Virology **299,** 38–47 (2002)
doi:10.1006/viro.2002.1450

The Genetic Diversity of European Type PRRSV Is Similar to That of the North American Type but Is Geographically Skewed within Europe

Roald Forsberg,* Torben Storgaard,†[1] Henriette S. Nielsen,† Martin B. Oleksiewicz,† Paolo Cordioli,‡ Giovanni Sala,‡ Jotun Hein,* and Anette Bøtner†

*Department of Ecology and Genetics, University of Aarhus, Denmark; †Danish Veterinary Institute for Virus Research, Lindholm, Denmark; and †Istituto Zooprofilattico Sperimentale, Brescia, Italy

Received August 7, 2001; returned to author for revision September 17, 2001; accepted February 28, 2002

Porcine reproductive and respiratory syndrome virus (PRRSV) is a recently emerged pathogen. Two PRRSV genotypes exist, North American and European, which are only 55–70% identical at the nucleotide level. Previous studies have shown high nucleotide diversity in the North American genotype and low nucleotide diversity in the European genotype. Here, we analyzed the ORF5 and ORF7 genes for a large number of new European type PRRSV isolates in conjunction with existing database sequences. This new analysis showed that contrary to previous assumptions, genetic diversity is at least as high in the European genotype as in the North American genotype. Furthermore, we showed that genetic diversity of European type PRRSV has a marked geographical pattern, with exceptionally high genetic diversity among Italian sequences. The geographical pattern of diversity in relation to the epidemiology of PRRSV in Europe is discussed. Discrepancies between ORF5- and ORF7-based genealogies were observed, and further analysis of the data set confirmed the presence of recombination. We were therefore able to report the first observation of recombination in wild-type isolates of European genotype PRRSV.

Key Words: molecular epidemiology; recombination; PRRSV; porcine reproductive and respiratory syndrome virus; genetic diversity; European genotype.

INTRODUCTION

Porcine reproductive and respiratory syndrome virus (PRRSV) is a new member of the *Arteriviridae* family and the causative agent of a novel disease in pigs which emerged recently, and almost simultaneously, in North America and Western Europe (Cavanagh, 1997; Keffaber, 1989; Meng, 2000; Snijder and Meulenberg, 1998; Wensvoort *et al.,* 1991). Even though disease symptoms are very similar on the two continents, and emergence was nearly synchronous, North American and European PRRSV isolates are surprisingly different, with only 55–70% nucleotide identity in the different viral genes (Gagnon and Dea, 1998; Mardassi *et al.,* 1994; Meng *et al.,* 1995a; Morozov *et al.,* 1995; Nelsen *et al.,* 1999). Thus, two genotypes of PRRSV have been defined, North American and European. Many open questions exist concerning the emergence, disease history, and epidemiology of the two genotypes, some of which might be answered through investigation of the geographical distribution of PRRSV diversity on the two continents. It has been amply demonstrated that a high genetic diversity exists among isolates belonging to the North American type of PRRSV (Allende *et al.,* 1999; Andreyev *et al.,* 1997; Gagnon and Dea, 1998; Goldberg *et al.,* 2000; Kapur *et al.,* 1996; Meng *et al.,* 1995a,b; Murtaugh *et al.,* 1998; Nelsen *et al.,* 1999; Pirzadeh *et al.,* 1998; Wesley *et al.,* 1998). In contrast, the average genetic diversity of European type PRRSV has been found to be much lower (Drew *et al.,* 1997; Le Gall *et al.,* 1998; Suarez *et al.,* 1994, 1996). One of the studies (Suarez *et al.,* 1996) on European PRRSV diversity, however, included a single Italian PRRSV isolate that was highly divergent when compared to the remaining European isolates. Furthermore, a recent study of the open reading frame (ORF) 3 gene from a group of purely Danish isolates likewise challenged the notion that European type PRRSV isolates may be less genetically diverse than the North American PRRSV isolates (Oleksiewicz *et al.,* 2000). Instead, we hypothesized that the distribution of genetic diversity in Western Europe might be skewed, so that Italian and Danish isolates harbor the majority of the genetic diversity. The primary objective of the current study was therefore twofold: (i) to determine the geographical pattern of genetic diversity in European type PRRSV; and (ii) to compare the genetic diversity of European type PRRSV to that of the North American type. To investigate these issues, we sequenced the full ORF5 and 7 genes from 46 isolates of European type PRRSV, including many isolates of Italian and Danish origin, and

[1] To whom correspondence and reprint requests should be addressed at Toxicology, Preclinical Development, Novo Nordisk A/S, Novo Nordisk Park, DK-2760 Måløv, Denmark. Fax: +45 44 43 45 58. E-mail: TStS@novonordisk.com.

0042-6822/02 $35.00

TABLE 1

ORF5 Diversity Measures

Geographical area	Genotype	Number of sequences	Average pairwise genetic distance (π)	Maximum pairwise distance (D_{max})
North America	North American	142	0.073	0.153
Europe	European	48	0.090	0.167
Europe, excluding Italy and Denmark	European	19	0.042	0.120
Denmark	European	17	0.060	0.112
Italy	European	12	0.120	0.157

combined these new results with existing ORF5 and 7 sequences from the databases. These two genes were chosen to allow comparison with most previous studies on European PRRSV diversity (Drew *et al.,* 1997; Le Gall *et al.,* 1998; Suarez *et al.,* 1994, 1996). This compilation of new and database sequences constitute the largest sample set of European type PRRSV studied to date. To address the latter of the two questions, we also included all available database sequences of the North American type PRRSV in our analysis and compared these to the data set of European type sequences.

