ideris, A., Tan, S.W., 2009. Development of SYBR green I based one-step real-time RT-PCR assay for the detection and differentiation of very virulent and classical strains of infectious bursal disease virus. J. Virol. Methods 161, 271–279.
Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174.
Lu, G., Yan, H., Yang, Y., Cui, S., Lü, Y., Zhang, X., Zhang, D., Yang, P., Huang, F., Wang, J., Wang, Q., 2013. Evaluation of two commercial real-time PCR kits for detection of pandemic (H1N1) 2009 virus in Beijing. J. Virol. Methods 188, 25–28.
Lunney, J.K., Benfield, D.A., Rowland, R.R., 2010. Porcine reproductive and respiratory syndrome virus: an update on an emerging and re-emerging viral disease of swine. Virus Res. 154, 1–6.
Lurchachaiwong, W., Payungporn, S., Srisatidnarakul, U., Mungkundar, C., Theamboonlers, A., Poovorawan, Y., 2008. Rapid detection and strain identification of porcine reproductive and respiratory syndrome virus (PRRSV) by real-time RT-PCR. Lett. Appl. Microbiol. 46, 55–60.
Mackay, I.M., 2007. Real-Time PCR in Microbiology: From Diagnosis to Characterization. Caister Academic press, Norfolk.
Martínez, E., Riera, P., Sitjà, M., Fang, Y., Oliveira, S., Maldonado, J., 2008. Simultaneous detection and genotyping of porcine reproductive and respiratory syndrome virus (PRRSV) by real-time RT-PCR and amplicon melting curve analysis using SYBR Green. Res. Vet. Sci. 85, 184–193.
Murtaugh, M.P., Stadejek, T., Abrahante, J.E., Lam, T.T., Leung, F.C., 2010. The ever-expanding diversity of porcine reproductive and respiratory syndrome virus. Virus Res. 154, 18–30.
Persia, D., Pacciarini, M., Cordioli, P.S., 2001. Evaluation of three RT-PCR assays for the detection of porcine and respiratory syndrome virus (PRRSV) in diagnostic samples. Proceedings of the 10th International Symposium of Veterinary Laboratory Diagnosticians and OIE Seminar on Biotechnology, pp. 440-441.
Pesch, S., Meyer, C., Ohlinger, V.F., 2005. New insights into the genetic diversity of European porcine reproductive and respiratory syndrome virus (PRRSV). Vet. Microbiol. 107, 31–48.
Pestana, E.A., Belak, S., Diallo, A., Crowther, J.R., Viljoen, G.J., 2010. Analysis and troubleshooting. In: Pestana, E.A., Belak, S., Diallo, A., Crowther, J.R., Viljoen, G.J. (Eds.), Early, Rapid and Sensitive Veterinary Molecular Diagnostics-Real Time PCR Applications. Springer Science + Business Media B.V., Dordrecht, NL, pp. 247–260.
Pol, F., Deblanc, C., Oger, A., Le Dimna, M., Simon, G., Le Potier, M.F., 2013. Validation of a commercial real-time PCR kit for specific and sensitive detection of Pseudorabies. J. Virol. Methods 187, 421–423.
Ririe, K.M., Rasmussen, R.P., Wittwer, C.T., 1997. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Anal. Biochem. 245, 154–160.
Schmitt, B., Henderson, L., 2005. Diagnostic tools for animal diseases. Res. Sci. Tech. 24, 243–250.
Shi, M., Lam, T.T., Hon, C.C., Hui, R.K., Faaberg, K.S., Wennblom, T., Murtaugh, M.P., Stadejek, T., Leung, F.C., 2010a. Molecular epidemiology of PRRSV: a phylogenetic perspective. Virus Res. 154, 7–17.
Shi, M., Lam, T.T., Hon, C.C., Murtaugh, M.P., Davies, P.R., Hui, R.K., Li, J., Wong, L.T., Yip, C.W., Jiang, J.W., Leung, F.C., 2010b. Phylogeny-based evolutionary, demographical, and geographical dissection of north American type 2 porcine reproductive and respiratory syndrome viruses. J. Virol. 84, 8700–8711.
Thanawongnuwech, R., Suradhat, S., 2010. Taming PRRSV: revisiting the control strategies and vaccine design. Virus Res. 154, 133–140.
Toplak, I., Rihtarič, D., Hostnik, P., Grom, J., Štukelj, M., Valenčak, Z., 2012. Identification of a genetically diverse sequence of porcine reproductive and respiratory syndrome virus in Slovenia and the impact on the sensitivity of four molecular tests. J. Virol. Methods 179, 51–56.
Wills, R.W., Doster, A.R., Galeota, J.A., Sur, J.H., Osorio, F.A., 2003. Duration of infection and proportion of pigs persistently infected with porcine reproductive and respiratory syndrome virus. J. Clin. Microbiol. 41, 58–62.

Journal of Virological Methods 202 (2014) 79–86

Contents lists available at ScienceDirect

Journal of Virological Methods

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

The impact of porcine reproductive and respiratory syndrome virus genetic heterogeneity on molecular assay performances

Michele Drigo[a,*], Giovanni Franzo[a], Alessandra Gigli[b], Marco Martini[a], Alessandra Mondin[a], Patrice Gracieux[c], Letizia Ceglie[b]

[a] Department of Animal Medicine, Production and Health (MAPS), Viale dell'Università 16, 35020 Legnaro, PD, Italy
[b] Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università 10, 35020 Legnaro, PD, Italy
[c] ADIAGENE-bioMerieux, 38 rue de Paris, 22000 Saint Brieuc, France

Article history:
Received 7 August 2013
Received in revised form 7 February 2014
Accepted 4 March 2014
Available online 15 March 2014

Keywords:
PRRSV
ORF7
RT-PCR
In-house
Commercial
Italy

A B S T R A C T

The remarkable economic losses due to porcine reproductive and respiratory syndrome (PRRS) have stated the control and eradication of this disease is one of the main issues of swine modern farming. The limited cross-protection of vaccine-induced immunity compelled the adoption of strict biosecurity measures that must be associated with the prompt diagnosis of infection. In our study four RT-PCR methods, a RT-PCR, a SYBR Green I and two hydrolysis probes, were compared to evaluate their respective benefits and disadvantages. One hundred and seventy samples originating from 50 farms located in northern Italy were tested with all assays and performances were evaluated using a Bayesian approach to deal with the absence of a Gold Standard. Sequencing the complete of ORF7, the segment targeted by all methods, allowed a gain of insight into the genetic variability of Italian strains and to investigate the role of mismatches on assay sensitivity. Our study evidenced that methods based only on primers-genome interaction better tolerate PRRSV genetic variability, demonstrating a greater sensitivity (Se): SYBR Green I (Se = 98.4%) and RT-PCR (Se = 99%) outperform both in-house (Se = 71.4%) and commercial (Se = 91.7%) probe-based methods. On the other hand, probe-based assays allowed an easier genotyping of PRRSV strains and implementation of the internal control system (IC). Phylogenetic analysis allowed demonstration of a presence of two clades circulating continuously in northern Italy since 1996, when their probable ancestors were collected.

1. Introduction

Porcine reproductive and respiratory syndrome virus (PRRSV) emerged as a disease of swine in the late 1980s. Although more than 20 years have elapsed since its discovery, PRRS is still the most prevalent swine disease, with a huge economic impact (Lunney et al., 2010). Great genetic heterogeneity has been demonstrated among different strains. Two main genotypes (type I and II) have been identified sharing only 50–70% nucleotides and 50–80% similarity of amino acids (Forsberg, 2005). In addition, mean nucleotide diversity within the genotype has been estimated to be about 12.5–15%, while a maximum genetic distance of 21–30% has been reported within genotype I and II, respectively (Cho and Dee, 2006; Murtaugh et al., 2010; Pesch et al., 2005; Shi et al., 2010a, 2010b, 2013). The persistent nature of infection, coupled with the limited efficacy of vaccines, has made the control of PRRS particularly problematic (Chand et al., 2012). Avoiding the introduction and minimizing vertical and horizontal spread of the virus within the farm, play a major role in control and eradication of PRRS (Rowland and Morrison, 2012). In order for this strategy to be effective, there needs to be accurate diagnostic tools for identification of infected herds or animals (Corzo et al., 2010; Thanawongnuwech and Suradhat, 2010). Several methods have been validated to detect PRRSV infection. Although serology represents a popular choice, this method is disadvantageous because immune response is highly variable and, even if such antibody titer may fall rapidly after infection, it is not possible to easily discriminate between present vs. past infection or active vs. passive immunity. Moreover, negative or low positive ELISA results do not rule out persistent infection (Batista et al., 2004). Virus isolation is time-consuming and requires a certain expertise. Immunohistochemistry or immunofluorescence assays are used for antigen detection but their sensitivity is still a concern (Martínez et al., 2008). PCR and real time PCR have been used widely for viral infection diagnosis and genotype

* Corresponding author. Tel.: +39 049 827 2968; fax: +39 049 827 2973.
E-mail address: michele.drigo@unipd.it (M. Drigo).

http://dx.doi.org/10.1016/j.jviromet.2014.03.006
0166-0934/

identification (Belák, 2007). This success is due to their usual enhanced sensitivity, larger dynamic range and reduced risk of cross contamination. The use of a specific probe facilitates an increased specificity compared to conventional agarose gel-based PCR assays. Besides the development of multi-color real-time PCR cyclers, it has also been made possible to combine several assays within a single tube. This allows the simultaneous detection and discrimination between different pathogens as well as genotypes of the same microorganism, while providing an easy implementation of internal control (IC) systems (Gunson et al., 2008; Hoffmann et al., 2006; Wu et al., 2008; Zheng et al., 2013). Several commercial diagnostic tests are also available, making it possible to avoid the complex phase of in-house method validation. A major challenge in RT–PCR and real time RT-PCR is represented by problems in designing specific primers/probes able to deal with the PRRSV genetic variability (Hoffmann et al., 2009). This issue is relevant particularly for commercial kits sold on a large scale. However, national and international animal trade is exposed constantly to new strains, imposing on local laboratories to update continuously. Therefore, it is not surprising that the lack of sensitivity in diagnosis of PRRSV infection of both in-house and commercial kits has been reported by previous studies and laboratory experiences (Toplak et al., 2012; Wernike et al., 2012).

The aim of this study is to compare performances, advantages and disadvantages of different in-house and commercial RT-PCR and real time RT-PCR methods in detecting PRRSV from samples collected from a large area in northern Italy. The sequence of field samples has been performed to gain insight into the genetic heterogeneity of Italian viruses and to explore its relation with assay sensitivity.

2. Material and methods

2.1. Field samples

A total of 170 aliquots originating from as many pigs (80 lungs and 90 sera) were selected on the basis of diagnostic activity results obtained using the RT-PCR method (see below) from stored samples delivered during the 2010–2012 time period to Istituto Zooprofilattico delle Venezie (IZSVe). One hundred and fifty-one positive and 19 negative samples have been included in the study to challenge the assays with a broader spectrum of PRRSV strains. The samples originated from 50 farms located in three regions of northern Italy, with unknown previous sanitary status.

Two hundred microliters of serum samples were extracted using High Pure Viral RNA Kit (Roche Diagnostics, Monza, Italy). Lung samples were homogenized after an addition of 10 ml of PBS for each gram of tissue. Two hundred microliters of homogenate were extracted using High Pure RNA Tissue Kit (Roche Diagnostics, Monza, Italy). Before extraction, each aliquot was added with 2 μl of solution containing 2×10^5 copies/μl of RNA Internal Control (IC) (Hoffmann et al., 2006) immediately following an addition of lysis buffer. The same extract was subdivided in four aliquots for successive testing. All samples and RNA aliquots were stored at −80 °C until processing.

2.2. Test samples with RT-PCR and real time RT-PCR

Each aliquot was tested with each of four RT-PCR-based methods. Two in-house real time RT-PCR assays (here defined as Probe and SYBR) described by Drigo et al. (2014) were used with minor modification. Briefly, for TaqMan-based One-Step qRT-PCR 2 μl of extracted RNA were added to a standard reaction mix containing 0.2 μl of SuperScript® III RT/Platinum®Taq Mix (Life Technologies™, Monza, Italy), 0.3 μM of primer PRRSVf1 and PRRSVr2, 0.4 μM of primer EGBP-1F and EGBP-2R, 0.2 μM of PRRSV specific probes (EU-1, EU-2, US) and 0.4 μM of EGBPp. Sterile nanopure water was added to bring the final volume to 10 μl. Cycling parameters were 50 °C for 15 min, 95 °C for 2 min, 40 cycle of 95 °C for 10 s and 60 °C for 30 s. The fluorescence signal was acquired for each cycle at the end of the extension phase. Similarly, for SYBR Green One-Step qRT-PCR 2 μl of extracted RNA were added to a standard reaction mix containing 0.2 μl of SuperScript® III RT/Platinum® Taq Mix (Life Technologies™, Monza, Italy), 0.4 μM of primer PRRSVf1 and PRRSVr2, 0.1 μM of primer EGBP-1F and EGBP-2R. Sterile nanopure water was added to bring the final volume to 10 μl. Cycling parameters were 50 °C for 15 min, 95 °C for 5 min, 40 cycles of 95 °C for 10 s and 60 °C for 30 s. The fluorescence signal was acquired for each cycle at the end of the extension phase. After incubation for 1 min at 40 °C, melting curves were performed by gradually raising temperatures from 64 °C to 95 °C with a continuous collection of fluorescence data.

A commercial real time RT-PCR kit (ADIAVET™), was included in the study in order to determine its diagnostic sensitivity with regard to the highly variable “Italian cluster” (Shi et al., 2010a).

In-house and commercial real time RT-PCRs were performed on a LightCycler®480 system (Roche, Monza, Italy) and 7900HT Fast Real Time PCR System (Life Technologies™, Monza, Italy), respectively.

All samples were tested again with RT-PCR (Persia et al., 2001) to avoid a bias in sensitivity and specificity estimate during assay comparisons due to time-dependent RNA degradation. Also in this case, genotyping was possible due to the use of specific primers resulting in different amplicon lengths. Specific primer and probes are summarized in Table 1.

2.3. Sequences and phylogenetic analysis

The complete ORF7 of all positive samples to at least one of the methods defined previously was amplified as described by Oleksiewicz et al. (1998). Both strands of each amplicon were sequenced using the Byg Dye terminator v3.1 sequencing kit.

Table 1
List of primers and probes used for in-house developed assays.

Primer/Probe	Oligonucleotides	Assay	Reference
RT	5′-TCGCCCTAAT-3′	RT-PCR	Persia et al. (2001)
US/EU-F	5′-ATGGCCAGCCAGTCAATC-3′	RT-PCR	Persia et al. (2001)
EU-R	5′-GATTGCAAGCAGAGGGAGCGTTC-3′	RT-PCR	Persia et al. (2001)
US-R	5′-GGCGCACAGTATGATGCGTAG-3′	RT-PCR	Persia et al. (2001)
PRRSVf1	5′-GGGGAATGGCCAGYCAGTCAA-3′	qRT-PCR	Lurchachaiwong et al. (2008)
PRRSVr2	5′-GCCAGRGGAAAATGKGGCTTCTC-3′	qRT-PCR	Lurchachaiwong et al. (2008)
US	5′-HEX-CTGGGTAAGATCATCGCCCAGGA-3′-IABkFQ	qRT-PCR	Lurchachaiwong et al. (2008)
EU-1	5′-FAM-TTGGCCTGTCCTCCCCTAGGTTG-3′-IABKFQ	qRT-PCR	Lurchachaiwong et al. (2008)
EU-2	5′-FAM-ATGATGAAATCCCAGCGCCAGCGGT-3′-IABkFQ	qRT-PCR	Present study
EGFP-1-F	5′-GACCACTACCAGCAGAACAC-3′	IC System	Hoffmann et al. (2006)
EGFP-2-R	5′-GAACTCCAGCAGGACCATG-3′	IC System	Hoffmann et al. (2006)
EGFP-Cy5	5′-Cy5-AGCACCCAGTCCGCCCTGAGCA-IBRQ-3′	IC System	Hoffmann et al. (2006)

Sequences were obtained using the ABI PRISM® 3100 Genetic Analyzer (Applied Biosystem®, Monza, Italy). Chromatograms were evaluated with FinchTV (http://www.geospiza.com) and consensus sequences were obtained using CromasPro (CromasPro Version 1.5). Phylogenetic analysis was performed on all obtained sequences plus complete and partial ORF7 sequences published by Forsberg et al. (2002) (Italian sequences), Pesente et al. (2006), Stadejek et al. (2008) and Toplak et al. (2012). Others sequences obtained from Pubmed, a representative for several countries, were also included (see supplementary data). All sequences obtained from literature were renamed with their accession number followed by country where they were collected. Only sequences obtained from Stadejek maintained their original numeration. Multiple sequence alignment was carried out using the software TranslatorX (Abascal et al., 2010). The alignment robustness was evaluated using Guidance (Penn et al., 2010). In both cases, MAFFT was chosen as a multiple alignment algorithm. Jmodeltest 2.1.2 (Darriba et al., 2012) was used to select model of evolution considering results of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Phylogenetic tree comprising all sequences was reconstructed using MrBayes 3.2.0 (Ronquist et al., 2012) assuming a general time-reversible nucleotide substitution model with four categories of gamma-distributed rate heterogeneity and a proportion of invariant sites (GTR + Γ4 + I).

Phylogenetic trees for separate clades were drawn using the Maximum Likelihood method implemented in PhyML 3.0 (Guindon et al., 2010) assuming the GTR + Γ4 + I nucleotide substitution model. Phylogenetic tree reliability was evaluated using a fast non-parametric version of the aLRT (Shimodaira–Hasegawa [SH]-aLRT), which was developed and implemented in the PhyML 3.0 (Anisimova et al., 2011).

The number of mismatches between primers and probes of in-house methods and sequences was calculated using Mega5 (Tamura et al., 2011). To evaluate the relation between the distance from Lelystad virus and diagnostic performances, a T-test was performed, comparing mean number of differences between a reference strain (M96262.2) and our sequences of positive/negative samples.

2.4. Assay comparison

As no method could be assumed as Gold standard, sensitivity (Se) and specificity (Sp) of all four assays were estimated simultaneously adopting the Bayesian approach implemented in BayesLatentClassModels (BLCM) software (Dendukuri et al., 2009; Joseph et al., 1995), posing True infectious status as the latent variable. A preliminary run was performed setting the poorly informative priors of Se, Sp and prevalence, chosen on the basis of raw data analysis. After a burn-in of 1000 iterations, the number of monitored iterations of the Gibbs Sampler was set to 20,000. Considering the stochastic nature of the Gibbs sampler algorithm, to ensure the independence of convergence values from the starting point, several runs were performed with different starting values of Se and Sp. The final results were obtained setting prior and initial values with the outcomes of previous runs. Identical protocol was used to estimate sensitivity and specificity at herd level. A herd was assumed to be positive to a particular test if it delivered at least one positive sample during the time window considered. Values are reported in terms of mean, median and credible interval (CI) 95%. Predictive values for each latent class (i.e. positive and negative infectious status) corresponding to different combinations of test results are also reported.

All tests were also evaluated through participation in a proficiency test (GD Animal Health PRRSV proficiency test 2012, Deventer, Netherlands); provided samples comprised of two negative and ten positive samples, including both field and vaccine strains of both genotypes.

Table 2
Counts at sample and herd level for each possible assays combination.

RT-PCR	Probe	SYBR	ADIAVET™	Sample[a]	Herd[b]
+	+	+	+	87	37
+	+	+	–	9	0
+	+	–	+	0	0
+	+	–	–	0	0
+	–	+	+	35	5
+	–	+	–	2	1
+	–	–	+	3	1
+	–	–	–	1	0
–	+	+	+	1	0
–	+	+	–	0	0
–	+	–	+	0	0
–	+	–	–	0	0
–	–	+	+	0	0
–	–	+	–	0	0
–	–	–	+	7	2
–	–	–	–	25	4
Total				170	50

[a] Diagnostic result at sample level with respect to assay combination.
[b] Diagnostic result at herd level with respect to assay combination.

3. Results

3.1. Samples and diagnostic performances

Of the 170 field samples 137 were confirmed PRRSV positive by RT-PCR, 97 by Probe method, 134 by SYBR and 133 by ADIAVET™ kit. At herd level, 44, 37, 43 and 45 farms were classified as positive using RT-PCR, Probe, SYBR and ADIAVET™ kit, respectively. The counts of different result combinations are briefly reported in Table 2.

Bayesian analysis was proven consistent as demonstrated by the rapid convergence to the same parameter estimate, independently from initial values. Sensitivity was estimated very high for RT-PCR (99%) and SYBR (98.4%) methods. The probe assay exhibited poor performances (71.4%). The ADIAVET™ kit obtained intermediate results (91.7%). Specificity was considerably high for all in-house methods, being 96.2%, 98.7% and 98.8% for RT-PCR, SYBR and Probe methods, respectively. Poor specificity was demonstrated by commercial kit (75%) (Table 3). However, five of the seven samples that tested positive with the ADIAVET™ kit but negative with other methods were weak positives during initial routine diagnostic testing. At herd level, performances resulted substantially better in terms of sensitivity, with particular regard to Probe (85%) and ADIAVET™ (98.2%) methods. Specificity displayed a remarkable decrease in all tests considered (Table 3). Ring test results depose in favor of the high specificity of all assays (100%; two of two negative). Sensitivity was also proven high for RT-PCR (100%) and ADIAVET™ (90%) while SYBR and Probe were able to detect 80% of positive samples. All genotypes were correctly classified with one exception in which the sample distinction based on Tm was doubtful.

3.2. Phylogenetic analysis

In total, 97 sequences belonging to the genotype I subtype I were obtained. All ORF7 sequences were 387 nucleotides long. The phylogenetic analysis demonstrates that these are part of two main groups, with only a few exceptions (Fig. 1). The first one defines a clade (Clade A) that comprises only Italian and Slovenian strain samples taken at different times. In addition to our 62 sequences (2010–2012), other Italian strains were collected several years prior. Sequence AY035976 (27_Italy) was sampled in

Table 3
Estimate of sensitivity and specificity at sample and herd level using a Bayesian approach. Results are reported in terms of median, mean and 95% credible interval (CI).