For the purpose of studying the molecular epidemiology of European type PRRSV, we sought to depict the relationship of the isolates in the sample set by constructing evolutionary trees. These trees, however, brought to our attention the presence of genetic recombination in the history of the PRRSV isolates, and hence we were able to report the first observation of evolution by recombination in wild-type PRRS virus of the European type.

RESULTS

Genetic diversity

When comparing the measures of genetic diversity of the ORF5 gene, the European type viruses had a larger average pairwise genetic distance and also a slightly larger maximal pairwise genetic distance than the North American type viruses (Table 1). This was despite the fact that the North American ORF5 sample set was about three times larger than the European sample set, and the result was therefore unlikely to be an artifact of biased sampling. For the ORF7 gene, both the average pairwise genetic diversity and the maximum pairwise genetic distances were of nearly equal size in the European and the North American genotype (Table 2).

When comparing the genetic diversity between different geographical areas in Europe (Tables 1 and 2), Italy had by far the largest average pairwise genetic distance and also the maximally divergent isolates of the whole sample set, when results from both genes were compared. When comparing Danish isolates to isolates from a large geographical area consisting of Belgium, France, Germany, the Netherlands, Spain, and United Kingdom (excluding Italy), the Danish isolates had the highest average pairwise genetic distance in both ORF5 and ORF7. The maximum pairwise genetic distance was also slightly higher in Denmark than in the remainder of Europe, excluding Italy, when comparing the ORF7 sequences, but was of similar size when comparing the ORF5 sequences.

Phylogenetic analysis

Genealogies of ORF5 and ORF7 constructed with the neighbor-joining algorithm are shown in Figs. 1 and 2 (the parsimony algorithm yielded identical results, not shown). For the ORF5 genealogy, bootstrap values above 90% are indicated on the inner branches. Branches in the ORF7 genealogy generally had low bootstrap values and these were omitted in the figure. The genealogies of both ORFs supported a well-defined clade containing the prototype Lelystad isolate and Lelystad-like isolates from a large number of Western European countries including

TABLE 2

ORF7 Diversity Measures

Geographical area	Genotype	Number of sequences	Average pairwise genetic distance (π)	Maximum pairwise distance (D_{max})
North America	North American	90	0.042	0.116
Europe	European	65	0.045	0.111
Europe, excluding Italy and Denmark	European	34	0.017	0.078
Denmark	European	17	0.045	0.093
Italy	European	13	0.070	0.103

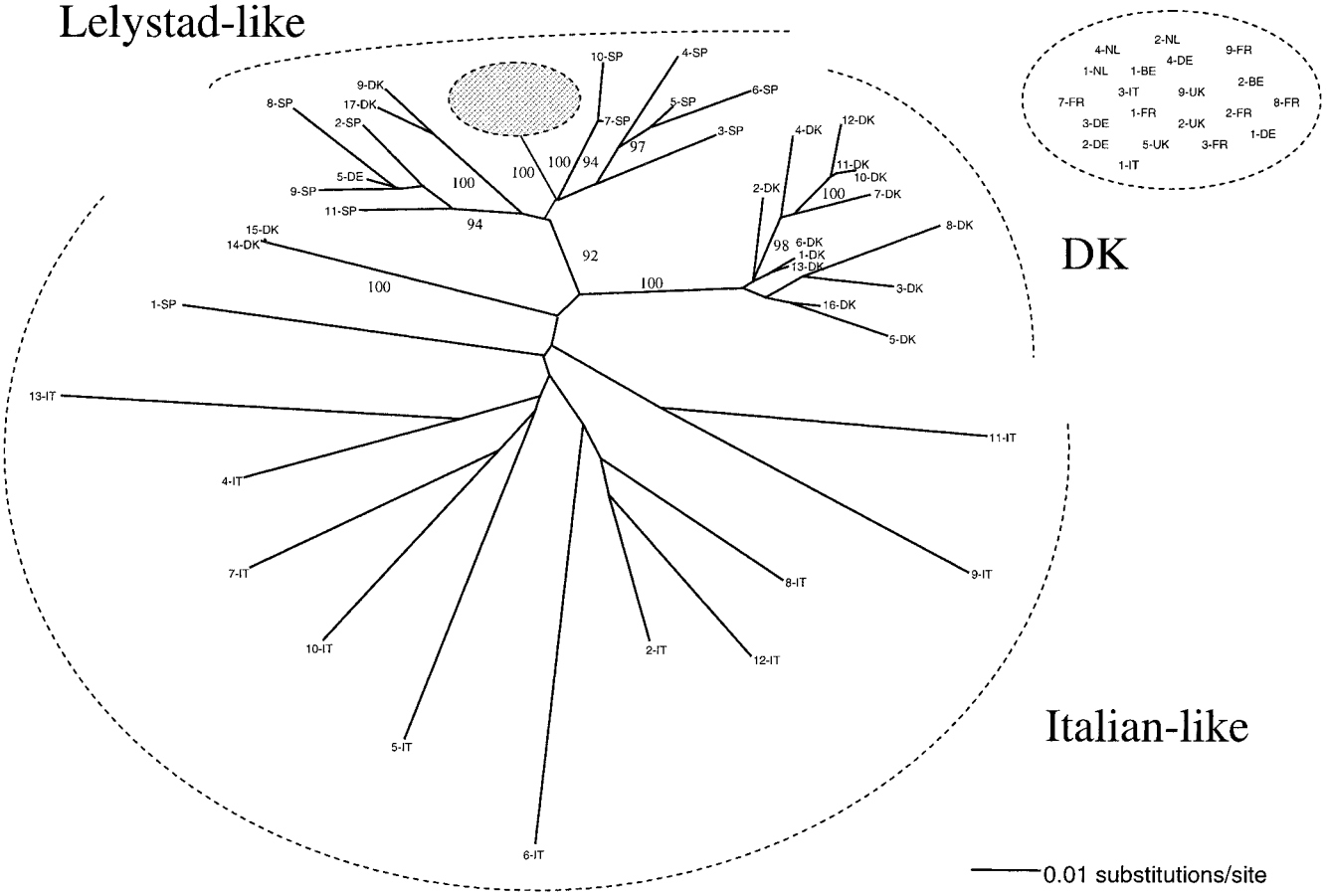

FIG. 1. Genealogy of European type ORF5 sequences. The genealogy was constructed using HKY corrected distances and the neighbor-joining algorithm. The sequence numbers refer to the list of isolates in Table 3, and the geographical origin of the sequences is indicated by the abbreviations: BE, Belgium; DE, Germany; DK, Denmark; FR, France; IT, Italy; NL, the Netherlands; SP, Spain; UK, United Kingdom. The hatched area represents a very tight clade, which is enlarged in the circle at the top right corner. The geographical groupings mentioned in the text are indicated on the figure and the bootstrap support of these groupings are written beside the branches.

the Netherlands, Belgium, Denmark, England, France, Germany, Italy, and Spain, and another well-defined clade of solely Danish isolates (Figs. 1 and 2). A third group with no bootstrap support was composed mainly of highly divergent Italian sequences. We refer to the members of this group as "Italian-like." It is interesting to note that Danish isolates appeared not only in the purely Danish clade but also in the clade of Lelystad-like isolates, and in conjunction with the group of Italian-like isolates. The majority of Italian isolates seemed to have no well-defined relationship, as indicated by the lack of bootstrap support in the Italian-like group, but similar to the Danish situation, Italian isolates also appeared within the well-defined Lelystad-like clade.

Recombination analysis

It was clear from the genealogies that the two genes (ORF5 and 7) had a coupled history in defining larger groups of related sequences. However, it was also clear that there were several discrepancies between the two trees (Fig. 1 and 2), but these occurred in regions of the genealogy that had low or no bootstrap support and were dominated by long external and short internal branches. This could be because the true genealogy was star-like as an effect of rapid population growth in the past, or in the case of the ORF7 gene, because of the little variation/signal in the data. Rapid population growth or low level of variation in the ORF7 gene might as such explain the minor rearrangements occurring between the two trees. However, there were also major rearrangements between the two trees in the large group of Italian isolates, where the ORF7 gene is quite variable. This could indicate that the evolutionary history of the two genes had been decoupled due to recombination events between the ancestors of the sample set. To examine whether recombination had occurred, we selected three isolates of Italian origin, which grouped differently in the ORF5 and 7 trees (Figs. 1 and 2, isolate 2-IT, 10-IT, and 12-IT). From these isolates, we sequenced the ORF6 gene to examine the whole ORF5–7 region for apparent chimeric sequences. This was done using the SimPlot software (Lole *et al.,* 1999) and the result is shown in Fig.