Diagnostic level	Performance	Assay	Mean	Median	95% credible interval	
					Lower bound	Upper bound
Sample	Sensitivity	RT-PCR	0.990	0.991	0.977	0.997
		Probe	0.714	0.714	0.657	0.766
		SYBR	0.984	0.986	0.956	0.998
		ADIAVET™	0.917	0.918	0.882	0.946
	Specificity	RT-PCR	0.962	0.970	0.882	0.997
		Probe	0.988	0.991	0.954	0.999
		SYBR	0.987	0.991	0.952	0.999
		ADIAVET™	0.750	0.752	0.636	0.848
Herd	Sensitivity	RT-PCR	0.978	0.985	0.920	1.000
		Probe	0.850	0.855	0.729	0.943
		SYBR	0.970	0.978	0.898	0.999
		ADIAVET™	0.965	0.973	0.893	0.998
	Specificity	RT-PCR	0.776	0.784	0.530	0.982
		Probe	0.890	0.917	0.644	0.997
		SYBR	0.838	0.861	0.570	0.994
		ADIAVET™	0.667	0.656	0.508	0.886

1996 while 10 samples, whose ORF7 were published by Pesente et al. (2006) were collected between 2002 and 2003. Slovenian sequences (2009–2010), were part of the cluster defined by Toplak et al. (2012) as a new group of PRRSV in Europe, with only a 92% nucleotide identity to the closest related ORF7 sequence. However, according to the results obtained by the same author for ORF5 sequences, strain representatives of this group (e.g. HQ213933) are clearly part of Italian Clade A, sharing among them at least 95% nucleotide identity (e.g. 503). Other Italian samples are part of a larger cluster including sequences collected from several European and non-European countries. Within this group, it is possible to define a second clade (Clade B) composed of closely related Italian strains (Fig. 1). Thirty-three sequences from our study (collected between 2010 and 2012), and two sequences, one published by Pesente (collected in 2003) and the other by Fosberg (collected in 1996), are part of this group. Clades A and B are characterized by relevant mean genetic distance between the groups (p-distance = 0.122). Within group, mean p-distance was 0.083 and 0.041 for group A and B respectively. The other two strains (441 and 2789/22) that did not cluster in these clades were identified (Fig. 1). Strains 441 and 2789/22 share 89.1% and 94.5% nucleotide identity with the Lelystad virus (M96262.2), respectively.

Globally the PRRSV sequenced in this study shared an 88.8–94.5% nucleotide identity with the Lelystad virus. Mean number of differences between Lelystad and our sequences was significantly higher ($p < 0.05$) in positive samples to Probe assay compared to negative ones. Opposite results were obtained for ADIAVET™ kit although the difference was not statistically significant ($p = 0.077$). However, both assays were able to detect strains both similar and distant from Lelystad (p-distance = 88.8–94.5%). Other assays could not be considered due to the absence of sequenced negative samples.

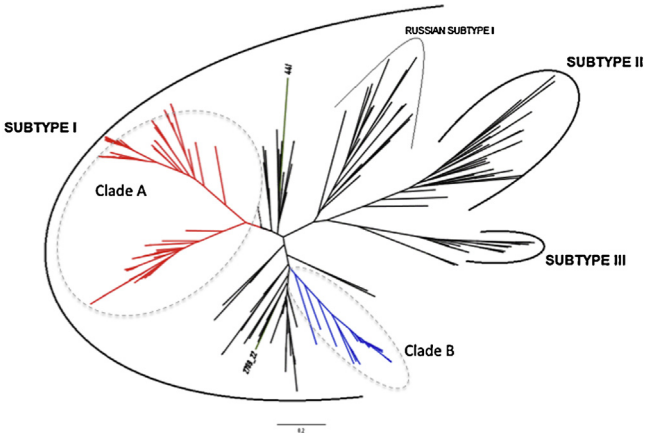

Fig. 1. Overall phylogeny constructed using MrBayes. Clade A and B are emphasized in red and blue, respectively. The names of "out of clade" sequences are displayed.

Relation between genetic diversity and diagnostic sensitivity is reported in Fig. 2a and b. Only three out of 62 sequenced samples belonging to Clade A were negative in Probe. On the other hand, Probe performed poorly in Clade B, where 21 of 33 samples were classified as negative.

The number of mismatches between sequences and primers of in-house methods is comprised between 0 and 3 for Probe and SYBR while it reaches a maximum of 4 differences for RT-PCR. Nonetheless all sequences were obtained from positive samples to both RT-PCR and SYBR.

A remarkable difference was observed in mismatched numbers between the sequence of strains belonging to Clade A or B and Probes EU-1 and EU-2. All strains of both clades displayed at least three to six mismatches compared to Probe EU-1, originally published by Lurchachaiwong et al. (2008), with only three exceptions: samples 282, 1241/33 and 1241/34 that exhibited only one or two mismatches (Fig. 2a and b). A similar conclusion can be drawn for Probe EU-2 with regard to Clade B strains. On the contrary, only 0–2 differences were found for Clade A with only three exceptions: sequences 174/2, 594/62 and 2128/29 (Fig. 2a). However, for these strains, the first mismatch was more than 5 bases apart from the 5′-end of Probe EU-2.

Of the two "out of clade" strains, one (i.e. 441), carrying five mismatches to each probe, resulted negative to Probe method.

4. Discussion

Diagnostic assays based on molecular biology are widely applied in the diagnosis of infectious diseases (Balka et al., 2009; Belák, 2007; Hoffmann et al., 2009). Several in-house methods have been published and many others are commonly used in diagnostic laboratories all over the world. Many commercial kits are also available and are gaining increased favor for practical reasons. However, susceptibility of these techniques to large-scale target variability cannot be underestimated. This issue is particularly true for such a variable virus as PRRS virus.

The aim of this work is to compare the performances of different assays, comprising commercial and in-house methods based on different chemistries. RT-PCR and real time RT-PCR typically display a greater sensitivity than other diagnostic assays hampering the

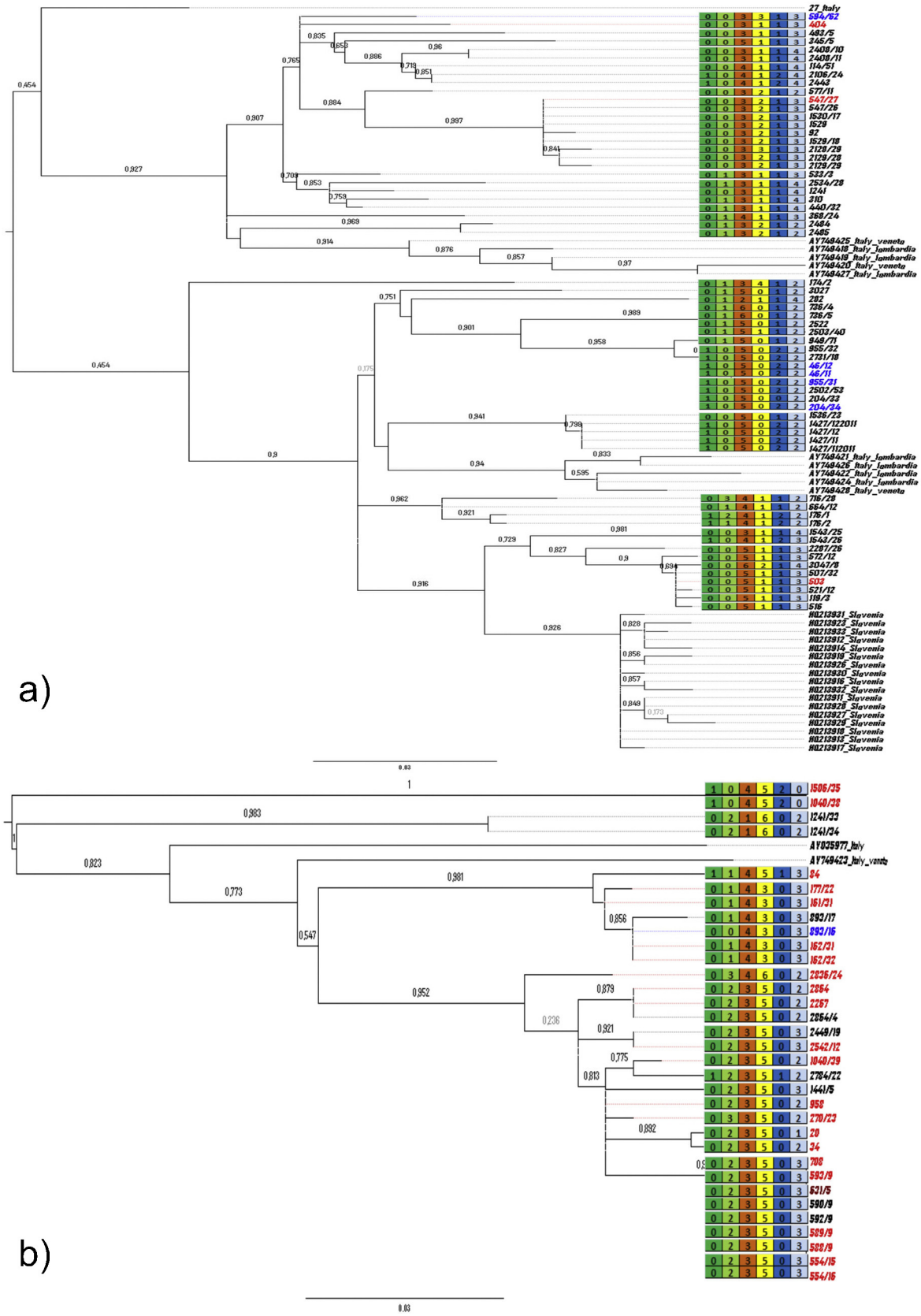

Fig. 2. Phylogenetic trees including all sequences belonging to Clade A (a) and Clade B (b). The number of mismatches between primers and probes used in "in-house methods" and sequences of virus collected in this study, is displayed using the following color coding: dark and light green = qRT-PCR forward and reverse primers; orange = Probe EU-1; yellow = Probe EU-2; dark and light blue = end point RT-PCR forward and reverse primers. Strains name of samples negative to Probe and ADIAVET™ kit are colored in red and blue, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Positive predictive value (PPV) and negative predictive value (NPV) corresponding to each combination of test results at sample and herd level.

Assay results[a]	Sample level				Herd level			
	Mean	Median	95% credible interval		Mean	Median	95% credible interval	
			Lower bound	Upper bound			Lower bound	Upper bound
PPV								
−−−−	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
−−−+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
−−+−	0.18	0.12	0.01	0.68	0.01	0.00	0.00	0.05
−−++	0.77	0.82	0.34	0.99	0.29	0.22	0.01	0.86
−+−−	0.01	0.00	0.00	0.08	0.00	0.00	0.00	0.01
−+−+	0.20	0.13	0.01	0.75	0.10	0.03	0.01	0.60
−++−	0.95	0.98	0.70	1.00	0.33	0.22	0.00	0.96
−+++	1.00	1.00	0.99	1.00	0.94	0.98	0.61	1.00
+−−−	0.11	0.06	0.00	0.54	0.01	0.00	0.00	0.04
+−−+	0.62	0.68	0.08	0.98	0.31	0.27	0.02	0.81
+−+−	1.00	1.00	0.98	1.00	0.67	0.75	0.08	0.99
+−++	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00
++−−	0.88	0.95	0.34	1.00	0.35	0.26	0.01	0.95
++−+	0.99	1.00	0.95	1.00	0.95	0.98	0.69	1.00
+++−	1.00	1.00	1.00	1.00	0.99	1.00	0.92	1.00
++++	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
NPV								
−−−−	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
−−−+	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
−−+−	0.82	0.88	0.32	0.99	0.99	1.00	0.95	1.00
−−++	0.23	0.18	0.01	0.66	0.71	0.78	0.14	0.99
−+−−	0.99	1.00	0.92	1.00	1.00	1.00	0.99	1.00
−+−+	0.80	0.87	0.25	0.99	0.90	0.97	0.40	1.00
−++−	0.05	0.02	0.00	0.30	0.67	0.78	0.04	1.00
−+++	0.00	0.00	0.00	0.01	0.06	0.02	0.00	0.39
+−−−	0.89	0.94	0.46	1.00	0.99	1.00	0.96	1.00
+−−+	0.38	0.32	0.02	0.92	0.69	0.73	0.19	0.98
+−+−	0.00	0.00	0.00	0.02	0.33	0.25	0.01	0.92
+−++	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
++−−	0.12	0.05	0.00	0.66	0.65	0.74	0.05	0.99
++−+	0.01	0.00	0.00	0.05	0.05	0.02	0.00	0.31
+++−	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.08
++++	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

[a] In each line tests are ordered as follow: RT-PCR, Probe, SYBR, ADIAVET™; "+" = positive; "−" = negative.

definition of a proper Gold standard (Mackay, 2007; Shaw et al., 2004). To deal with this issue a Bayesian approach was used, in order to estimate the assays diagnostic performances.

In terms of sensitivity, both at individual and herd level, RT-PCR and SYBR assays, based only on primers-genome interaction, displayed the best performances. In fact these methods seem to be only affected marginally by the presence of a limited number of mismatches: several samples were strongly positive also if there were up to four differences between primer and target sequence. Moreover, the use of degenerate primers as reported in SYBR assay, can clearly improve this aspect. Both TaqMan probe-based methods exhibited a lower sensitivity. As previously described (Balka et al., 2009; Chai et al., 2013), the need of an additional highly specific interaction severely hampers sensitivity of diagnostic tests in the presence of imperfect probes. The Probe assay displayed a remarkable susceptibility to Italian genetic heterogeneity as demonstrated by the poor sensitivity to Clade B strains. Considering that the primers pair was the same for SYBR and Probe assay, this lack of sensitivity is clearly attributable to probes inadequacy (i.e. mismatches number) to circulating viruses.

Although complete information on chemistry are not available, the better performance of ADIAVET™ kit are probably due to the use of several probes–primers combinations and to their continuous updating, as demonstrated by the comparison with the results obtained in a previous study (Toplak et al., 2012). Four out of six samples, negative to ADIAVET™ kit, clustered together suggesting a failure related to the specific viral sequence. However the producer declares the perfect complementarity of a part of diagnostic oligonucleotides to all negative samples. As a consequence, lack of sensitivity could be ascribed to the lack of binding due to causes independent from primary structure (e.g. poor quality template, RNA secondary structure) or to non-optimal probe-primers combinations affecting the efficiency of the PCR reaction. Specificity was optimal for all methods considered with only the exception of ADIAVET™ kit (Sp = 75%). Although it isn't possible to exclude incorrect binding, the triple specific interaction between genome, primers and probe makes this hypothesis unlikely (Hoffmann et al., 2009). Moreover it should be stressed that five out of seven discordant samples were represented by samples that had been previously classified as positive during IZSVe routine diagnostic activity. It is plausible that RNA degradation affected sensitivity of other diagnostic assays while ADIAVET™ kit was still able to detect PRRSV genome in these samples due to lower limit of detection (LoD), smaller amplicon size or other undetermined causes. Although no Gold Standard was assumed, agreement in lack of sensitivity by all other methods, can determine a bias also using a Bayesian approach.

Herd level performances display an increased sensitivity compared to single sample level. This aggregate measure is more affected by assays susceptibility to genetic heterogeneity, than by analytic sensitivity (i.e. LoD). Actually, different samples of the same farm, most likely with a different viral load, were considered as one. These results confirm the aforementioned hypothesis that in the presence of high viral load even samples with several mismatches can be detected. Similarly, it can also be inferred that imperfect oligonucleotide interaction, affected also by minor

genetic distance (e.g. *intra*-clade diversity), coupled with other causes of assay variability as reaction stochasticity, operator and machinery variation, can widely affect reaction efficiency, complicating the comparison between different methods as well as outcomes of the same test. As a consequence, particular caution should be used when interpreting quantification of PRRSV titer. When attempting to control/eradicate the disease or to maintain a free status, herd diagnosis is extremely relevant in managing commercial farms exchanges (Corzo et al., 2010). This study reveals that farm monitoring over time determined a greater sensitivity even using poor performance testing. As a consequence, commercial relations between farms should be founded on knowledge of sanitary status obtained through long-term virological surveillance, possibly coupled with clinical and serological evidences. Our study also reveals that multiple testing strategies, with different assays, determines a great increase in diagnostic accuracy (Table 4), as reported by several authors (King et al., 2006; Wernike et al., 2012).

Despite the fact that RT-PCR and SYBR assays revealed better performances in both single and herd level, they also hide some weaknesses: no IC system, necessary for ensuring adequate efficiency of RNA extraction and confirming the absence of PCR-inhibitors in each sample, was used in the RT-PCR assay evaluated. Genotyping was proven difficult and somehow subjective using SYBR, due to significant Tm variation among strains. Besides, IC detection in SYBR was relatively inconstant in positive samples due to limiting primer concentrations, making it difficult to account partial inhibition. On the contrary, the use of specifically labeled probes differentially allowed an easier discrimination between genotype I and II and implementation of an IC system.

Although genetic distance to the Lelystad virus is often reported as aggregated information by several diagnostic laboratories, during this study it was possible to demonstrate that it was an extremely poor indicator in predicting diagnostic test performance.

Phylogenetic analysis revealed the presence, within the genotype I subtype I, of two main clades, characterized by a relevant genetic distance. In both cases clusters have included also sequences collected several years earlier (1996 and 2002–2003). So it is possible to suppose that some of the strains that were present before 1996 gave rise to two independent clusters that have continuously circulated and evolved in northern Italy, gaining a significant intra-clade diversity, while the other became extinct. Furthermore, the presence of "out of clade" Italian sequences could be justified by the recent introduction of new strains from other countries through swine or semen trade (Corzo et al., 2010; Lambert et al., 2012; Rowland and Morrison, 2012). However, an underlying not-sampled genetic diversity cannot be excluded, as suggested by the strong phylogenetic relationship between strain 441 (collected in 2012) and two strains (i.e. AY035978 and AY035971) collected in Italy, one in 1998, the other unknown (data not shown). Their epidemiological evolution and interaction with endemic clade will need further investigation.

Moreover, a strong relation between the "new" Slovenian isolates (Toplak et al., 2012) and Clade A Italian strains was reported. Considering the prolonged circulation of these strains in Italy mentioned previously, and remembering that Slovenia was PRRSV free until 2004 (Valencak, 2004), it is highly probable that this Slovenian clade originated from Italian isolates, experiencing then an independent evolution.

5. Conclusion

This study reports a greater sensitivity of assays based only on primers/genome interaction. However, it would be reductive to discharge TaqMan-based real time RT-PCR as a diagnostic option due to lack of sensitivity. Each test examined in this study had some intrinsic disadvantages: the need of post-PCR manipulation and the absence of an internal control (RT-PCR), the difficulty to discriminate between genotypes due to variable Tm (SYBR), and the susceptibility to minor genomic mutations (TaqMan-based methods). This last issue is relevant particularly for commercial kits, aimed at being distributed on a large, international scale. All these factors and many others should be considered by every laboratory on the basis of peculiar expertise, instrumentation and knowledge of field epidemiologic condition. Besides, a continuous devotion should be reserved to the study of viral genetic diversity and to the sharing of related information. As demonstrated in our study, as well as Toplak's one (Toplak et al., 2012), effective collaboration between public and private enterprises on an international scale has revealed an invaluable approach to develop and easily upgrade accessible tools for controlling and taming this fastidious disease. At the same time, field veterinarians, regretting any uncritical choice, should keep up-to-date on the diagnostic techniques and their limits, potentialities and applications, allowing a problem oriented choice.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jviromet.2014.03.006.

References

Abascal, F., Zardoya, R., Telford, M.J., 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 38, W7–W13.

Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C., Gascuel, O., 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst. Biol. 60, 685–699.

Balka, G., Hornyák, A., Bálint, A., Benyeda, Z., Rusvai, M., 2009. Development of a one-step real-time quantitative PCR assay based on primer–probe energy transfer for the detection of porcine reproductive and respiratory syndrome virus. J. Virol. Methods 158, 41–45.

Batista, L., Dee, S.A., Rossow, K.D., Polson, D.D., Xiao, Z., Olin, M., Murtaugh, M.P., Molitor, T.W., Joo, H.S., Pijoan, C., 2004. Detection of porcine reproductive and respiratory syndrome virus in pigs with low positive or negative ELISA s/p ratios. Vet. Rec. 154, 25.

Belák, S., 2007. Molecular diagnosis of viral diseases, present trends and future aspects. A view from the OIE Collaborating Centre for the Application of Polymerase Chain Reaction Methods for Diagnosis of Viral Diseases in Veterinary Medicine. Vaccine 25, 5444–5452.

Chai, Z., Ma, W., Fu, F., Lang, Y., Wang, W., Tong, G., Liu, Q., Cai, X., Li, X., 2013. A SYBR Green-based real-time RT-PCR assay for simple and rapid detection and differentiation of highly pathogenic and classical type 2 porcine reproductive and respiratory syndrome virus circulating in China. Arch. Virol. 158, 407–415.

Chand, R.J., Trible, B.R., Rowland, R.R., 2012. Pathogenesis of porcine reproductive and respiratory syndrome virus. Curr. Opin. Virol. 2, 256–263.

Cho, J.G., Dee, S.A., 2006. Porcine reproductive and respiratory syndrome virus. Theriogenology 66, 655–662.

Corzo, C.A., Mondaca, E., Wayne, S., Torremorell, M., Dee, S., Davies, P., Morrison, R.B., 2010. Control and elimination of porcine reproductive and respiratory syndrome virus. Virus Res. 154, 185–192.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.

Dendukuri, N., Hadgu, A., Wang, L., 2009. Modeling conditional dependence between diagnostic tests: a multiple latent variable model. Stat. Med. 28, 441–461.

Drigo, M., Franzo, G., Belfanti, I., Martini, M., Mondin, M., Ceglie, L., 2014. Validation and comparison of different end point and real time RT-PCR assays for detection and genotyping of porcine reproductive and respiratory syndrome virus. J. Virol. Methods 201, 79–85.

Forsberg, R., 2005. Divergence time of porcine reproductive and respiratory syndrome virus subtypes. Mol. Biol. Evol. 22, 2131–2134.

Forsberg, R., Storgaard, T., Nielsen, H.S., Oleksiewicz, M.B., Cordioli, P., Sala, G., Hein, J., Bøtner, A., 2002. The genetic diversity of European type PRRSV is similar to that of the North American type but is geographically skewed within Europe. Virology 299, 38–47.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Gunson, R.N., Bennett, S., Maclean, A., Carman, W.F., 2008. Using multiplex real time PCR in order to streamline a routine diagnostic service. J. Clin. Virol. 43, 372–375.

Hoffmann, B., Beer, M., Reid, S.M., Mertens, P., Oura, C.A.L., van Rijn, P.A., Slomka, M.J., Banks, J., Brown, I.H., Alexander, D.J., King, D.P., 2009. A review of RT-PCR technologies used in veterinary virology and disease control: sensitive and specific diagnosis of five livestock diseases notifiable to the World Organisation for Animal Health. Vet. Microbiol. 139, 1–23.

Hoffmann, B., Depner, K., Schirrmeier, H., Beer, M., 2006. A universal heterologous internal control system for duplex real-time RT-PCR assays used in a detection system for pestiviruses. J. Virol. Methods 136, 200–209.

Joseph, L., Gyorkos, T.W., Coupal, L., 1995. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. Am. J. Epidemiol. 141, 263–272.