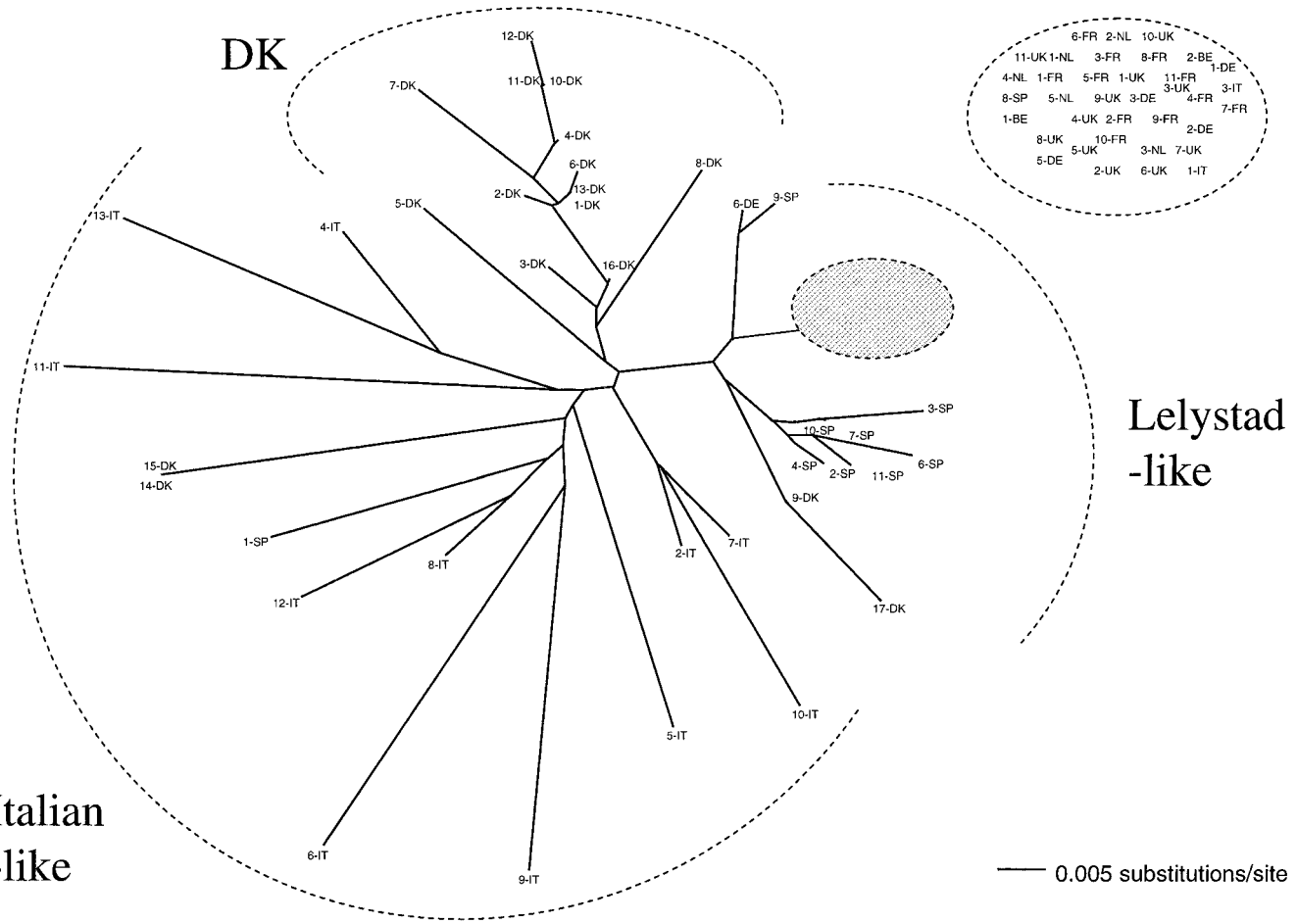

FIG. 2. Genealogy of European type ORF7 sequences. The genealogy was constructed using HKY corrected distances and the neighbor-joining algorithm. The sequence numbers refer to the list of isolates in Table 3, and the geographical origin of the sequences is indicated by the abbreviations: BE, Belgium; DE, Germany; DK, Denmark; FR, France; IT, Italy; NL, the Netherlands; SP, Spain; UK, United Kingdom. The hatched area represents a very tight clade, which is enlarged in the circle at the top right corner. The geographical groupings mentioned in the text are indicated on the figure.

3. For the major part of the ORF5 and 6 genes, the closest relative of the 2-IT isolate was the 12-IT isolate, but there was a sudden shift around the end of the ORF6 gene, where the Lelystad/Boxmeer consensus sequence became the closest relative instead.

Test for a molecular clock

We have previously demonstrated an accurate molecular clock for the ORF3 gene of the European type PRRSV (Forsberg *et al.,* 2001; Oleksiewicz *et al.,* 2000). However, when we performed likelihood ratio tests for the presence of a molecular clock in ORF5 and ORF7, we found that the clock hypothesis was rejected for the complete data set (ORF5+7) and for the two genes separately (results not shown).

DISCUSSION

Characterization and quantification of genetic diversity of PRRSV is of obvious relevance, for example, for the development of new vaccines and diagnostic tests (Meng, 2000). Comparison of genetic diversity within the European and the North American genotypes is also important for the understanding of PRRSV emergence. Finally, characterization of genetic diversity in relation to time and geography is a necessity for the use of sequence information in PRRSV epidemiology. However, comparing the genetic diversity of different viral populations is not always straightforward. It would have been preferable to build a specific population genetic model of the geographical structuring of PRRSV to correct for sampling variance and perform statistical tests of different hypotheses. Due to the nonrandom sampling of the isolates used in the current study, it was not possible to build such a model. Therefore, this study did not employ any rigid statistical testing scheme, but merely reported the genetic diversity data (Tables 1 and 2). An often used measure of the genetic diversity is the average pairwise genetic distance between isolates (see e.g., Kapur *et al.,*

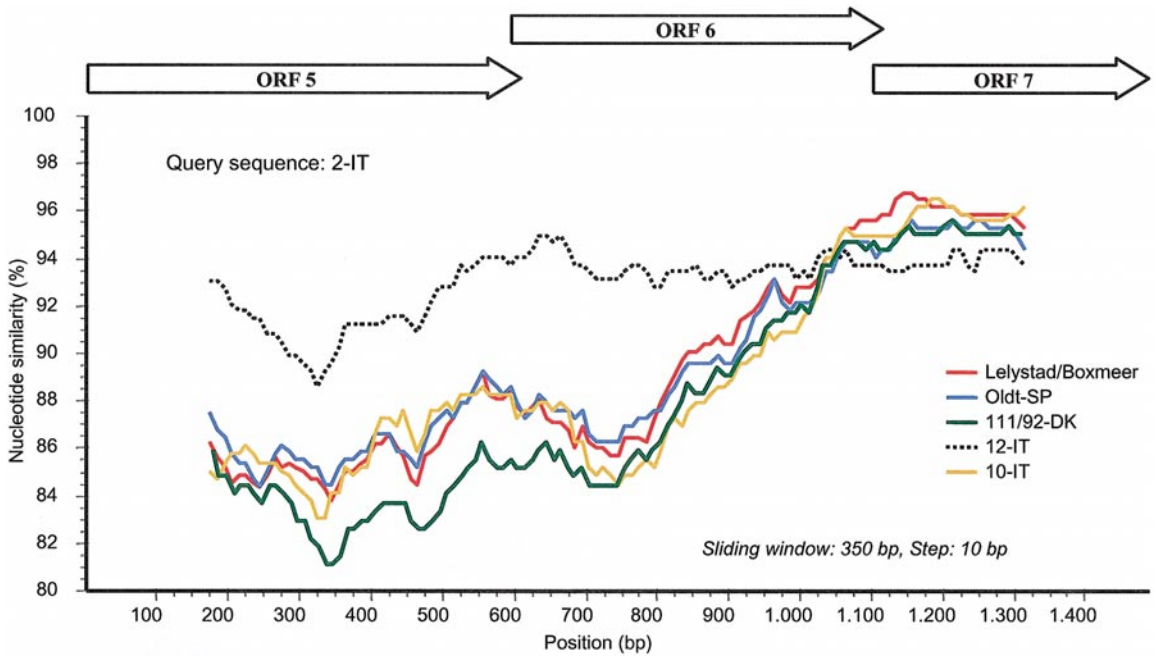

FIG. 3. SimPlot showing recombination point in European type PRRSV. The nucleotide identity between the query sequence and the selected sequences in a sliding window is plotted as a function of position along the PRRSV genome. The ORFs covered are drawn above the plot. The dashed line indicates the 12-IT sequence, which is an apparent chimera sequence that shows a marked shift in identity to the query sequence around the end of the ORF 6 gene.

1996) which we have also used here. However, the maximum pairwise genetic distance is less sensitive to sample bias and was therefore also included (see comments under Materials and Methods). Using both measures, we have shown, contrary to the general assumption, that the genetic diversity quantified by the average pairwise genetic distance is larger between European type PRRSV isolates than between North American type PRRSV and that the maximum genetic distances of isolates from the two populations are of comparable size. The lack of correction for sampling variance means that the measure of average pairwise distance should be compared with caution, but the more than triple size of the ORF5 sample set of North American type isolates makes us confident that the trend in the measure of maximum genetic distance will not be reversed by adding more North American type samples. A classical result from population genetic theory says that the subdivision of a population will yield an increase in the overall genetic diversity, as the reduced size of the subpopulations causes increased genetic drift (Wrights, 1931). Hence, a potential explanation for the larger average pairwise genetic distance found in the European type isolates may be that the European pig population is more structured due to less animal movement across borders and therefore offers fewer opportunities for the migration of virus between countries, leading to a substructured and more heterogeneous virus population. Alternative explanations also exist. For example, different pig breeds are known to exhibit different susceptibility to PRRSV (Christopher-Hennings *et al.,* 2001), which might result in different selective pressures on PRRSV in different countries. An attenuated North American type PRRSV vaccine has been shown to revert to virulence and to be transmitted between swineherds (Allende *et al.,* 2000; Bøtner *et al.,* 1997; Mengeling *et al.,* 1999; Nielsen *et al.,* 2001, 2002; Storgaard *et al.,* 1999). Therefore, it might also be that the lower average genetic distance of North American PRRSV isolates is caused by a large number of vaccine-derived isolates circulating in this population due to the extensive use of live-attenuated vaccines in North America. The similar use of European type live vaccines has been very limited in the time period studied and can therefore not have affected the average genetic distance of European type PRRSV.

Our hypothesis at the start of this study was that a high degree of substructuring in European type PRRSV populations might exist. Comparisons within a large sample of European type isolates confirmed a very pronounced geographical pattern of genetic diversity in Western Europe. The genetic diversity measured by both average pairwise genetic distance and maximum genetic distance was higher in Italy than in the remainder of Western Europe. Denmark appeared to be a geographical area that also harbored exceptionally diverse PRRSV isolates. Our findings in the current study are in agreement with the results of previous studies on European PRRSV diversity. Taken together, Suarez *et al.* (1994, 1996), Drew *et al.* (1997), and Le Gall *et al.* (1998) found very low levels of sequence variability in European type

PRRSV. However, these studies were confined to PRRSV isolates from Great Britain (Drew *et al.,* 1997), Belgium, France, Germany, The Netherlands, and Spain (Le Gall *et al.,* 1998; Suarez *et al.,* 1996). The new PRRSV sequences made in this study (Table 3) confirmed the low level of PRRSV diversity in the abovementioned countries, with the only exception being a single Spanish isolate (L56/2/91). Suarez *et al.* (1996) included a single Italian PRRSV isolate in their study. This isolate (2156) was highly divergent when compared to the rest of the European type PRRSV isolates. Our sequencing of 12 additional Italian PRRSV isolates confirmed that Italian isolates are consistently very different from other European type PRRSV isolates. However, our finding that the diversity of European type PRRSV is just as high or even higher than the diversity of North American type PRRSV is completely new and was only reached due to the inclusion of many new isolates, especially from Italy and Denmark (Table 3).