King, D.P., Ferris, N.P., Shaw, A.E., Reid, S.M., Hutchings, G.H., Giuffre, A.C., Robida, J.M., Callahan, J.D., Nelson, W.M., Beckham, T.R., 2006. Detection of foot-and-mouth disease virus: comparative diagnostic sensitivity of two independent real-time reverse transcription-polymerase chain reaction assays. J. Vet. Diagn. Invest. 18, 93–97.

Lambert, M.-È., Arsenault, J., Poljak, Z., D'Allaire, S., 2012. Correlation among genetic, Euclidean, temporal, and herd ownership distances of porcine reproductive and respiratory syndrome virus strains in Quebec, Canada. BMC Vet. Res. 8, 76.

Lunney, J.K., Benfield, D.A., Rowland, R.R.R., 2010. Porcine reproductive and respiratory syndrome virus: an update on an emerging and re-emerging viral disease of swine. Virus Res. 154, 1–6.

Lurchachaiwong, W., Payungporn, S., Srisatidnarakul, U., Mungkundar, C., Theamboonlers, A., Poovorawan, Y., 2008. Rapid detection and strain identification of porcine reproductive and respiratory syndrome virus (PRRSV) by real-time RT-PCR. Lett. Appl. Microbiol. 46, 55–60.

Mackay, I.M., 2007. Real-Time PCR in Microbiology: From Diagnosis to Characterization. Caister Academic Press, Norfolk.

Martínez, E., Riera, P., Sitjà, M., Fang, Y., Oliveira, S., Maldonado, J., 2008. Simultaneous detection and genotyping of porcine reproductive and respiratory syndrome virus (PRRSV) by real-time RT-PCR and amplicon melting curve analysis using SYBR Green. Res. Vet. Sci. 85, 184–193.

Murtaugh, M.P., Stadejek, T., Abrahante, J.E., Lam, T.T.Y., Leung, F.C., 2010. The ever-expanding diversity of porcine reproductive and respiratory syndrome virus. Virus Res. 154, 18–30.

Oleksiewicz, M.B., Bøtner, A., Madsen, K.G., Storgaard, T., 1998. Sensitive detection and typing of porcine reproductive and respiratory syndrome virus by RT-PCR amplification of whole viral genes. Vet. Microbiol. 64, 7–22.

Penn, O., Privman, E., Ashkenazy, H., Landan, G., Graur, D., Pupko, T., 2010. GUIDANCE: a web server for assessing alignment confidence scores. Nucleic Acids Res. 38, W23–W28.

Persia, D., Pacciarini, M., Cordioli, P.S., 2001. Evaluation of three RT-PCR assays for the detection of porcine and respiratory syndrome virus (PRRSV) in diagnostic samples. In: Proceedings of the 10th International Symposium of Veterinary Laboratory Diagnosticians and OIE Seminar on Biotechnology, pp. 440–441.

Pesch, S., Meyer, C., Ohlinger, V.F., 2005. New insights into the genetic diversity of European porcine reproductive and respiratory syndrome virus (PRRSV). Vet. Microbiol. 107, 31–48.

Pesente, P., Rebonato, V., Sandri, G., Giovanardi, D., Ruffoni, L.S., Torriani, S., 2006. Phylogenetic analysis of ORF5 and ORF7 sequences of porcine reproductive and respiratory syndrome virus (PRRSV) from PRRS-positive Italian farms: a showcase for PRRSV epidemiology and its consequences on farm management. Vet. Microbiol. 114, 214–224.

Ronquist, F., Teslenko, M., Van, D.M., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. Mrbayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.

Rowland, R.R., Morrison, R.B., 2012. Challenges and opportunities for the control and elimination of porcine reproductive and respiratory syndrome virus. Transbound. Emerg. Dis. 59, 55–59.

Shaw, A.E., Reid, S.M., King, D.P., Hutchings, G.H., Ferris, N.P., 2004. Enhanced laboratory diagnosis of foot and mouth disease by real-time polymerase chain reaction. Rev. Sci. Tech. 23, 1003–1009.

Shi, M., Holmes, E.C., Brar, M.S., Leung, F.C., 2013. Recombination is associated with an outbreak of novel highly pathogenic porcine reproductive and respiratory syndrome viruses in china. J. Virol. 87, 10904–10907.

Shi, M., Lam, T.T., Hon, C.C., Hui, R.K., Faaberg, K.S., Wennblom, T., Murtaugh, M.P., Stadejek, T., Leung, F.C., 2010a. Molecular epidemiology of PRRSV: a phylogenetic perspective. Virus Res. 154, 7–17.

Shi, M., Lam, T.T., Hon, C.C., Murtaugh, M.P., Davies, P.R., Hui, R.K., Li, J., Wong, L.T., Yip, C.W., Jiang, J.W., Leung, F.C., 2010b. Phylogeny-based evolutionary, demographical, and geographical dissection of North American type 2 porcine reproductive and respiratory syndrome viruses. J. Virol. 84, 8700–8711.

Stadejek, T., Oleksiewicz, M.B., Scherbakov, A.V., Timina, A.M., Krabbe, J.S., Chabros, K., Potapchuk, D., 2008. Definition of subtypes in the European genotype of porcine reproductive and respiratory syndrome virus: nucleocapsid characteristics and geographical distribution in Europe. Arch. Virol. 153, 1479–1488.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.

Thanawongnuwech, R., Suradhat, S., 2010. Taming PRRSV: revisiting the control strategies and vaccine design. Virus Res. 154, 133–140.

Toplak, I., Rihtarič, D., Hostnik, P., Grom, J., Štukelj, M., Valenčak, Z., 2012. Identification of a genetically diverse sequence of porcine reproductive and respiratory syndrome virus in Slovenia and the impact on the sensitivity of four molecular tests. J. Virol. Methods 179, 51–56.

Valencak, Z., 2004. Porcine reproductive and respiratory syndrome (PRRS) in Slovenia: evaluation of serology. Slov. Vet. Res. 41, 99–101.

Wernike, K., Bonilauri, P., Dauber, M., Errington, J., LeBlanc, N., Revilla-Fernández, S., Hjulsager, C., Isaksson, M., Stadejek, T., Beer, M., Hoffmann, B., 2012. Porcine reproductive and respiratory syndrome virus: interlaboratory ring trial to evaluate real-time reverse transcription polymerase chain reaction detection methods. J. Vet. Diagn. Invest. 24, 855–866.

Wu, C., Cheng, X., He, J., Lv, X., Wang, J., Deng, R., Long, Q., Wang, X., 2008. A multiplex real-time RT-PCR for detection and identification of influenza virus types A and B and subtypes H5 and N1. J. Virol. Methods 148, 81–88.

Zheng, L.L., Wang, Y.B., Li, M.F., Chen, H.Y., Guo, X.P., Geng, J.W., Wang, Z.Y., Wei, Z.Y., Cui, B.A., 2013. Simultaneous detection of porcine parvovirus and porcine circovirus type 2 by duplex real-time PCR and amplicon melting curve analysis using SYBR Green. J. Virol. Methods 187, 15–19.

International trades, local spread and viral evolution: the case of Porcine circovirus type 2 (PCV2) strains heterogeneity in Italy

Submitted to Veterinary Microbiology

Franzo Giovanni[1*#], Tucciarone Claudia[1*], Dotto Giorgia[1], Gigli Alessandra[2], Ceglie Letizia[2], Drigo Michele[1].

[1]Department of Animal Medicine, Production and Health (MAPS), Viale dell'Università 16, 35020 Legnaro (PD), Italy;

[2]Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università 10, 35020 Legnaro (PD), Italy.

*Equally contributing authors

[#]Corresponding author. Tel: +390498272968; fax: +390498272973;

email address: giovanni.franzo@studenti.unipd.it

Abstract

Porcine circovirus type 2 is one of the most widespread and economically relevant infections of swine. Four genotypes have been recognized, but currently, only three (PCV2a, PCV2b and PCV2d) are effectively circulating. The widespread livestock trade and rapid viral evolution have contributed to determining the high heterogeneity of PCV2 and the dispersal of potentially more virulent strains. Italian swine farming and the related processing industry are relevant in the national economy. Despite the noteworthy losses associated with direct and control measure costs, no data are currently available on the molecular epidemiology of PCV2 in Italy. Our study, which was intended to fill this gap, considered 75 completed genome PCV2 sequences, which were obtained from samples collected from the highly densely populated area of Northern Italy between 2007 and 2014. Phylogenetic analysis and comparison with reference sequences demonstrated the co-circulation, with different prevalences, of PCV2a, PCV2b and PCV2d within the national borders, with PCV2b being the most prevalent. Recombination between different genotypes was also proven to be frequent. Phylogeographic analysis demonstrated that the marked variability of Italian PCV2 strains can be attributable to multiple introduction events. The comparison of the phylogenetic analysis results, the location of different haplotypes and the international commercial routs of live pigs allow the speculation of

several links as well as the role of Italy as both an importer and exporter of PCV2 haplotypes, mainly from and to European and Asian countries. A similarly intricate contact network was demonstrated within national borders, with different haplotypes being detected in the same province and different provinces harbouring the same haplotype. Overall, this paper represents the first description of PCV2 in Italy and demonstrates that the high variability of circulating Italian strains is due to multiple introduction events, wide circulation within national boundaries and rapid viral evolution.

Keywords

PCV2, trades, evolution, recombination phylogeography, complete genome.

Introduction

Porcine circovirus type 2 (PCV2) is the aetiological agent of one of the most relevant infectious diseases of swine, causing severe economic losses due to clinical and subclinical syndromes and control-associated costs (Alarcon et al., 2013). PCV2 is a small non-enveloped ssDNA virus belonging to the family *Circoviridae* and the genus *Circovirus* and is recognised as the only member of this group to cause disease in mammals. This virus possesses a simple circular ambisense genome ranging from 1766 to 1768 nucleotides and encoding four major proteins. ORF1 encodes the viral replicase proteins (*Rep* and *Rep'*), ORF2 encodes the viral capsid protein (*Cap*), and ORF3 and ORF4 encode proteins involved in the regulation of viral replication and apoptosis (Mankertz et al., 2004; Karuppannan and Kwang, 2011; Gao et al., 2014; Lv et al., 2014). PCV2 was initially associated with a clinical disease that was described for the first time at the beginning of the 1990's (Ellis et al., 1998; Harding et al., 1998). Since then, several clinical conditions, collectively named porcine circovirus diseases (PCVD) have been associated with PCV2 infection, including postweaning multisystemic wasting syndrome (PMWS), porcine dermatitis and nephropathy syndrome (PDNS), reproductive disorders, enteritis, proliferative and necrotising pneumonia (PNP) and porcine respiratory disease complex (PRDC) (Segales. 2012). Since the first reports of PCV2 infection, it has been described throughout the world, and PMWS showed epidemic proportions in Europe and South East Asia by the late 1990's and in the Americas by 2004-05 (Segales et al., 2013). Based on phylogenetic studies, a classification scheme for PCV-2 was proposed which divides the viral strains into four major groups based on genotype: PCV2a, PCV2b, PCV2d and PCV2c (Franzo et al., 2014b). The first three groups have a substantial worldwide distribution (Franzo et al., 2014b),

while the last group was detected in samples from Danish archives only once (Dupont et al., 2008). International swine trade has been reported to play a major role in PCV2 strain dissemination in both developed and developing countries and has the potential to affect evolution of this virus (Firth et al., 2009; Drew. 2011; Vidigal et al., 2012; Segales et al., 2013). In Italy, swine farming is an important sector of the agricultural economy; on average, 9 million pigs are raised in Italy every year (FAOSTAT http://faostat3.fao.org/faostat-gateway/go/to/home/E). Remarkably, Italian pig farming is mainly targeted towards the production of Italian cured meats, which requires animals to be fattened until, at least, 9 months of age. This poses a great challenge for the control of infectious diseases because of the frequent presence of animals of different ages and infectious and immune statuses in the same farm. Moreover, Italy is part of a widespread network of international trading of live pigs. In 2013, approximately 800000 and 15500 live swine were imported and exported mainly from and to countries of the European Union (http://comtrade.un.org/data/ and http://www.inea.it/banchedati), respectively. This creates a risky combination of factors for the introduction and spread of viruses, especially those characterised by long viral shedding and typically moderate to absent symptomatology. However, there are currently no data available on PCV2 genotypes circulating in Italy and their genetic variability. Our study intended to fill this gap in knowledge; therefore, the full genomes of samples collected in Italy since 2007 were sequenced to investigate the spreading dynamics of PCV2 within national borders and to compare the strains from Italy with those of foreign countries.

Materials and Methods

Samples and DNA extraction

Public and private diagnostic laboratories provided a total of 96 archive samples (lungs, lymph nodes and sera) that were positive for PCV2 by IHC or real time-PCR. After diagnosis, all samples were stored at -20°C until DNA extraction. Before DNA extraction, each gram of tissue was homogenised in 5 mL of PBS. DNA was extracted from 200 µL of homogenate or 200 µL of serum using the DNeasy® Blood and Tissue kit (Quiagen) according to the manufacturer's instructions. Samples and extracted DNA were stored at -20°C until processing.

PCR and sequencing

The complete PCV2 genome was amplified using four overlapping PCRs. Different primer pairs were designed using Primer3 (21). Each primer was evaluated *in silico* for secondary structures and self- and hetero-dimers using OligoAnalyzer 3.1 (http://eu.idtdna.com/analyser/applications/oligoanalyser/). The PCR protocols were validated for each primer pair by testing different primers concentrations and thermal cycle settings to obtain the highest yield and the absence of nonspecific amplicons. All PCRs were performed using a 2720 Thermal Cycler® (Applied Biosystems®) and Platinum®*Taq* DNA *Polymerase* (Life Technologies) in a final reaction volume of 35 µL. Briefly, 5µL of DNA was added to a standard reaction mix of 1X PCR buffer, 0.5 µM of each primer (Table 1), 0.2 mM of each dNTP, 1.5 mM $MgCl_2$ and one unit of Platinum®Taq DNA polymerase. Two thermal protocols were used that differed only in the annealing temperature. After the activation step at 94°C for 2 minutes, amplification was carried out in 45 cycles of 94°C for 30 sec, 60°C (for segment A, C and D) or 65°C (for segment B) for 30 sec and 72°C for 50 sec. A final extension at 72°C for 2 min was also performed. Amplification and specificity of bands were visualised using a SYBR safer stained 2% agarose gel, and all positive PCR products were purified using Nucleospin®Gel and PCR clean-up (Macherey and Nagel GmbH & Co. KG). Amplicons were sequenced with the same primers, in both senses, using the Byg Dye terminator v3.1 sequencing kit. Sequences were obtained using ABI PRISM® 3100 Genetic Analyzer. Chromatograms were evaluated by FinchTV (http://www.geospiza.com), and consensus sequences were reconstructed using CromasPro (CromasPro Version 1.5).

Sequence analysis

Complete genome sequences were aligned using the MUSCLE method implemented in MEGA6. Sequences were then checked for recombination using RDP3. The RDP, GENECONV, MaxChi and 3Seq methods were selected for the primary scan, and all methods implemented were used for the recombination detection refinement. The settings of each method were adjusted to account for the database features following the RDP3 manual recommendations. A recombination event was accepted if it was detected by more than two methods with a significant p-value of $p<0.01$ with Bonferroni's correction.

Italian strains were classified into different genotypes according to the classification proposed by Franzo et al., (2014b). After the exclusion of sequences demonstrated to be recombinant within the Cap gene, Italian ORF2 sequences were aligned with the full

reference database proposed by Franzo et al., (2014b) using the program Translator X. To account for the coding nature of ORF2, the alignment was performed at the amino acid level using the MAFFT method, and then, the nucleotide sequence was superimposed. Phylogenetic reconstruction was accomplished using MrBayes executing two independent metropolis coupled Markov chain Monte Carlo (MCMCMC) runs of 10 million generations, including one cold and three hot chains each. A general time reversible (GTR) substitution model with 4 categories of gamma-distributed rate heterogeneity was selected according to the Akaike information criterion (AIC) score calculated using JModel Test 2.1.2. Parameters and trees, sampled every 1000 generations, were summarised setting the burn-in to 25%. Effective sample size (ESS), average standard deviation of split frequencies and potential scale reduction factor (PSRF) were used to monitor mixing and runs convergence. Phylogenetic trees were also reconstructed with the Neighbor Joining (NJ) method and MEGA6 using the substitution model with the better Bayesian information criterion (BIC) score given by the same MEGA software. The confidence of the internal branches was evaluated performing 1000 bootstrap pseudo-replicates of the aligned dataset. The distance between sequences pairs and between and within genotypes was calculated using MEGA6. To understand the epidemiological link between Italy and other countries, genotype specific haplotype networks were reconstructed using the median joining method implemented in Network 4.6.1.2. For this purpose, sequences were selected only when the collection country was known. In total, 97 PCV2a, 381 PCV2b and 166 PCV2d ORF2 sequences, were downloaded from GenBank (Supplementary table 1) and aligned with sequences obtained in this study according to their genotype. The spread of PCV2 throughout the Italian provinces was evaluated using the same approach.

Results

PCV2 circulation in Italy

Seventy-five complete genome sequences (1767-1768 nt) were obtained in this study. Briefly, sequenced samples were collected from 39 farms located in 11 Italian provinces, covering approximately 30,000 Km^2 of Northern Italy, between November 2007 and January 2014 (Supplementary table II). Based on the phylogenetic analysis of the ORF2 sequences (Supplementary figure I), 3, 56 and 8 sequences were clearly classified as PCV2a, PCV2b and PCV2d, respectively. The two tree reconstruction methods employed

provided concordant results. PCV2b was the most prevalent genotype for the entire duration of the study, with the only exception of 2014, when just one sample (PCV2d) was collected. PCV2a was identified only sporadically between 2008 and 2012, while PCV2d was first sampled in 2010 and was still present in 2014, at the end of the study (Supplementary figure 2).

Ten recombinant strains, grouped in 6 independent recombination events, were identified (Table 2). In 8 of those strains, recombination affected, at least partially, the ORF2 gene. Briefly, 6 recombination breakpoints were predicted within the ORF2 gene, 4 within the ORF1 gene and two in the major inter-genic region. Four out of 6 recombination breakpoints within ORF2 and all breakpoints within ORF1 affected the beginning and the ending of 25% of the genes. Even if the recombination between PCV2b and PCV2d was the most frequent (i.e., 4 out of 6 recombination events), all genotypes were involved in the recombination events and all possible combinations of parental strains were observed. Remarkably, sequences grouped within the same recombination event were sampled in a restricted geographic area and time window (data not shown).

Italian sequences displayed a relevant nucleotide genetic distance for ORF2 and the complete genome (i.e., range of 0-11% and 0-5.6%, respectively), reflecting the circulation of different genotypes. Additionally, within the same genotype, the genetic distance was relevant, and PCV2a displayed the highest mean p-distance and range at both the complete genome and ORF2 level. PCV2b and PCV2d displayed a lower but still significant genetic distance. Similar results were obtained when examining the Cap amino acid sequence (Table 3).

PCV2 Geographic Distribution and International Trade

Median joining networks, compared with phylogenetic trees, allows for the display of the most parsimonious trees contemporaneously and for the display of the frequency of each haplotype. Combining this with the possibility to mark the sequences according to a certain character, allows for an easier interpretation of the structure of PCV2's geographical distribution and spread. Haplotype nomenclature, distribution and frequency are reported in supplementary data II. In Italy, PCV2 was characterised by an intricate structure. PCV2a displayed a clear association between haplotype and provinces, with different provinces displaying distantly related strains. PCV2d was detected in only three provinces, and

closely related strains tended to be located in the same area. PCV2b, the most prevalent genotype, showed a totally different pattern, with different regions harbouring the same haplotype and different haplotypes being present in the same region (Figure 1); however, even for PCV2b, the same or closely related haplotypes were often displayed in the same region. Potential links with other countries were also evaluated. Italian PCV2a strains were part of two main clusters; two strains, 345/2012 (HA_23) and 751/45 (HA_22) belonged to a group including mainly European sequences, while strain 7d (HA_24) was part of a group that included North American and Asian sequences (Figure 2a). Due to the high number and high similarity of available PCV2b sequences, the generated haplotype network was highly intricate. To simplify the interpretation of the haplotype trends, a preliminary analysis was performed on only the haplotypes that displayed more than one sequence. The rationale behind the exclusion of sequences that were unique in the data set was that a simpler network that still displayed the main haplotype trends should be obtainable from groups instead of individuals. Furthermore, a group of identical sequences is less likely to include random errors (e.g., sampling, lab, typing) and low fitness mutants. Using this approach, we demonstrated that several strains sequenced in our work were identical to others circulating in either Europe, Asia or South America (Figure 2b): nineteen PCV2b Italian strains (33,92%) belonged to three widely distributed haplotypes, namely HB_0 (9 sequences) 3HB_1 (3 sequences) and HB_5 (7 sequences). Italian specific haplotypes comprising more than one strain were identified. Four of these haplotypes were closely related to haplotypes isolated in both Italy and foreign countries, while for other 2 related haplotypes were reported only in foreign countries. PCV2d was identified mainly in Asian countries, especially China. Italian PCV2d sequences were closely related to Chinese strains. Sequence 1077/2012 was identical to the most frequent and widespread PCV2d haplotype (HD_0), which includes sequences from China (20), Romania (1), Vietnam (1) and the USA (2). Sequence 961 (HD_14) originated directly from HD_0. Another 2 haplotypes, HD_3 (i.e., sequences 1779/2011, 1780/2011 and 1781/2011) and HD_4 (i.e., sequences 1053 and 2665/2012), included only Italian sequences and seemed to originate from other Chinese haplotypes that were related, but not identical, to the previous haplotype (i.e., HD_0). Another Italian sequence (i.e., 2864/6) constituted a fifth haplotype (HD_13) originating from HD_3 (Figure 3).

Discussion

The main objective of this work was to examine the molecular epidemiology of PCV2 strains circulating in Italy by evaluating the presence of different genotypes and their heterogeneity. All currently recognised genotypes (Franzo et al.,2014b) were identified in Italy, except for PCV2c. The structure of PCV2 genotypes in Italy is characterised by the clear predominance of PCV2b, which is in accordance with reports in other countries, where several authors recognised a genotype shift from PCV2a to PCV2b approximately in 2003 (Grau-Roma et al., 2011; Segales et al., 2013). In fact, the presence of only a few distantly related PCV2a strains indicates that this genotype is losing importance in our country. However, the limited timeframe of our study and the absence of any previous PCV2 epidemiological data, made it impossible to verify the occurrence of a genotype shift in Italy or to verify trends in genotype relationships. PCV2d, the most recently identified genotype, includes sequences that have been detected since 2005. However, the major PCV2d haplotypes, to which the Italian strains are related, were first identified in China in 2007 (Guo et al., 2010), approximately three years prior to the identification of the Italian PCV2d sequences. Our study also showed the presence of frequent recombination events, which is a well-documented phenomenon for ssDNA viruses (Lefeuvre et al., 2009) and for PCV2 in particular (Ma et al., 2007; Hesse et al., 2008; Huang et al., 2013; Franzo et a.,2014b). According to Lefeuvre et al., (2009), recombination mainly affects the inter-genic regions or the periphery of genes. More interestingly, the high recombination frequency observed in Italy, involving all genotypes identified, strongly indicates frequent co-infections. The frequency of this phenomenon in Northern Italy has already been highlighted in a recent study about another widespread swine RNA virus, PRRSV (Franzo et al., 2014a). The extensive pathogen circulation, confirmed by phylogeographic data, probably played a major role in this phenomenon and emphasises the actual limit of biosecurity and biocontainment strategies. Nevertheless, the true role of recombination in PCV2 evolution in Italy is still unknown. All the recombinant strains that originated from a single recombinant event were collected from limited geographic areas during relatively short timeframes (data not shown), which hardly support the occurrence of a relevant fitness gain of these recombinants compared to the parental strains. Still, even if the biological role of this phenomenon needs further studies, it cannot be overlooked from an epidemiological point of view. Actually, recombination may greatly affect strain

genotypization and bias results of epidemiologic studies, potentially resulting in misleading knowledge (Table 2).

The striking within-genotype genetic variability reported in Italy can hardly be justified by viral evolution and, instead, is probably attributable to multiple introduction events. This hypothesis is also supported by the topology of the phylogenetic tree, where different Italian sequences are part of different clades (Supplementary data I). Haplotype networks allowed us to explore the relationships between strains collected in and outside of Italy and between strains collected within Italy. Of note, an accurate link between countries is often hard to be establish due to a limited phylogenetic signal and high phylogenetic uncertainness, leading to multiple reticulations, particularly, between strains with a high percentage of identity. Sampling bias could also not be excluded due to the different sampling and sequencing activities carried out by different countries and to the different quality of the metadata available for the deposited sequences. Nevertheless, the existing information allowed us to depict an adequate overview of PCV2's molecular epidemiology in Italy, including the pattern of its introduction and subsequent spreading. Multiple introduction events probably characterised all genotypes. PCV2a was imported in separate events, potentially from Asia and Europe, resulting in the high variability of Italian sequences mirroring the heterogeneity of PCV2a throughout the world. The detection of different, distantly related viruses in different Italian provinces is probably attributable to the decreasing prevalence and circulation of this genotype, leading to the progressive geographical isolation of the strains. Currently, PCV2b is the most prevalent genotype, with a multitude of haplotypes having been recognised (Vidigal et al., 2012 and this study). Nevertheless, only 11 PCV2b haplotypes were identified in more than one country. Nineteen PCV2b Italian strains (33,92%) belong to three of these haplotypes, which mainly include viruses collected from Europe and Asia but also viruses collected from South America. The high similarity of the Italian PCV2 haplotypes might suggest progressive evolution from an initial introduction event; however, the presence of the same haplotype in distant countries conflicts with this hypothesis (Figure 2b). The occurrence of the same mutation at different time points and different geographical locations is extremely unlikely. A more parsimonious approach would propose a single mutational event followed by spread in different countries, including Italy, through international trade. Accordingly, sequences belonging to the same haplotypes as those of Italian sequences were mainly

sampled from countries from and to which Italy imports (i.e., Denmark, France and Netherlands) or exports (i.e., Eastern European countries) live swine (http://comtrade.un.org/data/ accessed 25/08/2014). The impact of Italian exportations on viral spread have already been stressed in a previous study by Drigo et al., which demonstrated the Italian origin of Slovenian PRRSV strains (Drigo et al., 2014b). Accordingly, the comparison of Italian PCV2 strains with a collection of partial ORF2 Slovenian sequences confirmed the close relationship among viruses collected on both sides of the Italy-Slovenia border (data not shown). Haplotypes detected only in Italy originated from either haplotypes already detected in Italy or from ones detected only in foreign countries (Figure 2b). Overall, those data indicated a wide network of commercial and epidemiological links that allowed for the introduction of multiple strains from other countries that then became established, gained prevalence and rapidly evolved. Within the national borders, a similar heterogeneity was demonstrated, with different haplotypes being detected in the same province and with different provinces harbouring the same haplotype, which reflects a diffuse viral circulation in Italy. A diffusion network among Italian provinces, substantially overlapping to those reported in this study, has also been identified for PRRSV (Franzo et al., 2014a), indicating that the co-circulation of these viruses is sustained by common epidemiological links and transmission routes. PCV2d is the most recently introduced genotype in Italy. It has been detected mainly in Asia (Ge et al., 2012), but has also already been reported in Europe and the USA (Figure 3). As for PCV2b, the relatively small genetic diversity and the high number of reticulations hampered the definition of the precise sources of Italian strains. However, all Italian strains are part of, or seem to originate from, haplotypes detected mainly in China, potentially as a result of separate introduction events. The geographical structure of PCV2d within Italy is simpler than that of PCV2b, with each haplotype present in just one province, even if the same province (i.e., Padova and Treviso) can harbour different haplotypes. Combining the results presented in Figure 1 and Figure 3, it could be speculated that within-province variability, at least for Padua, could be the result of multiple introductions. Interestingly, commercial exchanges of live swine between China and Italy are negligible and, more generally, China is essentially an importer of live swine from Europe (Vidigal et al., 2012and http://comtrade.un.org/data/ accessed 25/08/2014). The route of potential introduction of PCV2strains from China remains unknown, and other sources cannot be

excluded. An underestimation of PCV2d circulation in Europe due to lower sequencing activity is also possible.

Overall, this paper represents the first description of PCV2 in Italy. We have demonstrated that the variability of Italian strains is due to both multiple introduction events and rapid viral evolution. Foreign haplotypes, most likely imported from European but also potentially from Asian countries, became established at different time points and spread within national boundaries, progressively evolving as a consequence of rapid mutation rates (Firth et al., 2009) and, potentially, recombination. This paper highlights the limitation of current biosecurity strategies implemented in Italy for the control of PCV2 infection. Considering the subtle nature of this disease, a more intense diagnostic activity should be performed on imported animals or semen, and more rigorous control strategies should be implemented to reduce the risk of importing potential virulent variants.

References

Alarcon, P., Rushton, J., Wieland, B., 2013. Cost of post-weaning multi-systemic wasting syndrome and porcine circovirus type-2 subclinical infection in England - an economic disease model. Prev. Vet. Med. 110, 88-102.

Drew, T.W., 2011. The emergence and evolution of swine viral diseases: to what extent have husbandry systems and global trade contributed to their distribution and diversity? Rev. Sci. Tech. 30, 95-106.

Drigo, M., Franzo, G., Gigli, A., Martini, M., Mondin, A., Gracieux, P., Ceglie, L., 2014. The impact of porcine reproductive and respiratory syndrome virus genetic heterogeneity on molecular assay performances. J. Virol. Methods 202, 79-86.

Dupont, K., Nielsen, E.O., Baekbo, P., Larsen, L.E., 2008. Genomic analysis of PCV2 isolates from Danish archives and a current PMWS case-control study supports a shift in genotypes with time. Vet. Microbiol. 128, 56-64.

Ellis, J., Hassard, L., Clark, E., Harding, J., Allan, G., Willson, P., Strokappe, J., Martin, K., McNeilly, F., Meehan, B., Todd, D., Haines, D., 1998. Isolation of circovirus from lesions of pigs with postweaning multisystemic wasting syndrome. Can. Vet. J. 39, 44-51.

Firth, C., Charleston, M.A., Duffy, S., Shapiro, B., Holmes, E.C., 2009. Insights into the evolutionary history of an emerging livestock pathogen: porcine circovirus 2. J. Virol. 83, 12813-12821.

Franzo, G., Cecchinato, M., Martini, M., Ceglie, L., Gigli, A., Drigo, M., 2014a. Observation of high recombination occurrence of Porcine Reproductive and Respiratory Syndrome Virus in field condition. Virus Res. 194, 159-166.

Franzo, G., Cortey, M., Olvera, A., Novosel, D., Castro, A., Biaginni, P., Segalés, J., Drigo, M., 2014b. Revisiting the Taxonomical classification of PCV2: still a real challenge. J. Virol. (submitted) .

Gao, Z., Dong, Q., Jiang, Y., Opriessnig, T., Wang, J., Quan, Y., Yang, Z., 2014. ORF4-protein deficient PCV2 mutants enhance virus-induced apoptosis and show differential expression of mRNAs in vitro. Virus Res. 183, 56-62.

Ge, X., Wang, F., Guo, X., Yang, H., 2012. Porcine circovirus type 2 and its associated diseases in China. Virus Res. 164, 100-106.

Grau-Roma, L., Fraile, L., Segales, J., 2011. Recent advances in the epidemiology, diagnosis and control of diseases caused by porcine circovirus type 2. Vet. J. 187, 23-32.

Guo, L.J., Lu, Y.H., Wei, Y.W., Huang, L.P., Liu, C.M., 2010. Porcine circovirus type 2 (PCV2): genetic variation and newly emerging genotypes in China. Virol. J. 7, 273-422X-7-273.

Harding, J.C., Clark, E.G., Strokappe, J.H., Willson, P.I., Ellis, J.A., 1998. Postweaning multisystemic wasting syndrome: epidemiology and clinical presentation. Swine health and production 6, 249-256.

Hesse, R., Kerrigan, M., Rowland, R.R., 2008. Evidence for recombination between PCV2a and PCV2b in the field. Virus Res. 132, 201-207.

Huang, Y., Shao, M., Xu, X., Zhang, X., Du, Q., Zhao, X., Zhang, W., Lyu, Y., Tong, D., 2013. Evidence for different patterns of natural inter-genotype recombination between two PCV2 parental strains in the field. Virus Res. 175, 78-86.

Karuppannan, A.K., Kwang, J., 2011. ORF3 of porcine circovirus 2 enhances the in vitro and in vivo spread of the of the virus. Virology 410, 248-256.

Lefeuvre, P., Lett, J.M., Varsani, A., Martin, D.P., 2009. Widely conserved recombination patterns among single-stranded DNA viruses. J. Virol. 83, 2697-2707.

Lv, Q.Z., Guo, K.K., Zhang, Y.M., 2014. Current understanding of genomic DNA of porcine circovirus type 2. Virus Genes 49, 1-10.

Ma, C.M., Hon, C.C., Lam, T.Y., Li, V.Y., Wong, C.K., de Oliveira, T., Leung, F.C., 2007. Evidence for recombination in natural populations of porcine circovirus type 2 in Hong Kong and mainland China. J. Gen. Virol. 88, 1733-1737.

Mankertz, A., Caliskan, R., Hattermann, K., Hillenbrand, B., Kurzendoerfer, P., Mueller, B., Schmitt, C., Steinfeldt, T., Finsterbusch, T., 2004. Molecular biology of Porcine circovirus: analyses of gene expression and viral replication. Vet. Microbiol. 98, 81-88.

Segales, J., 2012. Porcine circovirus type 2 (PCV2) infections: clinical signs, pathology and laboratory diagnosis. Virus Res. 164, 10-19.

Segales, J., Kekarainen, T., Cortey, M., 2013. The natural history of porcine circovirus type 2: from an inoffensive virus to a devastating swine disease? Vet. Microbiol. 165, 13-20.

Vidigal, P.M., Mafra, C.L., Silva, F.M., Fietto, J.L., Silva Junior, A., Almeida, M.R., 2012. Tripping over emerging pathogens around the world: a phylogeographical approach for determining the epidemiology of Porcine circovirus-2 (PCV-2), considering global trading. Virus Res. 163, 320-327.

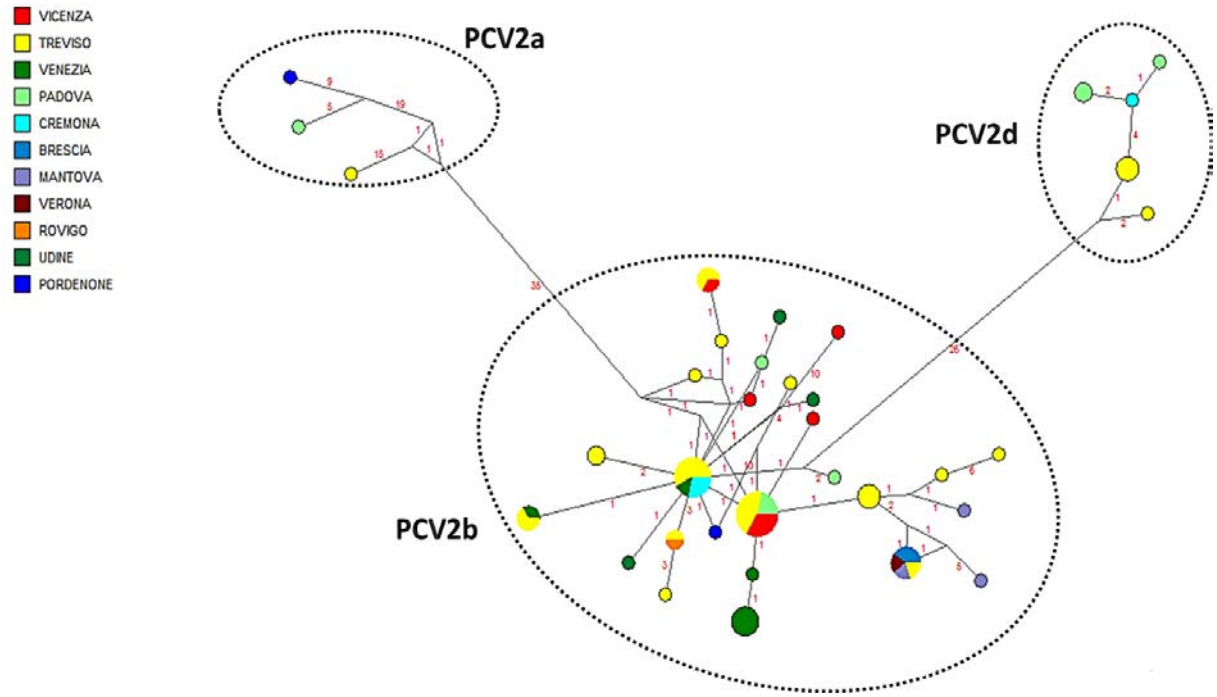

Figure 1. Haplotype network of PCV2. Median-joining (MJ) network of all Italian PCV2 haplotypes built using the Network 4.6.1.20 program. The size of the circumferences is proportional to the haplotype frequencies, while the number of mutations between haplotypes is reported near the branches. Provinces where samples were collected are represented as a color-coded pie chart with slices proportioned to provinces' frequency.

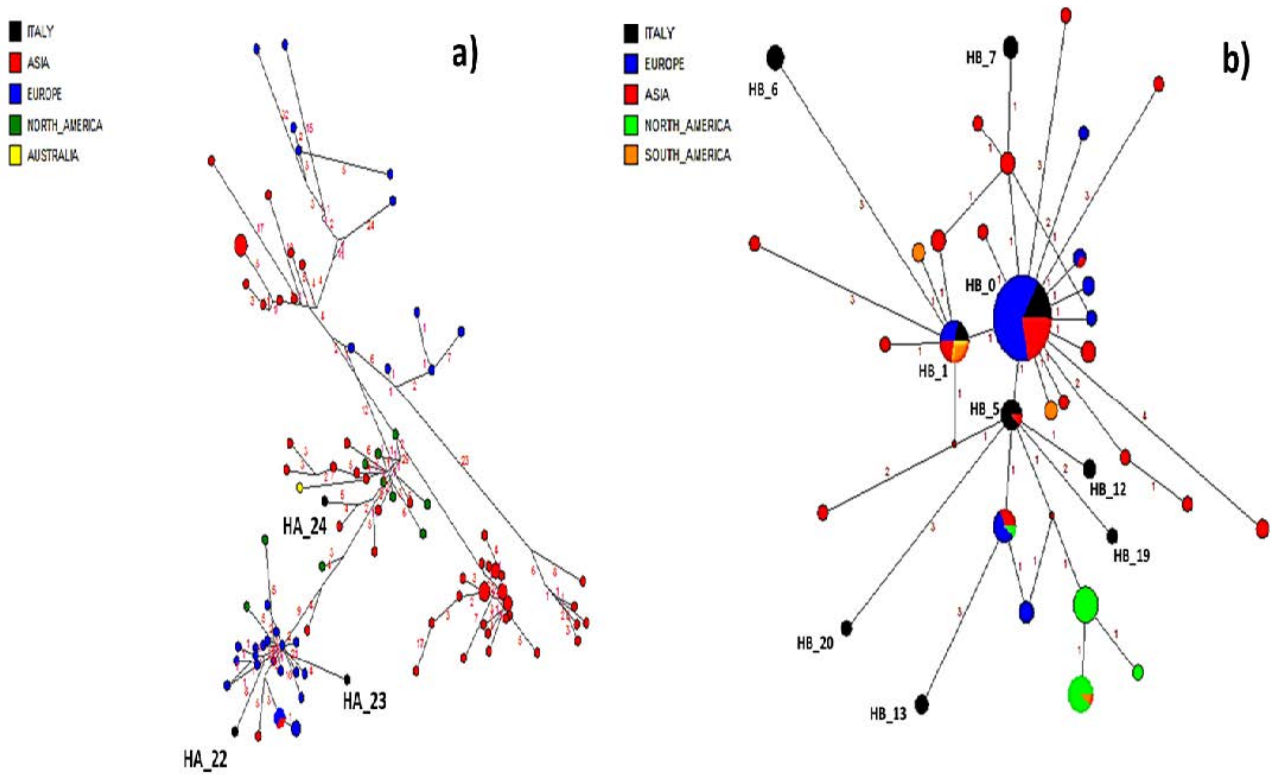

Figure 2. Median-joining (MJ) network of PCV2a (a) and b (b). The haplotypes are grouped according to genotype classification. The size of the circumferences is proportional to the haplotype frequencies while the number of mutations between haplotypes is reported near the branches. Haplotypes are color coded according to the continents where samples were collected while Italian samples are colored in black.

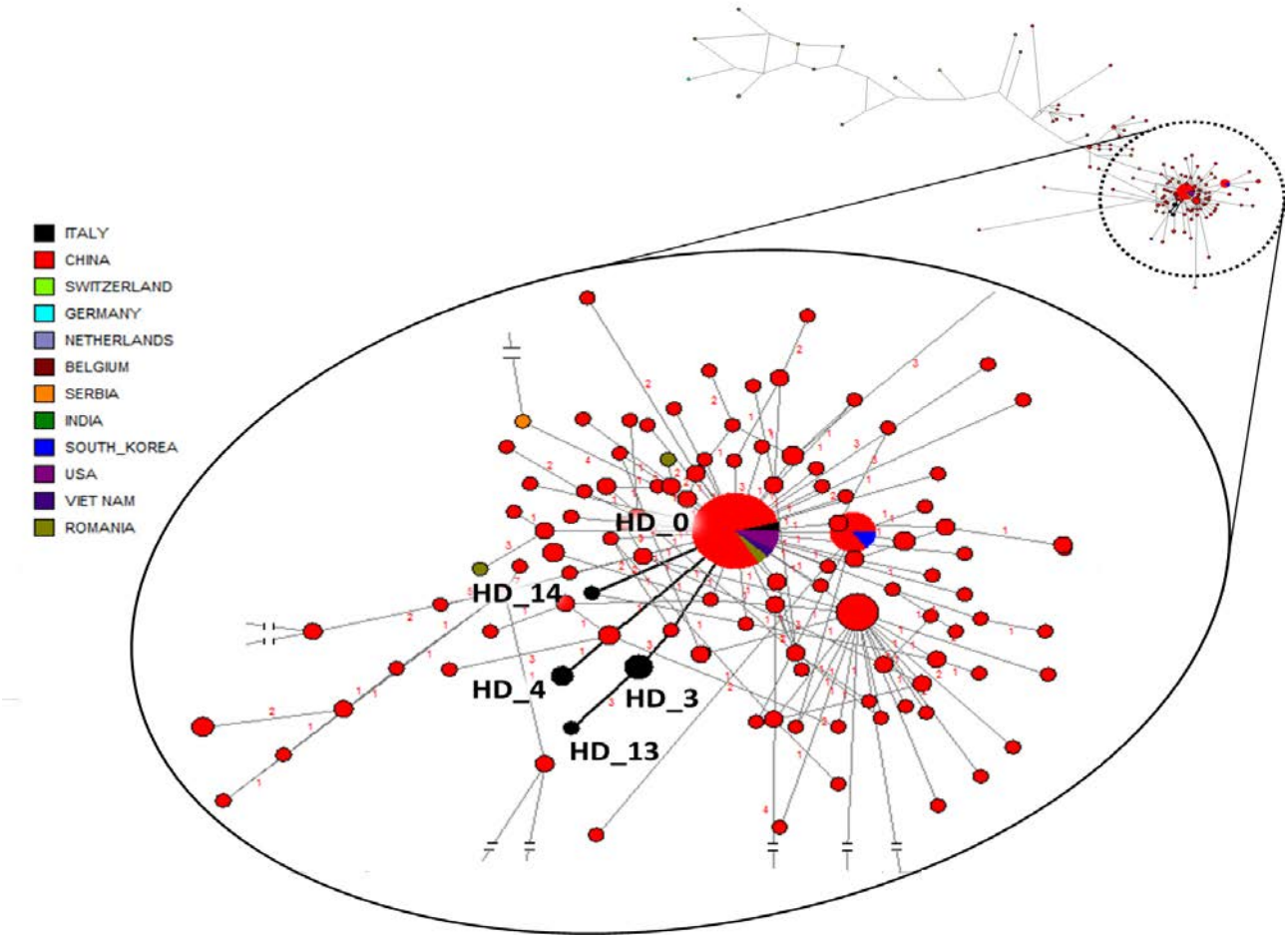

Figure 3. Median-joining (MJ) network highlighting PCV2d haplotypes related to Italian ones. The size of the circumferences is proportional to the haplotype frequencies the while the number of mutations between haplotypes is reported near the branches. Haplotypes are color-coded according to the country where samples were collected; Italian samples are colored in black.

Segment	Primer	Amplicon Length	Position[1]
A Fw	5'-CACCTCAGCAGCAACATGCC-3'	544	36-55
A Rv	5'-CAGGTGGCCCCACAATGACGTGTAC-3'	544	555-579
B Fw	5'-AGTACCTTGTTGGAGAGCGGGAGTC-3'	542	420-444
B Rv	5'-GGGGGGGAAAGGGTGACGAACTGG-3'	542	939-961
C Fw	5'-TCTACTGAGACTGTGTGATCG-3'	746	725-745
C Rv	5'-CTCTGTRCCCTTTGAATACT-3'	746	1449-1468
D Fw	5'-GGCGGGAGGAGTAGTTTACATAGGGTC-3'	675	1293-1320
D Rv	5'-ATTACCCTCCTCGCCAACAA-3'	675	181-200

[1] genomic position related to PCV2 sequence Accesion Number AY484412.

Table 1. Primers used for genome amplification (Fw: forward primer; Rv: reverse primer).