Despite phylogenetic inconsistencies between ORF5 and 7, we were still able to extract some information from the evolutionary trees of the two ORFs. The trees loosely defined three major groups of European type PRRSV isolates, namely a clade of Lelystad-like isolates covering all of Western Europe including Italy and Denmark, a clade of purely Danish isolates, and a group of highly diverse Italian-like isolates from Italy, Spain, and Denmark. Hence, the evolutionary trees helped to interpret the pattern of genetic diversity of European type PRRSV described above: Italy had much genetic diversity since Italian isolates belonged to two different phylogenetic groups (Italian-like and Lelystad-like) with the majority being included in a group of highly divergent isolates. Denmark held somewhat less diversity as Danish isolates fell into all three phylogenetic groups but with the majority belonging to the same close group. Finally, the remaining European countries only had a single isolate that fell outside the narrow clade of Lelystad-like isolates, a fact that limits the genetic diversity in these countries. The observation that the majority of Danish isolates belonged to a single and unique clade would indicate that the major Danish epidemic stems from an introduction from a single source. This corresponds well with epidemiologic information (Bøtner *et al.,* 1994) and previous phylogenetic analysis (Oleksiewicz *et al.,* 2000). Additionally, the large group of Lelystad-like isolates, which comprise the major Western European epidemic, would also appear to have originated from a single introduction of PRRS virus around the time of the emergence of PRRSV in Europe. Later, genetic material from the same reservoir as that which initiated the Italian and the major Western European epidemic may have been introduced into Denmark, causing the genetic diversity of this region to rise. Recently, a study of ORF5 diversity among PRRSV isolates from the Czech Republic (Indik *et al.,* 2000) reported the finding of two isolates (V-501 and V-503) that in our terminology would be classified as Italian-like. We have, however, chosen not to include these sequences in the current analysis since they were reported together with five Czech isolates from the time period 1995–1998 that all were almost identical to the Lelystad isolate from 1991. This surprising finding requires further analysis. Nevertheless, the finding of Italian-like isolates in the Czech Republic is very interesting and suggests that further sampling in Eastern Europe will be of great interest.

The inference of recombination events in the history of a population sample is a difficult issue (see Wiuf *et al.,* 2001 for a recent discussion) and no optimal method exists. It has previously been shown that recombination has occurred in the North American type PRRSV population (Yuan *et al.,* 1999). Recently, we have also shown that recombination can occur between diverse European type PRRSV sequences under cell culture conditions (van Vugt *et al.,* 2001). In the current study, the demonstration of phylogenetic inconsistency between the ORF5 and ORF7 genes, and the presence of apparently chimeric sequences in the data set, both indicated that recombination has occurred in the history of the European type PRRSV. Therefore, the finer details of European type PRRSV phylogenetic trees and the use of these trees in the inference of disease history should be taken with caution. Without a clear model of the recombination process and the dynamics of the viral population, it is not possible to quantify the amount of recombination in the sample set. The presence of well-defined geographical clades nevertheless indicated that recombination has not resulted in the transfer of a significant amount of genetic material between these clades, and hence, recombination may be a rare occurrence in European type PRRSV and may only affect isolates from areas where the movement of PRRSV-infected pigs is large enough to confer a significant probability of double infection with genetically diverse virus isolates.

We have recently shown that nucleotide substitutions in ORF3 sequences of European type PRRSV follow an accurate molecular clock (Forsberg *et al.,* 2001; Oleksiewicz *et al.,* 2000). In the current study, we found that the same is not true for the ORF5 and the ORF7 sequences. It has recently been demonstrated through simulations that even moderate levels of recombination may destroy molecular clocks (Schierup and Hein, 2000). Our finding of apparent recombination in the ORF5+7 data set may therefore explain the rejection of the molecular clock hypothesis for these genes. It might also be of relevance for the rejection of the molecular clock hypothesis that both ORF5 and ORF7 are significantly shorter than ORF3 and that both genes have more selective constraint on the amino acid sequence than the ORF3 gene (data not shown), and hence accumulate fewer substitutions.

In conclusion, by using data from two ORFs, we were able to detect recombination in European type PRRSV

TABLE 3

Details of the European Type PRRSV Isolates Used

Isolate	Country	Abbreviation	Isolation date	GenBank Accession No.		Journal reference
92V58	Belgium	1-BE	02/92	ORF5: AY035900	ORF7: AY035945	This study
AV30	Belgium	2-BE	06/92	ORF5: AY035901	ORF7: AY035946	This study
111/92	Denmark	1-DK	09/04/92	ORF5–7: AY035944		This study
12654	Denmark	2-DK	19/12/95	ORF5: AY035902	ORF7: AY035947	This study
12770/95	Denmark	3-DK	21/12/95	ORF5: AY035903	ORF7: AY035948	This study
12985	Denmark	4-DK	04/01/96	ORF5: AY035904	ORF7: AY035949	This study
14474B	Denmark	5-DK	04/03/96	ORF5: AY035905	ORF7: AY035950	This study
18794	Denmark	6-DK	26/03/93	ORF5: AY035906	ORF7: AY035951	This study
20567 A	Denmark	7-DK	26/03/97	ORF5: AY035907	ORF7: AY035952	This study
21191	Denmark	8-DK	09/05/97	ORF5: AY035908	ORF7: AY035953	This study
228 A	Denmark	9-DK	16/11/93	ORF5: AY035909	ORF7: AY035954	This study
24554/97	Denmark	10-DK	19/11/97	ORF5: AY035910	ORF7: AY035955	This study
25434/98	Denmark	11-DK	15/01/98	ORF5: AY035911	ORF7: AY035956	This study
28639/98	Denmark	12-DK	16/06/98	ORF5: AY035912	ORF7: AY035957	This study
32-10/92	Denmark	13-DK	17/03/92	ORF5: AY035913	ORF7: AY035958	This study
340-1	Denmark	14-DK	09/01/94	ORF5: AY035914	ORF7: AY035959	This study
361-4	Denmark	15-DK	18/01/94	ORF5: AY035915	ORF7: AY035960	This study
48/92-1	Denmark	16-DK	20/03/92	ORF5: AY035916	ORF7: AY035961	This study
5767-6	Denmark	17-DK	10/02/95	ORF5: AY035917	ORF7: AY035962	This study
5A	France	1-FR		ORF5: U40697		Suarez *et al.*, 1996
6A	France	2-FR		ORF5: U40698		Suarez *et al.*, 1996
8D	France	3-FR		ORF5: U40699		Suarez *et al.*, 1996
II/91	France	4-FR	11/91		ORF7: Z92526	Le Gall *et al.*, 1998
IV/92	France	5-FR	09/92		ORF7: Z92528	Le Gall *et al.*, 1998
IX/95	France	6-FR	01/95		ORF7: Z92529	Le Gall *et al.*, 1998
SDRPII7A	France	7-FR	29/11/91	ORF5: AY035918	ORF7: AY035963	This study
SDRPIV4A	France	8-FR	23/09/92	ORF5: AY035919	ORF7: AY035964	This study
SDRPV4A	France	9-FR	17/05/93	ORF5: AY035920	ORF7: AY035965	This study
V/93	France	10-FR	05/93		ORF7: Z92706	Le Gall *et al.*, 1998
VI/93	France	11-FR	05/93		ORF7: Z92707	Le Gall *et al.*, 1998
2.25	Germany	1-DE	1993	ORF5: AY035921	ORF7: Z92538	Le Gall *et al.*, 1998 and this study
2.35	Germany	2-DE	1993	ORF5: AY035922	ORF7: AY035966	This study
2.46	Germany	3-DE	1993	ORF5: AY035923	ORF7: AY035967	This study
2.72	Germany	4-DE	1993	ORF5: AY035924	ORF7: Z92537	Le Gall *et al.*, 1998 and this study
2.96	Germany	5-DE	1993	ORF5: AY035925	ORF7: AY035968	This study
1/93	Italy	1-IT	19/01/93	ORF5: AY035926	ORF7: AY035969	This study
1142/97	Italy	2-IT	17/02/97	ORF5–7: AY035941		This study
1751/93	Italy	3-IT	23/02/93	ORF5: AY035927	ORF7: AY035970	This study
1828	Italy	4-IT		ORF5: AY035928	ORF7: AY035971	This study
1999/93	Italy	5-IT	23/06/93	ORF5: AY035929	ORF7: AY035972	This study
2029/97	Italy	6-IT	08/04/97	ORF5: AY035930	ORF7: AY035973	This study
2156	Italy	7-IT		ORF5: U40696	ORF7: AY035974	Suarez *et al.*, 1996 and this study
2481/97	Italy	8-IT	24/04/97	ORF5: AY035931	ORF7: AY035975	This study
2567/96	Italy	9-IT	07/05/96	ORF5: AY035932	ORF7: AY035976	This study
3391/93	Italy	10-IT	24/11/93	ORF5–7: AY035942		This study
3943/96	Italy	11-IT	21/06/96	ORF5: AY035933	ORF7: AY035977	This study
7571/96	Italy	12-IT	14/11/96	ORF5–7: AY035943		This study
974/98	Italy	13-IT	10/02/98	ORF5: AY035934	ORF7: AY035978	This study
Boxmeer 10	The Netherlands	1-NL		ORF2–7: L04493		Conzelmann *et al.*, 1993
Lelystad	The Netherlands	2-NL	1991	ORF1–7: M96262		Meulenberg *et al.*, 1993
NL2.2	The Netherlands	3-NL	15/11/91		ORF7: Z92533	Le Gall *et al.*, 1998
NL3.1	The Netherlands	4-NL	1992	ORF5: U40695		Suarez *et al.*, 1996
NL4.1	The Netherlands	5-NL			ORF7: Z92534	Le Gall *et al.*, 1998
L56/2/91	Spain	1-SP	26/10/91	ORF5: AY035935	ORF7: AY035979	This study
2228	Spain	2-SP	1992	ORF5: U40688		Suarez *et al.*, 1996
3211	Spain	3-SP	1992	ORF5: U40689		Suarez *et al.*, 1996
4606	Spain	4-SP	1991	ORF5: U40690		Suarez *et al.*, 1996
5710	Spain	5-SP	1991	ORF5: U40702		Suarez *et al.*, 1996
5999	Spain	6-SP	1991	ORF5: U40691		Suarez *et al.*, 1996
65/2/91	Spain	7-SP	25/06/91	ORF5: AY035936	ORF7: AY035980	This study
705	Spain	8-SP	1993	ORF5: U40692		Suarez *et al.*, 1996