Event	Identification n	Full Genome	ORF2	Recombinating genotypes	Beginning Recombination breakpoint	Ending recombination breakpoint
1 (2 seq)	1053 2665/2012	PCV2d	PCV2d	2a-2d	40	897
2 (2 seq)	2542/18 2542/16	PCV2d	PCV2d	2b-2d	1389	795
3 (2 seq)	26d 24d	PCV2a/PCV2b	PCV2a/PCV2b	2a-2b	1437	1758
4 (2 seq)	28758 28031	PCV2d	PCV2b	2b-2d	817	1459
5 (1 seq)	1748/35/2011	PCV2d	PCV2b	2b-2d	1354	1641
6 (1 seq)	2436/76	PCV2d	PCV2d	2b-2d	817	1369

Table 2. Recombination events list and sequence classification according to ORF2 and full genome reconstructed trees.

	ORF2						Full genome		
	Nucleotide			Amino acid					
	p-distance	*range*	*Ntd number*	*p-distance*	*range*	*AA number*	*p-distance*	*range*	*ntd*
PCV2a	0,047	0,020-0,063	699	0,054	0,026-0,073	233	0,023	0,010-0,030	1768
PCV2b	0,007	0-0,033	699-702	0,007	0-0,034	233-234	0,01	0-0,025	1767
PCV2d	0,005	0-0,011	702	0,004	0-0,013	234	0,003	0,001-0,006	1767
All sequences	0,026	0-0,11	699-702	0,028	0-0,133	233-234	0,02	0-0,056	-
Inter-genotype 2a-2b	0,085	-	-	0,099	-	-	0,045	-	-
Inter-genotype 2a-2d	0,101	-	-	0,114	-	-	0,053	-	-
Inter-genotype 2b-2d	0,055	-	-	0,058	-	-	0,039	-	-

Table 3. Mean p-distance values, variability ranges and lengths calculated at nucleotide level for ORF2 and complete genome. The same values were calculated also at amino acid level for ORF2 only.

Revisiting the taxonomical classification of Porcine Circovirus type 2 (PCV2): still a real challenge

Submitted to Journal of Virology

Giovanni Franzo[1]*#, Martí Cortey[2]*, Alex Olvera[3], Dinko Novosel[4], Alessandra Castro[5], Philippe Biagini[6], Joaquim Segalés[7,8]‡, Michele Drigo[1]‡

[1] Department of Animal Medicine, Production and Health (MAPS), University of Padua. Viale dell'Università 16, 35020 Legnaro, Padua, Italia [2]The Pirbright Institute, Pirbright, Woking, UK,[3]irsiCaixa - HIVACAT, Badalona, Spain, [4]Department of Pathology, Croatian Veterinary Institute, Zagreb, Croatia, [5]Department of Preventive Veterinary Medicine and Animal Health, College of Veterinary Medicine, University of São Paulo, Brazil, [6]Unité de Virologie Moléculaire, Emergence et Co-évolution Virale UMR CNRS 7268, Marseille, France, [7]Centre de Recerca en Sanitat Animal (CReSA), Fundació UAB-IRTA, 08192 Bellaterra, Barcelona, Spain, [8]Departament de Sanitat i Anatomia Animals, Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Barcelona, Spain.

#Address correspondence to Giovanni Franzo, giovanni.franzo@studenti.unipd.it

*, G Franzo and M Cortey contributed equally to this work.

‡, J Segalés and M Drigo contributed equally to this work.

Running title: PCV2 intraspecific classification revisited

Abstract

PCV2 has emerged as one of the most devastating viral infections of swine farming, causing a relevant economic impact due to direct losses and control strategies expenses. Epidemiological and experimental studies have evidenced the role of genetic diversity in potentially affecting PCV2 virulence. The growing number of PCV2 complete genomes and partial sequences available at GenBank questioned the accepted PCV2 classification. Nine hundred seventy five PCV2 complete genomes and 1,270 ORF2 sequences available from GenBank were subjected to recombination, PASC and phylogenetic analyses and results were used for comparison with previous classification scheme. The outcome of these analyses favor the recognition of four genotypes on the basis of ORF2 sequences, namely PCV2a, PCV2b, PCV2c and PCV2d-mPCV2b. To deal with the difficulty of founding an unambiguous classification and accounting the impossibility to define a p-distance cut-off, a set of reference sequences and marker positions that could be used in further phylogenetic studies for PCV2 genotyping was established. Being aware that extensive phylogenetic analysis are time-consuming and often impracticable during routine diagnostic activity, those nucleotide marker positions in the reference sequences for the ORF2 gene, adequately conserved, should be useful to perform a quick genotype differentiation.

Importance

The mandate of the International Committee on Taxonomy of Viruses (ICTV) does not include any consideration below the species level; therefore, a definition for any taxon below such level has been left to the initiative of specialist groups. Since now, PCV2 has been traditionally classified in genotypes assuming as criteria the genetic distance among strains, based on a genotype classification proposal of 2008. However the rapid evolution of PCV2, coupled with the increased number of submitted sequences, has challenged previous classification schemes and criteria. In this work the current PCV2 genotype classification is revisited, investigating different approaches, on the basis of an updated database. A different and fast approach based on unambiguous marker positions for every genotype is proposed.

Introduction

Members of the family *Circoviridae*, genus *Circovirus*, are icosahedral, non-enveloped viruses with single-stranded ambisense circular genomes. Two members of this genus have been reported to infect mammals; *Porcine circovirus type 1* (PCV1) and *Porcine circovirus type 2* (PCV2) (1). PCV1, initially designated as porcine circovirus (PCV), was first discovered in 1974 as a permanent contaminant of continuous cell culture PK15 and is considered non-pathogenic (2). At the beginning of 1990s a new syndrome, named Postweaning Multisistemic Wasting Syndrome (PMWS), was described by Clark and Harding (3,4) and the etiological agent, recognized as a *Circovirus* different from PCV1, was first isolated in 1998 (5). Since the first reports, PCV2 infection was reported all over the world and PMWS showed epidemical proportion in Europe and South East Asia by the late 1990's and in the Americas by 2004-05 (6). Progressively, several clinical manifestations, collectively named porcine circovirus diseases (PCVD), have been associated with PCV2 infection and are responsible of a relevant economic impact to the pig industry due to direct losses and control measures' costs. PCV2 display a simple ambisense genome ranging from 1766 to1768 nucleotides. Three major Open Reading Frames (ORF1, ORF2 and ORF3) have been recognized (7), encoding for proteins involved respectively in viral genome replication (Rep and Rep'), capsid protein (Cap) and possibly in pathogenesis (ORF3) (8, 9, 10). At the intraspecific level, two major PCV2 groups were initially defined (i.e.11, 12). Applying PASC analyses (13) and linearized phylogenetic trees (14) under the assumption of strict molecular clock, two nucleotide diversity cut-offs for ORF2 (3.5%) and the complete genome (2.0%) between them were proposed (15). Accordingly, PCV2 genotypes a and b were designated under the European project no 513928 of the Sixth Framework Programme (www.pcvd.eu)(16). A third genotype retrospectively reported from Denmark in the 1980s was described and named PCV2c (17). Two additional genotypes were proposed after analyzing several sequences from the People's Republic of China (18); although subsequent analyses did not support the genotype status for those strains (19). Another study, restricted to Chinese isolates, also proposed a fourth group, which fitted with the existing definition of genotype and was named PCV2d (7). This late clade was apparently more virulent (20), it was also detected in the USA, but it was designated as mPCV2b (21). The mPCV2b-PCV2d genotype will be referred in this paper as PCV2d. Although initial studies did not find any link between PCV2 sequences and the

occurrence of PMWS (22), crescent amount of literature has evidenced a putative higher virulence of PCV2b compared to PCV2a. Even if comparison between PCV2a and b in experimental infection was difficult and has provided inconclusive results (8, 23), several epidemiological studies afforded evidences of an increased pathogenicity of genotype PCV2b (15). Besides, a genotype shift from the older PCV2a to the new variant PCV2b has been reported during a time window that coincided with an increase in severe outbreaks of PCVD worldwide: Canada (11, 24, 25), China (18), Denmark (17), Spain (26), Sweden (27), Switzerland (28) and USA (29). Interestingly, the oldest PCV2b strains were reported at the beginning of 90' (30), contemporaneously to the first report of PMWS. All these elements suggest the viral genotype as one of the major roles of in the appearance of the disease (6). A growing number of PCV2 complete genomes and partial sequences are available at GenBank (more than 3,300 in July 2014), most of them updated after 2008 when the PCV2 genotype proposal was published. Since then, several new genotypes have been proposed, but also highly divergent sequences have been reported elsewhere (i.e. 31). Therefore, the aim of this paper is to revisit the intraspecific taxonomy of PCV2 and the genotype definition to check its current validity, to unify nomenclature and avoid further misconceptions.

Materials and Methods

Dataset

Nine hundred seventy five PCV2 complete genomes and 1,270 ORF2 sequences available from GenBank (www.ncbi.nlm.nih.gov) were downloaded in January 2014. Complete genome sequences were aligned at nucleotide level. Considering the coding nature of ORF2 sequences, multiple sequence alignment was carried out at amino acid level and was then used to generate the corresponding nucleotide sequences. The MUSCLE algorithm (32) implemented in MEGA6 (33) was employed in both cases. Alignments were visually inspected and edited to remove poorly aligned sequences or those with modification in the reading frame or premature stop codon, highly suggestive of sequencing errors. The alignments robustness was evaluated using Guidance (34) assuming MAFFT (35) as multiple alignment algorithms.

Recombination analyses

Recombination analysis was performed on both complete genome and ORF2 alignments using RDP3 (36). The RDP, GENECONV, MaxChi and 3Seq methods where selected as primary scan, while all the methods implemented in RDP3 where used for recombination detection refinement. Settings for each method were adjusted considering the database features following the recommendations of the RDP3 manual. A recombination event was accepted if detected by more than two methods with a significance p-value of $p<0.01$ with Bonferroni's correction. All sequences identified as recombinant where excluded from further analysis. Topological incongruence was investigated performing manual Bootscan and distance plot analysis using RDP3. To minimize the "noise" effect of intra genotype variability, 95% consensus sequences were reconstructed for each genotype. The whole ORF2 database was implemented in the analysis to benefit of the higher sequences number.

Phylogenetic analyses

Phylogenetic trees for both complete genome and ORF2 were reconstructed using the maximum likelihood (ML) method implemented in PhyML (37). Substitution model was selected according to Bayesian Information Criterion (BIC), calculated using Jmodeltest 2.1.2 (38). A combination of Nearest Neighbor Interchange (NNI) and Sub-tree Pruning and Regrafting (SPR) was used as tree rearrangement strategy. The phylogenetic tree reliability was evaluated using the Shimodaira–Hasegawa [SH]-aLRT (39) likelihood-based measures of branch supports implemented in PhyML. Phylogenetic trees were also reconstructed with the Neighbor Joining (NJ) method and MEGA6 using the substitution model with the better BIC score given by the same MEGA software. The confidence of the internal branches was evaluated performing 1000 bootstrap pseudo-replicates of the aligned dataset.

PASC analyses

Pairwise p-distances among sequences for every dataset were calculated with MEGA6. Distances were ordered and a histogram of pairwised differences was constructed to perform a PASC analysis using Microsoft Excel 2010.

Rates of substitution

Estimations for the rate of substitution were calculated for the nucleotide sequences of the PCV2 genomes and the ORF2 using a Bayesian Markov chain Monte Carlo (MCMC) approach implemented in BEAST v.1.8.0 package (40). Three independent runs of MCMC

per dataset were performed under a strict and a relaxed molecular clock model, using the General Time Reversible model of sequence evolution and the remaining default parameters in the prior's panel. To account for different population dynamics through time a Bayesian Skygrid (41) was chosen as tree prior. The MCMC run was 5×10^7 steps long and the posterior probability distribution of the chains was sampled every 1000 steps. Convergence was assessed by visually inspecting the runs' trace plot and on the basis of an effective sampling size greater than 200 after a 10% burn-in using Tracer software, version 1.6 (42). The estimations are the mean values obtained for the three runs, combined using LogCombiner v1.8.0 (part of the BEAST 1.8. package). The mean substitution rate and the 95% CI were calculated, and the best-fitting models were selected by a Bayes factor using marginal likelihoods implemented in Tracer (43) and using Stepping Stone and Path Sampling approach.

Results

Dataset

After removal of poor quality sequences and exclusion of predicted recombinant sequences detected by RDP3, 595 complete genomes and 954 ORF2 were maintained (Appendix I). Recombination traces were reported in a substantial proportion of the PCV2 whole genome (37.7%) and ORF2 (24.3%) sequences downloaded from GenBank. Recombination was pervasive and affected several genome fragments, intra- and inter-genes, involving strains belonging to both closely and distantly related clades (i.e. different genotypes according to previous classification). A majority of the positions in the alignments were variable, in the complete PCV2 genomes (52.8%) and especially in the ORF2 (76.4% nucleotide, 87.2%aminoacid).

Phylogenetic analysis and genotype definition

Phylogenetic trees reconstructed from PCV2 ORF2 using NJ (Fig1a) and ML (Fig1b) methods displayed very similar topologies and four main clades were identified. These four clades substantially corresponded to the previously defined PCV2a, PCV2b, PCV2c and PCV2d genotypes. Very few strains (n=6) showed contradictory clustering between clades PCV2a and PCV2d in the ORF2 trees (Appendix I). The same clustering in four major groups was obtained rooting the tree using PCV1 sequences as outgroup (data not shown). On the

contrary, phylogenetic trees reconstructed using complete genome showed a different topology due to the exclusion of many recombinants and to the closer relationship between remaining PCV2d and PCV2a strains (named according to ORF2 classification). Actually, the 50 PCV2d genomes clustered with the PCV2a strains using ORF1 and the whole sequence, but form a single clade, closely related with PCV2b, when using the ORF2.

Bootscan analysis

A salient feature of the bootscan and distance plot analyses is that, differently from RDP3, provide evidences of recombinant origin of the whole PCV2d clade (Fig2). Several marker positions along the genome, both in ORF1 and ORF2, are related with PCV2a, but equal to PCV2b in a short fragment of the ORF2. Briefly, along the PCV2d ORF1 nine marker positions agreed with PCV2a. According to the ORF2 alignment provided in the Appendix II, from position 1 to 222-230 six (23, 27, 31, 169, 188, 222) positions of PCV2d mostly match PCV2a; from position 229 to 330 nine marker positions (255, 256, 258, 262, 263, 264, 265, 271, 273) of PCV2d coincide with PCV2b (Table 1, Appendix II); and from position 331 to the end of the alignment (Appendix II) three marker positions equals PCV2a (362, 408, 630). Roughly, the genome of PCV2d is largely similar to PCV2a in the whole ORF1 and the two ends of ORF2; except a central fragment of 100 bp in the ORF2 mimicking PCV2b that carries a large number of marker positions.

PASC analysis

PASC analyses based on the PCV2 complete genome and the ORF2 (Fig3) displayed a multimodal curve. In both cases the definition of a single cut-off value to define PCV2 genotypes appeared complicate. Applying the previously reported cut-off values to classify complete genomes (0.02) and ORF2 (0.35) to the PASC analyses, 27 and 18 genotypes should be defined in PCV2 considering whole genomes and ORF2, respectively. For complete genomes (Fig3a), a threshold between 0.034 and 0.042 would separate PCV2a and PCV2c from PCV2b/PCV2d, but a second threshold of 0.038 would be necessary to differentiate PCV2b and PCV2d. The picture is even more complicated for ORF2 (Fig3b). The first value pointed by the PASC analysis (around 0.068) would be meaningless for the genotype definition and a second value around 0.090 would only differentiate PCV2c from PCV2a/PCV2b/PCV2d. Alternatively, a theoretical cut-off around 7.8% would separate PCV2a, PCV2c and PCV2b/PCV2d, but this value was not robust according to the pairwise

distribution. Consequently, a clear overlap between the p-distances calculated using sequences belonging to the same genotype and the sequences belonging to different genotypes was reported (Appendix III). Even the sequences corresponding to the highly divergent PCV2c genotype displayed a p-distance with those belonging to the PCV2a and PCV2b comprised between 0.087 and 0.165, within the range of intra-genotype distances of the PCV2a (0-0.103) and PCV2b (0-0.12).

Rate of substitution estimation

The substitution rates and the 95% CI for the previously defined PCV2a, PCV2b, PCV2c and PCV2d genotypes based on the ORF2 are shown in Figure 4. The estimated rate for PCV2a (0.00141 subs·site^{-1}·year^{-1}) is significantly higher compared with PCV2b (0.00078) and PCV2d (0.00077). The same picture is reported when the whole PCV2 genome is used, with PCV2a sowing higher substitution rates than the others. Actually, when all genotypes were analyzed together, the marginal likelihood scores indicated that a relaxed molecular clock fit much better in the model than a strict one (data not shown), pointing that the variation among genotypes and among lineages is affecting the goodness of the model.

Reference dataset

In view of the confused results reported in the PASC analysis, an alternative methodology for genotyping PCV2 based on marker positions is proposed. Table 1 summarizes the ORF2 marker positions consistently (>95%) different among PCV2 genotypes. Taking as a reference genotype PCV2b, 9 unique positions were identified to define genotype PCV2a, thirty-three for PCV2c and three for PCV2d (Table 1). In addition, a more robust ORF2 reference dataset was built selecting only those sequences unequivocally part of one genotype on the basis of the topology of both ML and NJ trees (Appendix II).

Discussion

A unified criterion for PCV2 genotyping is paramount to allow the comparison of molecular epidemiology data worldwide (19). Historically, the intraspecific classification of PCV2 has been controversial (19, 20, 22, 44). In 2008, the EU consortium on Porcine Circovirus Diseases proposed a standardized nomenclature for PCV2 genotype definition based on pairwise sequence comparisons (16). The PASC analyses applied to PCV2 complete and capsid (ORF2) nucleotide sequences defined two distance thresholds at 0.020 and 0.035,

respectively (15). Since then, a huge number of PCV2 sequences were deposited in GenBank, and several new genotypes were proposed (i.e. 7), though not always were validated (18, 19, 44). A significant proportion of those sequences have a recombinant origin and in some cases they have been circulating with increasing prevalence in several Asian countries and USA (23, 45, 46). Bearing this in mind, the intraspecific taxonomy of PCV2 and the genotype definition have been revisited, to check its validity benefiting of the large amount of PCV2 sequences published in recent years, to unify nomenclature and avoid further misconceptions.

Based on the results obtained, one of the main PASC and linearized phylogenetic trees assumptions is unequivocally violated. These techniques assume equal rates of evolution among clades, and according to the BEAST estimations for the four main groups of the phylogenetic trees (Fig1) they are significantly different (Fig4). Moreover, according to the Bootscan analysis, the whole PCV2d clade could have a recombinant origin (Table 1, Fig2). These evidences, coupled with the enormous amount of new sequences available and the amount of new diversity described, have important implications for PCV2 intraspecific classification. Mostly, the thresholds applied since 2008 to define PCV2 genotypes using PASC and linearized phylogenetic trees are currently not applicable to all PCV2 strains and therefore these methods should be changed. Since the accepted scheme is no longer valid, an alternative method to unambiguously assign a PCV2 sequence to a genotype is proposed.The suggested approach for classification criteria selection tried to account different theoretical and practical issues. The first challenge when dealing with PCV2 classification is related to the presence of several recombinant strains that as a matter of fact belongs to more than one genotype. Considering the high recombination frequency reported in this study and by other authors, and given the tendency to display a higher frequency of breakpoints between or at the periphery of the genes (47), complete genome sequences poses indubitably a greater challenge for PCV2 classification. The higher percentage of identity of ORF1 gene provides lower phylogenetic signal, limiting its applicability to phylogenetic inference and recombination detection. Finally, complete genome sequencing is laborious and expensive, so that many laboratories currently base their analysis on ORF2 sequences. Taking into account all these factors, a classification approach based on the ORF2 gene is proposed. Aiming to offer an unambiguous classification scheme, several reference sequences, whose classification was clear and

concordant for the two phylogenetic reconstruction methods, were selected (AppendixI and II). This allows recognizing and classifying the four genotypes using both phylogenetic reconstruction and differential marker positions for every genotype. Current classification agrees with that proposed by Olvera et al. (19) with the remarkable difference that strains previously classified into PCV2b Clade 1C are now part of PCV2d genotype. Despite the attempt to exclude all recombinants sequences from our databases using a combination of different approaches with RDP software, current methods are still perfectible and their results are somewhat dependent on databases features, initial settings and subjective refinement of the results. So, it can't be excluded that some recombinant sequences were not identified or misclassified, affecting the results. In practice is infeasible to perform an extensive recombination analysis on routinely basis and results would probably be quite different among different analysis and operators. Actually, BootScan and distance plot analysis has provided evidences, lost using automated RDP analysis, that the whole PCV2d genotype could have a recombinant origin. This ancestral recombination, even if partially masked by the subsequent evolution and by potential superimposed recombinant events, is supported by the differential clustering of PCV2d within the PCV2a clade considering the whole genome (and ORF1) and closer to the PCV2b genotype considering the ORF2. Recombination is a frequent event for PCV2 and, besides PCV2d other recombinant sequences were identified circulating, particularly, but not only, in the PR of China (46, 48, 49, 50, 51).

Nevertheless, considering the epidemiological importance of the PCV2d clade, its wide and increasing distribution (52) and its potential increased virulence (20) we propose to define this group as an actual genotype. Ignoring it would severely complicate the future investigations dealing with those strains and lead to further inconsistencies among different studies. Additionally, the recognition of this phenomenon poses new questions and challenges in the study of recombination role as a driving force of PCV2 evolution. The substantial agreement between the more rapid NJ method and the more accurate ML tree reconstruction methods represents a remarkable advantage. In addition to substantiate the robustness of this classification, it suggests the possibility to use a quicker approach and obtain reliable results. Being aware that extensive phylogenetic analysis are time consuming and often impracticable during routine diagnostic activity, conserved nucleotide marker positions in the ORF2 are proposed to perform a quick genotype

differentiation. These markers consistently (>95%) present in the ORF2 gene among the four main clades, coincident with the four accepted genotypes, are depicted in Table 1. The marker positions can be used as a reference to assign a certain sequence to one of the newly proposed genotypes.

In summary, the present study confirms and validates the variability of viral sequences and the extensive intra- and inter-genotype recombination existing within PCV2 strains, and highlights the difficulty of defining an unequivocal genotype definition of this virus. Considering that the method based on genetic distance seems to be no longer valid and has generated some misclassification through time, it is suggested that an approach based on the reference sequences and/or identification of marker positions can be used for routinely diagnostic activity.

Conflict of interest statement

None of the authors of this paper has a financial or personal relationship with other people or organizations that could inappropriately influence or bias the content of the paper.

References

1. Delwart E, Linlin Li. 2012. Rapidly expanding genetic diversity and host range of the Circoviridae viral family and other Rep encoding small circular ssDNA genomes. Virus Res. 164:114– 121.
2. Tischer I, Gelderblom H, Vettermann W, Koch MA. 1982. A very small porcine virus with circular single-stranded DNA. Nature 295:64-66.
3. Harding J. 1996. Post-weaning multisystemic wasting syndrome: preliminary epidemiology and clinical findings, p. 21. Proceedings of the Western Canadian Association of Swine Practitioners, Saskatoon, Canada.
4. Clark E. 1997. Postweaning multisystemic wasting syndrome, p. 499–501. Proceedings of the Annual Meeting of the American Association of Swine Practitioners, Quebec City, Canada.

5. Ellis J, Hassard L, Clark E, Harding J, Allan G, Willson P, Strokappe J, Martin K, McNeilly F, Meehan B, Todd D, Haines D. 1998. Isolation of circovirus from lesions of pigs with postweaning multisystemic wasting syndrome. Can. Vet. J. 39:44-51.

6. Segalés J, Kekarainen T, Cortey M. 2013. The natural history of porcine circovirus type 2: From an inoffensive virus to a devastating swine diseases? Vet. Microbiol. 165:13–20.

7. Guo LJ, Lu YH, Wei YW, Huang LP, Liu CM. 2010. Porcine circovirus type 2 (PCV2): genetic variation and newly emerging genotypes in China. Virol. J. 7:273.

8. Trible BR, Rowland RR. 2012. Genetic variation of porcine circovirus type 2 (PCV2) and its relevance to vaccination, pathogenesis and diagnosis. Virus Res. 164:68-77.

9. Liu J, Zhu Y, Chen I, Lau J, He F, Lau A, Wang Z, Karuppannan AK, Kwang J. 2007. The ORF3 protein of porcine circovirus type 2 interacts with porcine ubiquitin E3 ligase Pirh2 and facilitates p53 expression in viral infection. J. Virol. 81:9560-9567.

10. Karuppannan AK, Kwang J. 2011. ORF3 of porcine circovirus 2 enhances the in vitro and in vivo spread of the of the virus. Virology 410:248-256.

11. Gagnon CA, Tremblay D, Tijssen P, Venne MH, Houde A, Elahi SM. 2007. The emergence of porcine circovirus 2b genotype (PCV-2b) in swine in Canada. Can Vet. J. 48:811-819.

12. Olvera A, Cortey M, Segalés J. 2007. Molecular evolution of porcine circovirus type 2 genomes: phylogeny and clonality. Virology 357:175–185.

13. Bao Y, Kapustin Y, Tatusova T. 2008. Virus classification by pairwise sequence comparison (PASC). pp. 342–348. In Mahy BWJ, Van Regenmortel MHV (ed), Encyclopedia of Virology, 5 vols. Elsevier, Oxford,UK.

14. Takezaki N, Rzhetsky A, Nei M. 1995 .Phylogenetic test of the molecular clock and linearized trees. Mol. Biol. Evol. 12:823-833.

15. Grau-Roma L, Crisci E, Sibila M, Lopez-Soria S, Nofrarias M, Cortey M, Fraile L, Olvera A, Segales J. 2008. A proposal on porcine circovirus type 2 (PCV2) genotype definition and their relation with postweaning multisystemic wasting syndrome (PMWS) occurrence. Vet. Microbiol. 128:23–35.

16. Segalés J, Olvera A, Grau-Roma L, Charreyre C, Nauwynck H, Larsen L, Dupont K, McCullough K, Ellis J, Krakowka S, Mankertz A, Fredholm M, Fossum C, Timmusk S, Stockhofe-Zurwieden N, Beattie V, Armstrong D, Grassland B, Baekbo P, Allan G. 2008. PCV-2 genotype definition and nomenclature. Vet. Rec. 162:867–868.

17. Dupont K, Nielsen EO, Baekbo P, Larsen LE. 2008. Genomic analysis of PCV2 isolates from Danish archives and a current PMWS case-control study supports a shift in genotypes with time. Vet. Microbiol. 128:56-64.

18. Wang F, Guo X, Ge X, Wang Z, Chen Y, Cha Z, Yang H. 2009. Genetic variation analysis of Chinese strains of porcine circovirus type 2. Virus Res. 145:151–156.

19. Cortey M, Olvera A, Grau-Roma L, Segalés J. 2011. Further comments on porcine circovirus type 2 (PCV2) genotype definition and nomenclature. Vet. Microbiol. 149:522–523.

20. Guo L, Fu Y, Wang Y, Lu Y, Wei Y, Tang Q. 2012. A porcine circovirus type 2 (PCV2) mutant with 234 amino acids in capsid protein showed more virulencein vivo, compared with classical PCV2a/b strain. PLoS One 7:e41463.

21. Xiao C, Halbur PG, Opriessnig T. 2012. Complete genome sequence of a novel porcinecircovirus type 2b variant present in cases of vaccine failures in the UnitedStates. J. Virol. 86:12469

22. Grau-Roma L, Fraile L, Segalés J. 2011. Recent advances in the epidemiology, diagnosis and control of diseases caused by porcine circovirus type 2. Vet. J. 187:23-32.

23. Opriessnig T, Gerbera PF, Xiao CT, Mogler M, Halbur PG. 2014. A commercial vaccine based on PCV2a and an experimental vaccinebased on a variant mPCV2b are both effective in protecting pigsagainst challenge with a 2013 U.S. variant mPCV2b strain. Vaccine 32:230– 237.

24. Carman S, Cai HY, DeLay J, Youssef SA, McEwen BJ, Gagnon CA, Tremblay D, Hazlett M, Lusis P, Fairles J, Alexander HS, van Dreumel T. 2008. The emergence of a new strain of porcine circovirus-2 in Ontario and Quebec swine and its association with severe porcine circovirus associated disease--2004-2006. Can. J. Vet. Res. 72:259-268.

25. Gagnon CA, Tremblay D, Tijssen P, Venne MH, Houde A, Elahi SM. 2007. The emergence of porcine circovirus 2b genotype (PCV-2b) in swine in Canada. Can. Vet. J. 48:811–819.

26. Cortey M, Pileri E, Sibila M, Pujols J, Balasch M, Plana J, Segalés J. 2011. Genotypic shift of porcine circovirus type 2 from PCV-2a to PCV-2b in Spain from 1985 to 2008. Vet. J. 187:363-368.

27. Timmusk S, Wallgren P, Brunborg IM, Wikström FH, Allan G, Meehan B, McMenamy M, McNeilly F, Fuxler L, Belák K, Põdersoo D, Saar T, Berg M, Fossum C. 2008. Phylogenetic analysis of porcine circovirus type 2 (PCV2) pre- and post-epizootic postweaning multisystemic wasting syndrome (PMWS). Virus Genes. 36: 509-520.

28. Wiederkehr DD, Sydler T, Buergi E, Haessig M, Zimmermann D, Pospischil A, Brugnera E, Sidler X. 2009. A new emerging genotype subgroup within PCV-2b dominates the PMWS epizooty in Switzerland. Vet. Microbiol. 136:27-35.

29. Cheung AK, Lager KM, Kohutyuk OI, Vincent AL, Henry SC, Baker RB, Rowland RR, Dunham AG. 2007. Detection of two porcine circovirus type 2 genotypic groups in United States swine herds. Arch. Virol. 152:1035-1044.

30. Firth C, Charleston MA, Duffy S, Shapiro B, Holmes EC. 2009. Insights into the evolutionary history of an emerging livestock pathogen: porcine circovirus 2. J. Virol. 83:12813–12821.

31. Novosel D, Tuboly T, Csagola A, Lorincz M, Cubric-Curik V, Jungic A, Curik I, Segalés J, Cortey M, Lipej Z. 2014. Origin of porcine circovirus type 2 (PCV2) from swine affected by PCV2-associated diseases in Croatia. Vet. Rec. 174:431.

32. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nuc. Ac. Res. 32:1792–1797.

33. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30:2725-2729.

34. Penn O, Privman E, Ashkenazy H, Landan G, Graur D, Pupko T. 2010. GUIDANCE: a web server for assessing alignment confidence scores. Nuc. Ac. Res. 38:W23-W28.

35. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772-780.

36. Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P. 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26:2462–2463.

37. Guindon S, Gascuel O. 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol.52:696-704.

38. Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9:772.

39. Anisimova M, Gil M, Dufayard J, Dessimoz C, Gascuel O. 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst. Biol. 60:685

40. Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7 Mol. Biol. Evol. 29:1969-1973

41. Mandev S, Lemey GP, Faria NR, Rambaut A, Shapiro B, Suchard MA. 2013. Improving Bayesian Population Dynamics Inference: A Coalescent-Based Model for Multiple Loci. Mol. Biol. Evol. 30:713–724.

42. Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v1.6, Available from (http://beast.bio.ed.ac.uk/Tracer)

43. Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. 2012. Improving the Accuracy of Demographic and Molecular Clock Model Comparison While Accommodating Phylogenetic Uncertainty. Mol. Biol. Evol.29:2157–2167.

44. Zhao GH, Cheng W, PJ Zhang, Han YS, Chen DK. 2010. Novel genotypes of type 2 porcine circovirus (PCV2) in PMWS pigs in China between 2008 and 2009. J. An. Vet. Adv. 9:3083-3091.

45. Cai L, Ni J, Xia Y, Zi Z, Ning K, Qiu P, Li X, Wang B, Liu Q, Hu D, Yu X, Zhou Z, Zhai X, Han X, Tian K. 2012. Identification of an emerging recombinant cluster in porcine circovirus type 2. Virus Res. 165:95-102.

46. Wei C, Zhang M, Chen Y, Xie J, Huang Z, Zhu W. 2013. Genetic evolution andphylogenetic analysis of porcine circovirus type 2 infections in southern Chinafrom 2011 to 2012. Infect. Genet. Evol. 17:87–92.

47. Lefeuvre, JM, Lett A, Varsani D, Martin P. 2009. Widely Conserved Recombination Patterns among Single-Stranded DNA Viruses. J. Virol. 83:2697–2707

48. Ma CM, Hon CC, Lam TY, Li VY, Wong CK, de Oliveira T, Leung FC. 2007. Evidence for recombination in natural populations of porcine circovirus type 2 in Hong Kong and mainland China. J. Gen. Virol. 88:1733-1737.

49. Mu C, Yang Q, Zhang Y, Zhou Y, Zhang J, Martin DP, Xia P, Cui B. 2012. Genetic variation and phylogenetic analysis of porcine circovirus type 2 infections in central China. Virus Genes 45:463–473.

50. Huang Z, Su S, Wei C, Xie J, Zhu W, Chen J, Wang H, Zhang G. 2012. Complete Genome Sequence of a Novel Field Strain of Rearranged Porcine Circovirus Type 2 in Southern China. J. Virol. 86:10895.

51. Huang Y, Shao M, Xu X, Zhang X, Du Q, Zhao X, Zhang W, Lyu Y, Tong D. 2013. Evidence for different patterns of natural inter-genotype recombination between two PCV2 parental strains in the field. Virus Res. 175:78-86.

52 Ge X, Wang F, Yang H. 2012. Porcine circovirus type 2 and its associated diseases in China.Virus Res. 164:100-106.

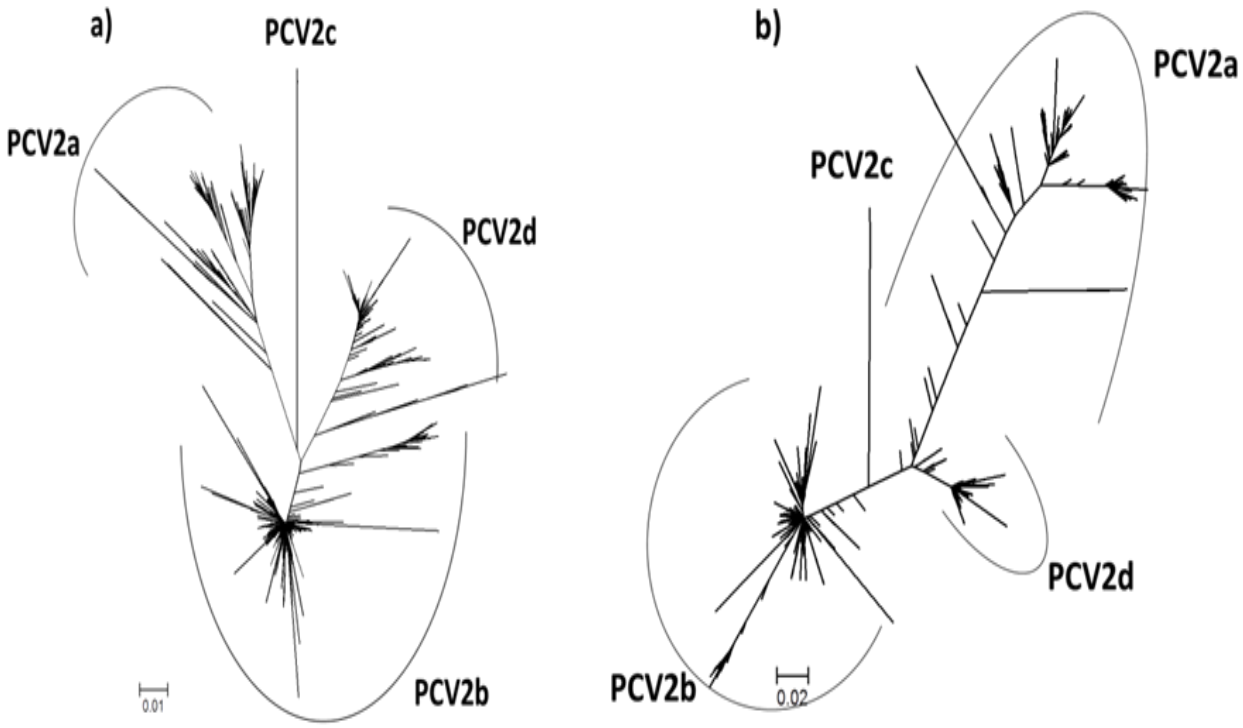

Fig. 1: Phylogenetic trees reconstructed using Neighbor Joining (a) and Maximum likelihood (b) methods based on the ORF2 database after removing the recombinant strains detected by RDP.

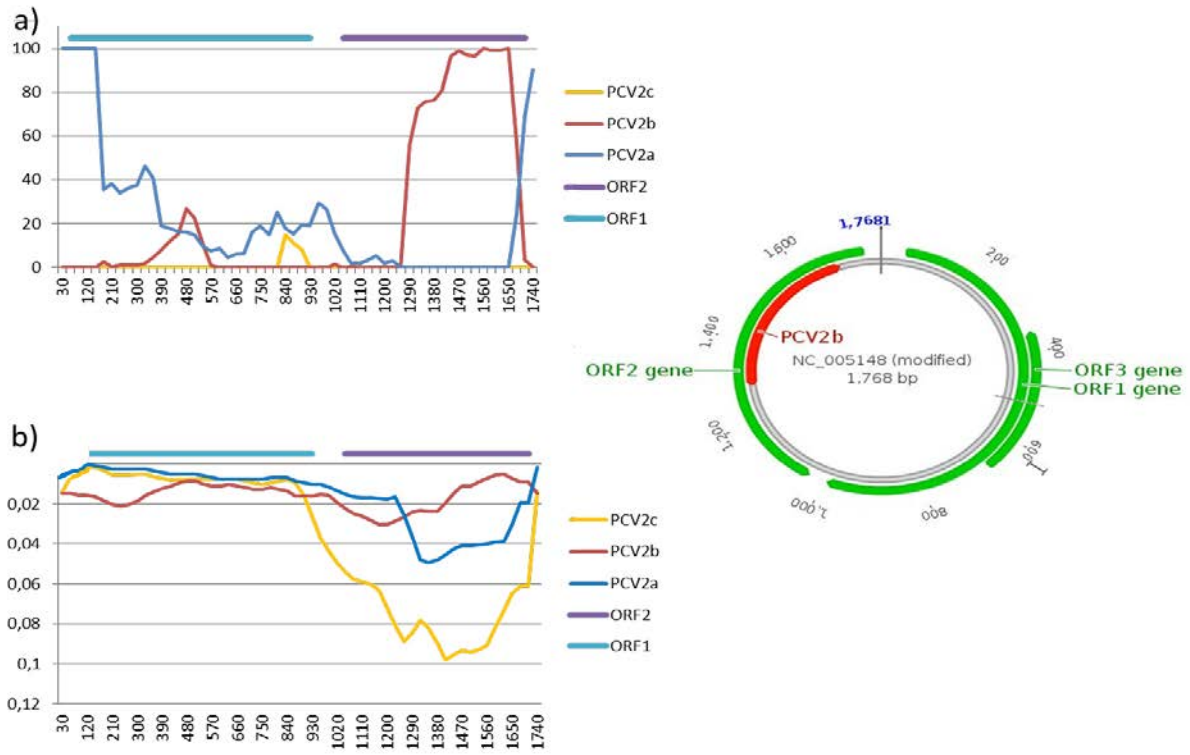

Fig. 2: Bootscan analysis (a) and distance plot (b) analysis performed on the complete genome 95% consensus sequence of each genotype. Bootstrap support (a) and of PCV2d clustering with PCV2a,PCV2b and PCV2c and the respective genetic distance (b) is plotted for different alignment positions. The analysis was performed using a sliding window of 300bp and a step size of 30.

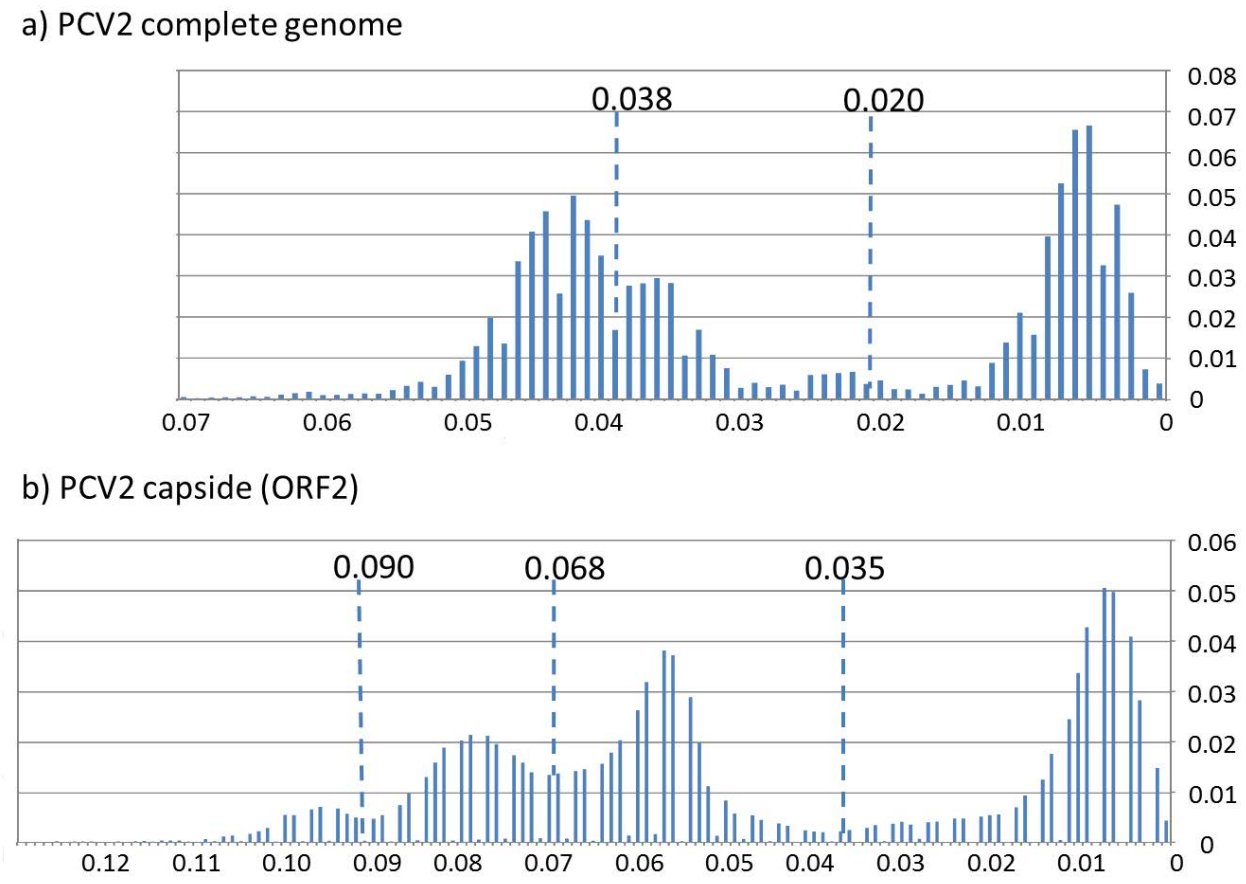

Fig. 3: PASC analysis of complete ORF2 database (a) and of selected reference sequences (b). The percentage of pairwise p-distances comprise within a 0.01 p-distance bin is reported.

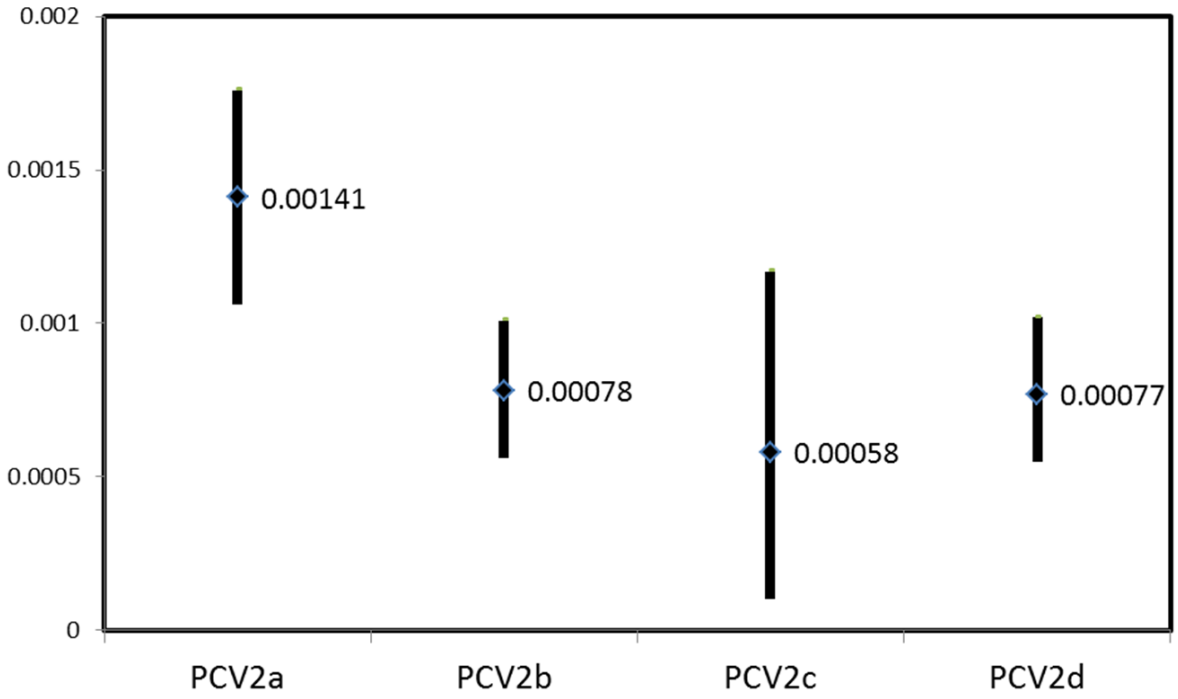

Fig. 4: Substitution rates and 95% CI for the PCV2 genotypes obtained using BEAST.

Table 1. Marker nucleotide positions (95% CI) in the ORF2 among PCV2 proposed genotypes.

Genotype /Position	PCV2a	PCV2b	PCV2c	PCV2d
129	C	C	A	C
138	C	C	T	C
154	A	A	T	A
157	T	T	T	A
159	C	C	T	C
161	G	G	T	G
162	A	A	T	T
168	T	T	C	T
178	A	A	T	A
180	C	C	A	C
181	A	A	C	A
182	C	C	A	C
187	A	A	T	A
190	A	A	C	A
198	C	C	T	C
232	G	G	C	G
255	G	C	C	C
256	A	T	T	T
258	C	A	A	A
262	A	C	C	C
263	A	C	C	C
264	A	C	C	C
265	A	C	C	C
271	A	G	G	G
273	A	G	G	G
306	G	G	A	G
309	T	T	G	T
317	G	G	T	G
318	G	G	T	G
319	C	C	G	C
322	T	T	A	T
324	C	C	A	C
339	G	G	A	G
513	T	T	T	C
585	C	C	T	T
597	C	C	A	C
598	A	A	C	A
599	C	C	A	C
600	T	T	C	T
606	C	C	T	C
607	G	G	C	G
619	T	T	A	T
621	C	C	T	C
623	A	A	C	A
629	A	A	C	A
637	A	A	G	A
643	G	G	G	A

Genetic characterisation of porcine circovirus type 2 (PCV2) strains from feral pigs in the Brazilian Pantanal: an opportunity to reconstruct the history of PCV2 evolution

Submitted to Veterinary Microbiology

Giovanni Franzo[a*#], Martí Cortey[b*], Alessandra Marnie Martins Gomes de Castro[c], Ubiratan Piovezan[d], Matias Pablo Juan Szabo[e], Michele Drigo[a], Joaquim Segalés[f], Leonardo José Richtzenhain[c]

[a] University of Padua, Padua, Italy; [b] The Pirbright Institute, Pirbright, Woking, UK; [c] Department of Preventive Veterinary Medicine and Animal Health, College of Veterinary Medicine, University of São Paulo, São Paulo, SP, Brazil; [d] Empresa Brasileira de Pesquisa Agropecuária - Embrapa Pantanal, Corumbá, MS, Brazil; [e] Laboratório de Ixodologia, Faculdade de Medicina Veterinária, Universidade Federal de Uberlândia, MG, Brazil; [f] Centre de Recerca en Sanitat Animal (CReSA), UAB-IRTA, Campus de la Universitat Autònoma de Barcelona, and Departament de Sanitat i d'Anatomia Animals, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain.

*, equally contributing authors

[#]Corresponding author: Giovanni Franzo

E-mail: giovanni.franzo@studenti.unipd.it / giovanni.franzo1@gmail.com

Full postal address: Department of Animal Medicine, Production and Health (MAPS), University of Padua. Viale dell'Università 16, 35020 Legnaro (Padua), Italy

Tel.: +390498272968

Fax: +390498272973

Abstract

Since its discovery, porcine circovirus type 2 has emerged as one of the most relevant swine infectious diseases, causing relevant economic losses for the pig industry. While four genotypes were identified, only three (PCV2a, PCV2b and PCV2d) are currently circulating and display a worldwide distribution. Another genotype, PCV2c, has been described only once in Danish archive samples collected between 1980 and 1990. In addition to commercial pigs, PCV2 has been demonstrated to infect wild boars and other wild species, which can potentially serve as a reservoir for domestic populations. In this study, eight sequences obtained from feral pigs in the Pantanal region (Mato Grosso do Sul State, Brazil) were compared with reference sequences and other Brazilian sequences, and the results revealed remarkable genetic diversity, with all four genotypes currently recognised being detected (PCV2a, PCV2b, PCV2c and PCV2d). This finding represents a remarkable discovery, as it is the first detection of PCV2c since 1990 and the first-ever detection of PCV2c in live animals. The peculiar population history and ecological scenario of feral pigs in the Pantanal coupled with the complex, and still only partially known, relationships of feral pigs with other PCV2 susceptible species (i.e., domestic pigs, wild boars and peccaries) open exciting questions concerning PCV2 origin and evolution. Overall, the results of the present study led us to form the following hypothesis: these PCV2 strains found in feral pigs may be the last descent of the strains that circulated among European pigs in the past, or they may have infected these feral pigs more recently through a bridge species.

Keywords: porcine circovirus type 2 (PCV2), Brazil, feral pigs, phylogeny

Introduction

Porcine circovirus type 2 (PCV2) belongs to the family Circoviridae and the genus Circovirus, which includes the smallest viruses currently known to autonomous replicate in eukaryotic cells (Delwart and Li, 2012). This group comprises non-enveloped viruses, with a single-stranded, ambisense, circular genome of less than 2000 bp, that infect several avian species, mammals and, probably, fish (Delwart and Li, 2012). PCV2 is the only member of the genus that is recognised to be pathogenic in mammals. It has a genome of 1767-1768 bp, encoding at least four open reading frames (ORFs). Two ORFs are fundamental for viral infectivity: ORF1 encodes two replicase Rep (314 AA) and Rep'(178 AA) proteins (Mankertz et al., 2004), and ORF2 encodes the only protein constituting the viral capsid (233-234 AA) (Guo et al., 2010). ORF3 and the recently discovered ORF4 code for two proteins involved in pathogenesis, probably through the regulation of viral replication and apoptosis (Jialing et al., 2013).

PCV2 was initially associated with postweaning multisystemic wasting syndrome (PMWS), which was first described in Canada in 1996 (Clark 1997; Harding et al., 1998). Currently, several clinical conditions, collectively named porcine circovirus diseases (PCVD) (Segalés 2012), have been associated with PCV2 and represent a relevant challenge to the swine industry worldwide (Segalés et al., 2013). Indeed, PCV2 infection was reported throughout the world, and PMWS showed epidemical proportions in Europe and South East Asia by the late 1990's and in the Americas by 2004-05 (Segalés et al., 2013). Similar to other ssDNA viruses, PCV2 display a rapid evolution, due to a combination of high substitution rates, frequent recombination and a large population size (Firth et al., 2009). To date, four genotypes have been recognised based on phylogenetic analysis (Franzo et al., 2014). PCV2a was the most prevalent genotype until approximately 2003, when PCV2b became the most widespread genotype (Olvera et al., 2007; Carman et al., 2008; Dupont et al., 2008; Li et al., 2010; Cortey et al., 2011). A third genotype, PCV2c, was described only in Danish archive samples collected between 1980 and 1990 (Dupont et al., 2008). Recently, a fourth genotype, PCV2d, has been identified (Guo et al., 2010).

In Brazil, the disease was first identified in 2000, and retrospective studies have demonstrated PCV2 circulation since 1988 (Ciacci-Zanella et al., 2009). Several studies have reported a relevant circulation of PCV2 (PCV2a and PCV2b) in both commercial swine

farms (de Castro et al., 2007; 2012; Chiarelli-Neto et al., 2009; Ciacci-Zanella et al., 2009) and wild boars (Barbosa et al., 2006; Castro et al., 2012). However, no data are currently available on the circulation of PCV2 strains in feral pigs, despite their potential role as a reservoir of several diseases (Desbiez et al., 2011).

Feral pigs (*Sus scrofa, Artiodactyla, Suidae*) are among the most invasive animal species worldwide, with populations found throughout the world (Engeman et al., 2013). Brazilian feral pigs (Fig. 1a) in the Pantanal area, a huge freshwater wetland (estimated area of 138.183 km^2), are the descendants of domestic pigs imported by European colonizers that became wild after escaping from farms abandoned during the Paraguay War (1865–1870) (Desbiez et al., 2011). They have been present in the Brazilian Pantanal since that time, and understanding their population ecology is important for conservation and economic purposes. Although the ecological consequences of the invasion of feral pigs in the Pantanal are poorly understood, they may have benefited both native wildlife and local people by serving as a replacement species for native wildlife hunting and as potential seed dispersers (Desbiez et al., 2011). However, when present in high densities, herds of feral pigs cause extensive damage to pastures and the breeding of domestic and feral animals, and, most importantly, feral pigs have been found to be potential reservoirs of several diseases (Desbiez et al., 2011). Therefore, the main objective of this paper is to characterise the PCV2 strains circulating among feral pigs in Brazil and to compare them with the strains described in domestic pigs and wild boars.

Material and Methods

Samples

This study was conducted in Nhecolândia, a southeast sub-region of Pantanal, Mato Grosso do Sul State, Brazil (Fig.2). This area is limited by the river Paraguay (West), river Taquari (North) and Rio Negro river (Southeast). Samples were collected in 2010 from fourteen animals with various body conditions, age and sex (Table 1). Feral pigs were captured in forestall areas or by water bodies using locally traditional method with the aid of dogs and lasso. Captured animals were sedated, tagged and released afterwards as described in Ramos et al. (2014) . Capture and handling as described was authorized by federal environmental authorities (SISBIO 21416-1, process number CA: 85625338)

DNA extraction and PCV2 detection

DNA was extracted from serum using a phenol-chloroform and proteinase K protocol. Quantification of PCV2 were performed using SYBR green chemistry and carried out in a StepOne Real Time PCR system (Applied Biosystems, Canada) under universal conditions using primers previously described (Yang et al., 2007). The number of copies of viral DNA was determined by comparison with a standard curve and the viral concentration was expressed as number of PCV2 DNA copies per microliters of serum. Extracted samples were also tested to β-actin as a polymerase chain reaction (PCR) control as previously described Hui et al. (2004) in an attempt to identify potential false-negative results due to failures in DNA extraction or the presence of PCR inhibitors.

PCV2 complete genome sequencing

Samples positive for PCV2 PCR were sequenced as described by An et al. (2007). Primers used in the study are detailed in Table 2. Amplified products were excised from 1.5% agarose gel and purified with a commercial kit (Concert; Gibco-BRL). Bidirectional sequencing reactions, by means of the same primers used for PCR, were performed using the BigDye Terminator kit (version 3.1; Applied Biosystems, Norwalk, Connecticut, USA) and were run on 3500 genetic Analyzers (Applied Biosystems, USA). Assembly of consensus sequences was performed by means of PHRED/PHRAP and the CAP3 program (http://asparagin.cenargen.embrapa.br/phph), with an analysis quality point of 20. All the PCV2 genomes obtained have been deposited and are available in GenBank (Table 1).

Sequence analysis

Brazilian strains were classified in different genotypes following the procedures proposed by Franzo et al. (Franzo et al., 2014). The ORF2 sequences obtained in the present work were aligned with 8 reference sequences representing the 4 PCV2 genotypes accepted nowadays, 2 sequences each (Franzo et al., 2014). In addition, 49 ORF2 Brazilian sequences and 19 complete genomes obtained from different studies (de Castro et al., 2007; Chiarelli-Neto et al., 2009; Ciacci-Zanella et al., 2009; Castro et al., 2012) were added to the database (SupMat I). Alignment was performed using the MAFFT method implemented in Guidance and confidence score evaluated performing 100 bootstrap replicates. Phylogenetic trees were reconstructed using the Neighbor-Joining (NJ) and Maximum Likelihood (ML)

methods implemented in MEGA6 and PhyML. The substitution model was selected on the basis of the Akaike information criterion (AIC) calculated using JmodelTest2.1.1. Confidence for each clade of an observed tree was evaluated performing 1000 bootstrap replicates. The same approach was used to reconstruct the phylogenetic tree based on PCV2 complete genome sequences (when available). Percentage of identity between Brazilian sequences was calculated using MEGA6.

Recombination Analysis

Presence of recombinant strains was investigated on the Complete Genome database using RDP3. RDP, GENCONV MaxChi, Bootscan and 3Seq methods were used for preliminary scan, while all the remaining ones were used for secondary scan. Settings for each method were adjusted to fit the database features according to RDP3 manual. Only events detected by more than two methods with a significance level lower then p-value<0.001 were accepted. Further confirmation was obtained by visually inspecting the recombination event using both RDP3 and Simplot.

Results

Eleven of the fourteen samples tested were positive for PCV2 with viral loads ranging from 4.19×10^3 to 9.06×10^6 copies of PCV2 DNA/mL of serum. Eight complete genomes of PCV2 were obtained (Table 1) from the 11 PCR positive samples.

According to the classification proposed by Franzo et al. (2014), 3 strains (166, 165 and 168) belonged to the genotype PCV2a, two strains (167 and 172) belonged to the genotype PCV2b, and two strains (176 and 177) belonged to the genotype PCV2d (Fig. 3a). Remarkably, strain 163 was unequivocally classified as PCV2c. Considering the phylogenetic analysis based on the complete genome sequence, the topology remained substantially unaltered, with the single exception of strain 167, which displayed an intermediate position between the PCV2b and PCV2d genotypes (Fig. 3b). Comparable topologies were obtained with the NJ and ML approaches.

Recombination analysis confirmed that strain 167 was a recombinant strain (breakpoints approximately in positions 1,152 and 1,712), and strains 172 (PCV2b) and 177 (PCV2d) were identified as the most probable parent strains of strain 167, with the latter providing

nearly the entire ORF2, except its two ends (Fig. 4). All the strains analysed in this study displayed noteworthy genetic diversity among each other (p-distance range 0-0.1197; complete genome p-distance range 0-0.058) and compared to other Brazilian strains (ORF2 p-distance range 0.0114-0.1273; complete genome p-distance range 0.012-0.056).

Discussion

The 2012 census of live pigs in Brazil is one of the highest worldwide (more than 38 million in 2012, FAOSTAT). Wild boar (Fig. 1b) is an exotic species in Brazil that has been present since 1989 (Fig. 2), expanding Northwards from the Uruguay border mainly in the state of Rio Grande do Soul (Pereira das Neves, 2007). In parallel, there are several wild boar breeding farms that are rigorously controlled in the states of Minas Gerais, São Paulo, Paraná, Rio Grande do Sul and Santa Catarina (Barbosa et al., 2014), although some individuals were released for hunting purposes or escaped in the past (Britto and Patrocínio, 2006). Feral pigs are restricted to the Pantanal region, in Mato Grosso do Sul State (Fig. 2). PCV2 is a major cause of economic loss for swine farming throughout the world and in Brazil. PCV2 has been reported to also infect wild boars (54-89% of prevalence in Brazil, Barbosa et al. 2014). It has been speculated that the wild boar might serve as a PCV2 potential reservoir; however, the real role of these populations in PCV2 epidemiology, in contrast to other diseases, is controversial (Ruiz-Fons et al., 2008). In Brazil, a third population that is susceptible to PCV2 is feral pigs (Sollero et al., 2009). This work investigated the circulation of PCV2 in these animals and demonstrated remarkable genetic diversity. Despite the low number of individuals analysed, all four genotypes currently recognised (3 PCV2a, 2 PCV2b, 1 PCV2c and 2 PCV2d) were detected. Although it is impossible to formulate definitive conclusions due to the limited number of sequences, it is interesting to report the different genotype frequencies observed between commercial and wild populations of Sus Scrofa. PCV2b was reported to be the most prevalent genotype in commercial pig populations in Brazil (de Castro et al., 2007; 2012; Chiarelli-Neto et al., 2009), while PCV2a was the most prevalent in wild boars (Barbosa et al., 2006; Castro et al., 2012). The present study is the first analysis of PCV2 in feral pigs in Brazil. This study was the first to identify PCV2d in Brazil and the first to identify a contemporary PCV2c genotype worldwide (PCV2c was only detected in Danish archive samples from 1980s and 1990s, Dupont et al., 2008). Interestingly, one recombinant sequence (strain 167) was reported and appeared to be a recombination of strains 172 (PCV2b) and 176 (PCV2d),

both of which circulate in the same region of the Pantanal. This evidence suggests that PCV2 circulation was high enough to support relatively frequent co-infection and subsequent recombination.

The relatively high genetic distance between viruses identified in feral pigs and previously sequenced viruses obtained from domestic and wild boar in Brazil indicate a certain geographic isolation of these feral pig populations. It is highly likely that PCV2 circulated for a prolonged time in the population, which suggests a limited or negligible influence of feral pigs in commercial farms. In addition, the transport of live pigs between the Pantanal wetland and surrounding plateaus has been restricted in the Mato Grosso do Sul state since 2003 (Agência Estadual de Defesa Sanitária Animal e Vegetal, 2003). The detection of PCV2c in feral pigs in Brazil poses exciting questions concerning PCV2 origin and evolution: how and when were feral pigs infected with PCV2c, and how has PCV2c been maintained in the population? Furthermore, were PCV2c genotypes recently introduced to Brazil or were they present previously? The origin of this PCV2c virus is unknown. In this study, false positives for PCV2c caused by laboratory contamination can be excluded due to the absence of any contact with laboratories possessing the PCV2c genotype. A plausible explanation is that PCV2c strains descended from PCV2c strains circulating among European pigs in past centuries that were imported to Brazil and have survived until now, with unknown prevalence in feral pigs due to their limited contact with commercial pigs and wild boars. Although this hypothesis is not in agreement with the estimation of the spread of PCV2's most recent common ancestor (MRCA) (Firth et al., 2009), the probable underestimation of ancestor age using a molecular clock based approach is still a highly debated issue for rapidly evolving viruses (Holmes. 2003). A similar bias could also be true for PCV2, especially considering the limited timeframe for which sequences are available compared to that for the MRCA estimate. Nevertheless, the high similarity between Danish and Brazilian PCV2c is incompatible with the PCV2 substitution rate. However, it is possible that PCV2c displays a slower evolution rate or, more generally, that different dynamics characterise PCV2c evolution. In particular, current data on PCV2 evolution are mainly based on sequences of strains belonging to genotypes PCV2a and b collected from commercial farms. Due to the extremely limited number of available PCV2c sequences (i.e., 3 before this report) and the limited number of collection times (i.e., 1980, 1987 and 1990),

the current analysis of PCV2 MRCA might be biased by the disproportionate weights of different genotypes.

A second hypothesis may advocate a broad circulation of PCV2c in Europe in the first part of the 20th century that was then progressively superseded by more recently emerged genotypes. PCV2c always appears at the root of PCV2 phylogenies, and it is considered to represent an ancient PCV2 genotype, with PCV2a and PCV2b being more modern (Dupont et al., 2008). According to this hypothesis, European pigs could have acted as vectors of PCV2c transmission to Suidae of other countries until more recent times. In support of this hypothesis, at the beginning of the 20th century, naturalised Brazilian breeds were gradually substituted by a massive import of commercial breeds, mainly from Europe (Cavalcanti, 2000). However, in opposition to this hypothesis, in the Pantanal area, contact is not thought to occur between feral pigs and commercial domestic pigs. Nevertheless, small populations of households pigs (Fig. 1c) are raised in the Pantanal, often as free range animals, which allows for contact with feral pigs and potential pathogen introduction. Although the domestic populations in the Pantanal are substantially isolated from intensive pig farms, sporadic indirect contacts or unidirectional animal introduction from commercial farms to rural ones are likely to exist. In addition, there is some evidence of recent pig releases made by farmers that brought domestic animals from the city to the Pantanal. Similarly, it can be hypothesised that another bridge species, such as the wild boar, could act as a Trojan horse for the introduction of PCV2 (including PCV2c) strains from commercial pigs to feral pigs. Importation of wild boar was reported from Europe to South America, including Brazil, during the 1960´s and 1990´s. Some wild boars were released for hunting purposes or escaped (Britto and Patrocínio, 2006). Furthermore, wild boars spread northwards in Brazil from an initial population in Uruguay by crossing the Jaguarão River during the summer of 1989. Currently, these "javalís" (local name for wild boars) have formed free ranging populations, mainly in the southern area of Brazil: Rio Grande do Sul, Santa Catarina, Paraná, São Paulo, Minas Gerais, Mato Grosso do Sul, Mato Grosso, Goiás and Bahia (Deberdt and Scherer, 2007) (Fig. 2). However, wild boar populations have still not been recorded inside the Pantanal Wetland, and there are no records of crossbreeding between wild boars and feral pigs (Deberdt and Scherer, 2007). Alternatively, several species of Peccary (Family Tayassuidae) (Fig. 1d), which are widely distributed throughout South America, may also act as intermediate host species. The

collared peccary (Pecari tajacu) and white-lipped peccary (Tayassu pecari) are commonly present in the Pantanal of Brazil (Desbiez. 2007; Desbiez et al., 2011) and share their distribution ranges with feral pigs. It was recently shown that peccaries are infected with PCV2 (de Castro et al., 2014). Accordingly, peccaries harbouring PCV2 may have infected feral pigs. However, the inadequate data regarding domestic-rural pig and wild boar distributions and PCV2 prevalence and genotype distribution in peccaries makes a definitive hypothesis with a reasonable degree of confidence impossible to formulate.

Overall, this study reveals an unexpected PCV2 variability within the previously neglected population of feral pigs in the Pantanal (Matto Grosso do Sul, Brazil). Importantly, this is the first report of currently existing PCV2c strains in *Suidae* species. These results open new avenues in the reconstruction of PCV2 history and in the understanding of the role of host and ecological niches and their interactions.

Acknowledgment

The sequences data of this work was supported by the São Paulo Research Foundation (FAPESP), São Paulo State, Brazil (grant 2007/57115-3).

References

Agência Estadual de Defesa Sanitária Animal e Vegetal. 2003. *Estabelece medidas para o trânsito intraestadual de suídeos vivos no Estado de Mato Grosso do Sul*. PORTARIA/IAGRO/MS 607/03.

An, D.J., Roh, I.S., Song, D.S., Park, C.K., Park, B.K., 2007. Phylogenetic characterization of porcine circovirus type 2 in PMWS and PDNS Korean pigs between 1999 and 2006. Virus Res. 129, 115-122.

Barbosa, C.N., Martins, N.R., Freitas, T.R., Lobato, Z.I., 2014. Serological Survey of Porcine circovirus-2 in Captive Wild Boars (Sus scrofa) from Registered Farms of South and South-east Regions of Brazil. Transbound Emerg. Dis.(doi: 10.1111/tbed.12247).

Britto, M., Patrocínio, D., 2006. A fauna de espécies exóticas no Paraná: Contexto nacional e situação atual. In Unidades de conservação: ações para valorização da biodiversidade. , 53-94.

Carman, S., Cai, H.Y., DeLay, J., Youssef, S.A., McEwen, B.J., Gagnon, C.A., Tremblay, D., Hazlett, M., Lusis, P., Fairles, J., Alexander, H.S., van Dreumel, T., 2008. The emergence of a new strain of porcine circovirus-2 in Ontario and Quebec swine and its association with severe porcine circovirus associated disease--2004-2006. Can. J. Vet. Res. 72, 259-268.

Castro, A.M., Castro, F.G., Jr, Budino, F.E., Baldin, C.M., Silva, S.O., Brandao, P.E., Richtzenhain, L.J., 2012. Detection of Genetic characterization of Porcine circovirus 2 (PCV2) in Brazilian wildlife boars. Braz J. Microbiol. 43, 1022-1025.

Cavalcanti, S.S., 2000. *Suinocultura Dinâmica. Escola de Veterinária da UFMG*. 1ª ed., Rome.

Chiarelli-Neto, O., Yotoko, K.S., Vidigal, P.M., Silva, F.M., Castro, L.A., Fietto, J.L., Silva, A., Jr, Almeida, M.R., 2009. Classification and putative origins of Brazilian porcine circovirus 2 inferred through phylogenetic and phylogeographical approaches. Virus Res. 140, 57-63.

Ciacci-Zanella, J.R., Simon, N.L., Pinto, L.S., Viancelli, A., Fernandes, L.T., Hayashi, M., Dellagostin, O.A., Esteves, P.A., 2009. Detection of porcine Circovirus type 2 (PCV2) variants PCV2-1 and PCV2-2 in Brazilian pig population. Res. Vet. Sci. 87, 157-160.

Cortey, M., Pileri, E., Sibila, M., Pujols, J., Balasch, M., Plana, J., Segalés, J., 2011. Genotypic shift of porcine circovirus type 2 from PCV-2a to PCV-2b in Spain from 1985 to 2008. Vet. J. 187, 363.

de Castro, A.M., Cortez, A., Heinemann, M.B., Brandao, P.E., Richtzenhain, L.J., 2007. Genetic diversity of Brazilian strains of porcine circovirus type 2 (PCV-2) revealed by analysis of the cap gene (ORF-2). Arch. Virol. 152, 1435-1445.

de Castro, A.M., Cruz, T.F., Salgado, V.R., Kanashiro, T.M., Ferrari, K.L., Araujo, J.P., Jr, Brandao, P.E., Richtzenhain, L.J., 2012. Detection of porcine circovirus genotypes 2a and 2b in aborted foetuses from infected swine herds in the State of Sao Paulo, Brazil. Acta Vet. Scand. 54, 29.

de Castro, A.M., Brombila, T., Bersano, J.G., Soares, H.S., Silva, S.O., Minervino, A.H., Ogata, R.A., Gennari, S.M., Richtzenhain, L.J., 2014. Swine infectious agents in Tayassu pecari and Pecari tajacu tissue samples from Brazil. J. Wildl. Dis. 50, 205-209.

Deberdt, A.J., Scherer, S.B., 2007. O javali asselvajado: ocorrência e manejo da espécie no Brasil. Natureza & Conservação 5, 23-30.

Delwart, E., Li, L., 2012. Rapidly expanding genetic diversity and host range of the Circoviridae viral family and other Rep encoding small circular ssDNA genomes. Virus Res. 164, 114-121.

Desbiez, A., 2007. Wildlife conservation in the Pantanal: habitat alteration, invasive species and bushmeat hunting. Thesis PhD.

Desbiez A., Keuroghlian, A., Piovezan, U., Bodmer, R.E., 2009. Population ecology of feral pigs in the Brazilian Pantanal. Embrapa Documentos 106, Corumba, Brazil, 44 pp.

Desbiez, A.L., Keuroghlian, A., Piovezan, U., Bodmer, R.E., 2011. Invasive species and bushmeat hunting contributing to wildlife conservation: the case of feral pigs in a Neotropical wetland. Oryx 45, 78-83.

Dupont, K., Nielsen, E.O., Baekbo, P., Larsen, L.E., 2008. Genomic analysis of PCV2 isolates from Danish archives and a current PMWS case-control study supports a shift in genotypes with time. Vet. Microbiol. 128, 56-64.

Engeman, R.M., Massei, G., Sage, M., Gentle, M.N., 2013. Monitoring wild pig populations: a review of methods. Environ. Sci. Pollut. Res. Int. 20, 8077-8091.

Firth, C., Charleston, M.A., Duffy, S., Shapiro, B., Holmes, E.C., 2009. Insights into the evolutionary history of an emerging livestock pathogen: Porcine circovirus 2. J. Virol. 83, 12813.

Franzo, G., Cortey, M., Olvera, A., Novosel, D., Castro, A., Biaginni, P., Segalés, J., Drigo, M., 2014. Revisiting the Taxonomical classification of PCV2: still a real challenge. J. Virol. (submitted) .

Guo, L.J., Lu, Y.H., Wei, Y.W., Huang, L.P., Liu, C.M., 2010. Porcine circovirus type 2 (PCV2): Genetic variation and newly emerging genotypes in China. Vir. J. 7, 273.

Harding, J.C.S., Clark, E.G., Strokappe, J.H., Willson, P.I., Ellis, J.A., 1998. Postweaning multisystemic wasting syndrome: Epidemiology and clinical presentation. Journal of Swine Health and Production 6, 249.

Hui, R.K.H., Zeng, F., Chan, C.M.N., Yuen, K.Y., Peiris, J.S.M., Leung, F.C.C., 2004. Reverse transcriptase PCR diagnostic assay for the coronavirus associated with severe acute respiratory syndrome. J. Clin. Microbiol. 42, 1994-1999.

Jialing, H.J., Cao, J., Zhou, N., Jin, J., Wu, J., Zhou, J., 2013. Identification and functional analysis of the novel ORF4 protein encoded by porcine circovirus type 2. J. Virol. 87, 1420.

Holmes, E.C., 2003. Molecular clocks and the puzzle of RNA virus origins. J. Virol. 77, 3893-3897.

Li, W., Wang, X., Ma, T., Feng, Z., Li, Y., Jiang, P., 2010. Genetic analysis of porcine circovirus type 2 (PCV2) strains isolated between 2001 and 2009: genotype PCV2b predominate in postweaning multisystemic wasting syndrome occurrences in eastern China. Virus Genes 40, 244-251.

Mankertz, A., Çaliskan, R., Hattermann, K., Hillenbrand, B., Kurzendoerfer, P., Mueller, B., Schmitt, C., Steinfeldt, T., Finsterbusch, T., 2004. Molecular biology of Porcine circovirus: Analyses of gene expression and viral replication. Vet. Microbiol. 98, 81.

Olvera, A., Cortey, M., Segales, J., 2007. Molecular evolution of porcine circovirus type 2 genomes: phylogeny and clonality. Virology 357, 175-185.

Pereira das Neves, M., 2007. Levantamento populacional de javalis asselvejados no Estado do Rio Grande do Sul. IBAMA, Internal Report .

Ramos V.N., Piovezan, U., Franco, A.H., Osava, C.F., Herrera, H.M., Szabo, M.P., 2014. Feral pigs as hosts for Amblyomma sculptum (Acari: Ixodidae) populations in the Pantanal, Mato Grosso do Sul, Brazil. Exp. Appl. Acarol. 64, 393-406.

Ruiz-Fons, F., Segalés, J., Gortázar, C., 2008. A review of viral diseases of the European wild boar: effects of population dynamics and reservoir rôle. Vet. J. 176, 158-69.

Segalés, J., 2012. Porcine circovirus type 2 (PCV2) infections: Clinical signs, pathology and laboratory diagnosis. Virus Res. 164, 10.

Segalés, J., Kekarainen, T., Cortey, M., 2013. The natural history of porcine circovirus type 2: from an inoffensive virus to a devastating swine disease? Vet. Microbiol. 165, 13-20.

Sollero, B., Paiva, S., Faria, D., Guimaraes, S., Castro, S., Egito, A., Albuquerque, M., Piovezan, U., Bertani, G., Mariante, A.d.S., 2009. Genetic diversity of Brazilian pig breeds evidenced by microsatellite markers. Livestock Sci. 123, 8-15.

Fig. 1: Images of (a) a Feral pig, (b) a Wild boar, (c) a rural domestic pig, and (d) a collared peccary (*Pecari tajacu*) in the Pantanal region (Taken from Desbiez et al. 2009).

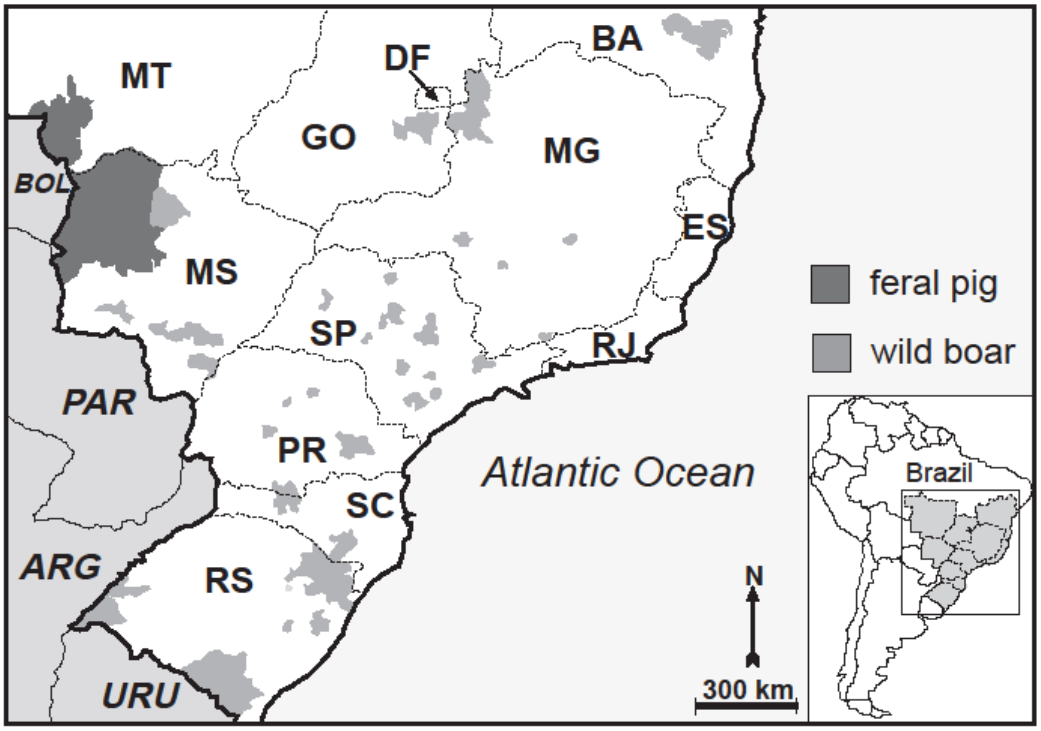

Fig. 2: Map of the southern states of Brazil, indicating the range distribution of feral pigs and wild boars. Collared peccary (*Pecari tajacu*) and white-lipped peccary (*Tayassu pecari*) are widely distributed in all the areas. Bordering countries: BOL, Bolivia; PAR, Paraguay; ARG, Argentina; URU, Uruguay. Brazilian states: MT, Mato Grosso; MS, Mato Grosso do Sul; PR, Paraná; RS, Rio Grande do Sul; SC, Santa Catarina; RJ, Rio de Janeiro; ES, Espírito Santo; BA, Bahia; MG, Minas Gerais; GO, Goiás; DF, Distrito Federal. (Adapted from Debert and Scherer, 2007).

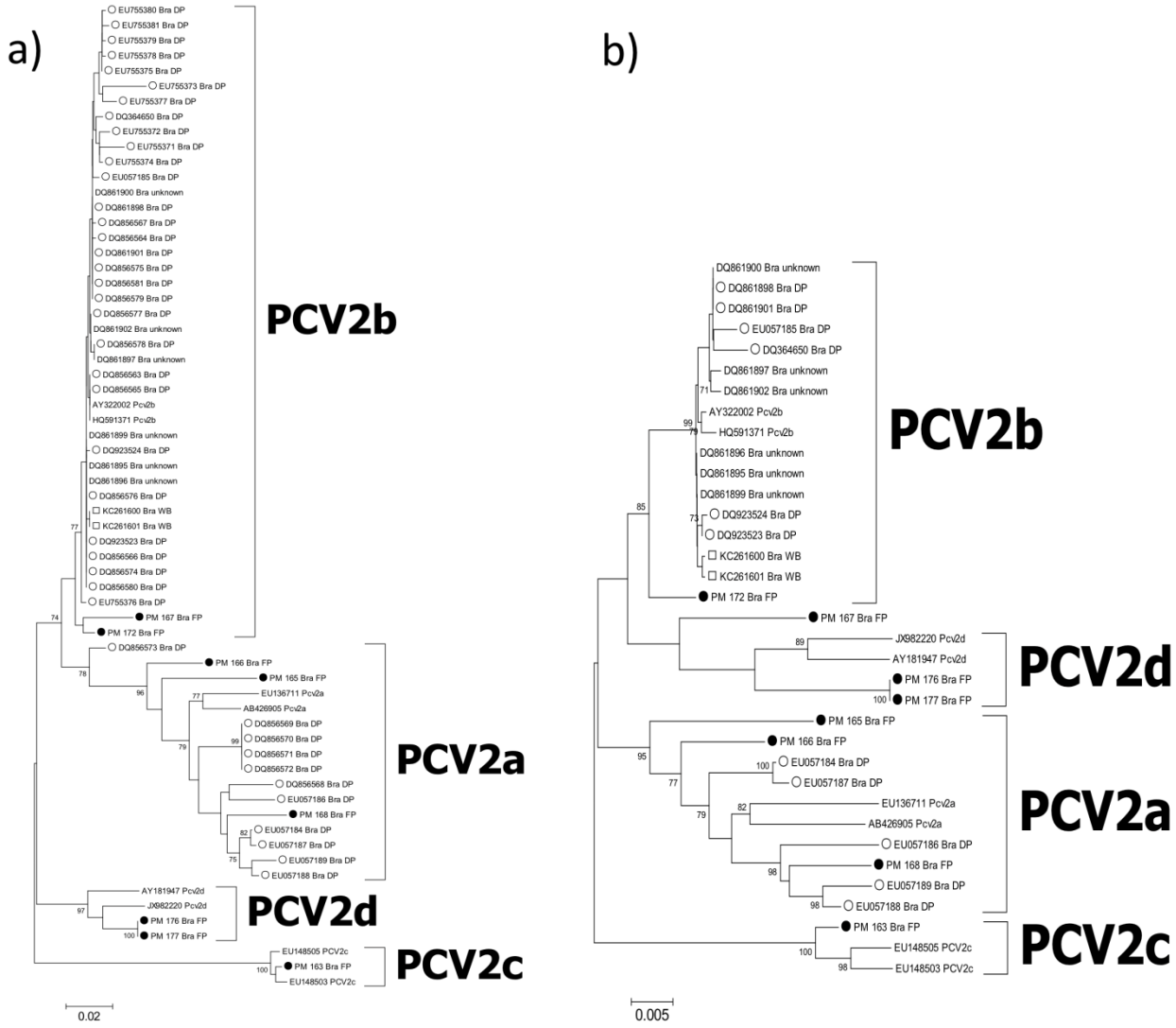

Fig. 3: Phylogenetic trees reconstructed using the Neighbor Joining method with 1,000 bootstrap replicates based on ORF2 (a) and complete genome (b) databases. Brazilian feral pigs sequences are represented by full circles while domestic pigs and wild boar sequences are coded as empty circles and squares, respectively.

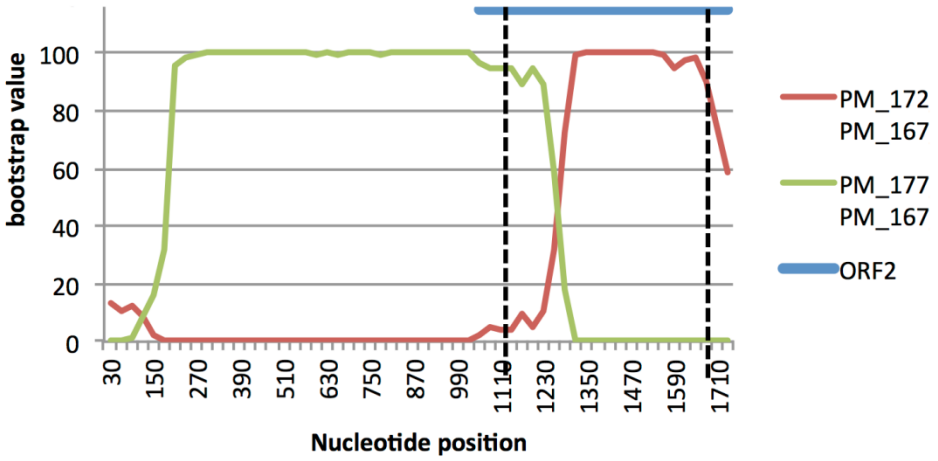

Fig. 4: Bootscan graph obtained by plotting the bootstrap value for strain 176 clustering with strains 172 and 177. Values have been calculated on a sliding window of 400 nucleotides moving through the alignment by steps of 20 nucleotides. Recombination breakpoints are represented as dashed lines. ORF2 position is provided in the upper part of the graph.

Table 1. Identification and features of feral pigs from which samples were obtained. Sequence availability is also reported**.**

Feral pig ID	Age (months)	Weight (Kg)	Sex	Sequence obtained	Accession Number
163	42	62.5	Female	Yes	KJ094599
164	180	64.5	Male	-	
165	48	30	Female	Yes	KJ094600
166	90	65	Male	Yes	KJ094601
167	36	31	Male	Yes	KJ094602
168	72	51	Female	Yes	KJ094603
170	54	81	Male	-	
171	36	39	Male	-	
172	36	33	Male	Yes	KJ094604
173	84	75	Male	-	
174	60	54	Female	-	
175	2	2	Male	-	
176	108	35	Female	Yes	KJ094605
177	24	16	Male	Yes	KJ094606

Table 2. Primers used for amplification and sequencing of full length PCV2 genome.

Primer identification	Sequence (5'-3')	References
P1	TAATCCTTCCGAAGACGAGC	An et al., 2007
P2	CGATCACACAGTCTCAGTAG	An et al., 2007
P3	CAGAAGCGTGATTGGAAGAC	An et al., 2007
P4	ATGTAGACCACGTAGGCCTC	An et al., 2007
P5	AGAAGCTCTTTATCGGAGGA	An et al., 2007
P6	AAGCGAACCACAGTCAGAAC	An et al., 2007
P6-2	CCTTTGAATACTACAGAATAAG	Cortey M (unpublished data)
P7	CTAGAATAACAGCACTGGAG	An et al., 2007
P8	GTTCGTCCTTCCTCATTACC	An et al., 2007
P7-2	TATGGCGGGAGGAGTAGTT	Cortey M (unpublished data)

GENERAL CONCLUSIONS

Field veterinarians and health organizations are continuously required to make decisions whose consequences will have a radical impact on public health and and on the income of producers

In the highly competitive economic system of the Western and rapidly developing countries the profitability of an industry is highly affected by the ability to understand the current situation and predict future trends. Infectious diseases, particularly those rapidly evolving, perfectly fit this scenario and severely challenge the managerial strategies and health plans because of their marked dynamicity. Unfortunately, the lack of adequate information leads to a decision making process based more on the personal experience and on anecdotal knowledge than on factual data.

The set-up of an effective control of infectious diseases must be founded on the information availability about the epidemiological scenario, on the ability to monitor it and to rapidly update current knowledge, and on the competence to implement adequate control measures consciously .

The studies carried out during the PhD programme work deals with several of these issues, providing, at first, a grounded and updated picture of molecular epidemiology of some major avian (aMPV and IBV) and swine (PCV2 and PRRSV) viral diseases in Italy. Additionally, the use of different analysis methods has allowed to provide insights into various aspects of virus ecology and epidemiology, ranging from virus origin and evolutionary forces that shape virus evolution to the role of international trades in affecting viral genetic heterogeneity. The use of different approaches based on mathematical modelling and bioinformatics has been proven effective at evaluating these issues and to provide a more objective interpretation of the data obtained. Far from being merely speculative issues, these information can provide further knowledge to set up more aware control strategies and to evaluate the effectiveness of present plans.

Depending on the specific pathogen and on its peculiar features, further practical implication of rapid viral evolution has been investigated. The issue of rapid and accurate diagnosis of infectious diseases was addressed dealing with the validation and evaluation of diagnostic assays based on molecular biology. The benefits and disadvantages of

different assays were also considered, discussed and their performances evaluated with respect to the Italian epidemiological features.

The results of these studies repeatedly highlighted the limitation of biosecurity measures implemented in our country. In this context, vaccination is considered the main strategy to control the infectious diseases-associated losses. In this sense, this work provides useful information to plan vaccination schemes based on the knowledge of the current epidemiological scenario. Nevertheless, it was also possible to emphasize the importance of the risks and implications that improperly planned vaccination strategies imply.

This work represents an effort to highlight the essential link and interdependence between viral evolution, infection epidemiology in animal populations and the everyday veterinary practice. Indeed, just to report few examples, our work helped a private enterprise in refining their vaccination strategy and allowed to reduce the cost without affecting the sanitary status of the flocks. Similarly, collaborating also with others public and private diagnostic laboratories, the weaknesses of a commercially available diagnostic kit were demonstrated, contributing to its following improvement.

Only the extensive collaboration between these two fields and the continuous devotion to the sharing of information between research institutes and farm companies can generate a virtuous circle that will ultimately contribute to the gaining of new knowledge and to the well-being of this economical sector.

SUPPLEMENTARY MATERIAL

Supplementary material are available at:

Phylodynamic analysis of Porcine reproductive and respiratory syndrome virus (PRRSV) in Italy action of selective pressures and interactions between different clades :

https://drive.google.com/file/d/0ByWa6C2ZM5GCZDlSNWxKVTZEN2M/view?usp=sharing

International trades, local spread and viral evolution: the case of Porcine circovirus type 2 (PCV2) strains heterogeneity in Italy

https://drive.google.com/file/d/0ByWa6C2ZM5GCNEhZcUpWXzBsVGc/view?usp=sharing

Revisiting the taxonomical classification of Porcine Circovirus type 2 (PCV2): still a real challenge

https://drive.google.com/file/d/0ByWa6C2ZM5GCYWFzR21OMEw1aVk/view?usp=sharing

Genetic characterisation of porcine circovirus type 2 (PCV2) strains from feral pigs in the Brazilian Pantanal: an opportunity to reconstruct the history of PCV2 evolution

https://drive.google.com/file/d/0ByWa6C2ZM5GCRXF5M1U5RUNMN3c/view?usp=sharing