TABLE 3—*Continued*

Isolate	Country	Abbreviation	Isolation date	GenBank Accession No.		Journal reference
L51/2/92	Spain	9-SP		ORF5: AY035937	ORF7: Z92531	Le Gall *et al.,* 1998 and this study
Olot/91	Spain	10-SP		ORF2–7:X92942		Duran *et al.,* 1997
P035	Spain	11-SP	1995	ORF5: U40694		Suarez *et al.,* 1996
BE1	United Kingdom	1-UK	18/02/93		ORF7: L77913	Drew *et al.,* 1995
H2-D768	United Kingdom	2-UK	30/10/91	ORF5: AY035938	ORF7: AY035981	This study
H3	United Kingdom	3-UK	21/06/91		ORF7: L77915	Drew *et al.,* 1995
HA1	United Kingdom	4-UK	18/06/92		ORF7: L77917	Drew *et al.,* 1995
L1-D767	United Kingdom	5-UK	14/02/92	ORF5: AY035939	ORF7: AY035982	This study
L2	United Kingdom	6-UK	5/10/92		ORF7: L77919	Drew *et al.,* 1995
LE1	United Kingdom	7-UK	11/05/92		ORF7: L77921	Drew *et al.,* 1995
NO1	United Kingdom	8-UK	25/03/92		ORF7: L77923	Drew *et al.,* 1995
NY3-D769	United Kingdom	9-UK	22/10/92	ORF5: AY035940	ORF7: Z92536	Le Gall *et al.,* 1998 and this study
NY4	United Kingdom	10-UK	22/06/94		ORF7: L77925	Drew *et al.,* 1995
Ox1	United Kingdom	11-UK			ORF7: L77927	Drew *et al.,* 1995

field isolates. We also demonstrated that the notion of European type PRRSV being less diverse than the North American genotype is misleading. Furthermore, we have shown that the genetic diversity in European type PRRSV is skewed geographically and that Italy in particular seems to hold more genetic diversity than the remainder of Western Europe. While our study represents the largest compilation of European type PRRSV sequences, only eight European countries were sampled. Given the great practical importance of PRRSV genetic diversity for diagnostic and vaccination (Meng, 2000), and the anticipated increase in trading across the European borders due to the eastward opening of the European Communities (EC), it seems highly relevant to continue genetic characterization of PRRSV isolates from other European countries.

MATERIALS AND METHODS

Virus isolation

All Danish and Italian PRRSV isolates originated from clinical material submitted to the Danish Veterinary Institute for Virus Research (DVIVR) or to the Istituto Zooprofilattico Sperimentale, Laboratorio Virologia Specializzata, Italy. PRRSV isolates originating from Belgium, France, Germany, the Netherlands, Spain, and the United Kingdom were obtained through an EC Concerted Action on PRRSV serodiagnostic (AIR3-CT92-0399). All isolates were grown in primary porcine pulmonary alveolar macrophages (PPAM) and stored at −80°C.

RT-PCR and DNA sequencing

Total RNA was extracted from cell culture material by binding to silica particles in guanidine thiocyanate. RNA extraction and RT-PCR amplification of ORF5 and ORF7 were done as described previously (Oleksiewicz *et al.,* 1998). Four of the isolates were also selected for ORF6 sequencing. Briefly, cDNA was produced using a cocktail of PRRSV-specific primers and subsequently used as template in separate PCR reactions specific for ORF5, ORF6, and ORF7. The PCR primers for ORF5 (5′CAATGAGGTGGGCIACAACC and 5′TATGTIATGCTAAAGGCTAGCAC), ORF6 (5′AGGACTTCGGCTGAGCAATGG and 5′ATCAGGCGCACTGTATGAGCAAC), and for ORF7 (5′GCCCCTGCCGAICAC and 5′TCGCCCTAATTGAATAGGTGA) were also used for cycle sequencing of gel-purified (not cloned) amplicons using the BigDye cycle sequencing kit as recommended by the manufacturer (Applied Biosystems, Nærum, Denmark). To ensure complete sequence overlap, ORF6 amplicons were in addition sequenced using the primer 5′TGATTGACTGGCTGGCCATTCC. Individual sequences were determined by electrophoresis on an ABI prism 310 Genetic Analyzer (Applied Biosystems) and contigs were generated using SeqMan II (DNASTAR, Madison, WI). All sequences were deposited in GenBank with Accession Nos. AY035900–AY035982. Table 3 lists details of the European type isolates used in the study. The American type sequences in the present study were all the ORF5 and all the ORF7 sequences which were available through GenBank as of January 2001, a total of 142 ORF5 sequences and 90 ORF7 sequences. A complete list of GenBank Accession Nos. for the North American type sequences used in the analysis can be obtained from the authors.

Analysis of genetic diversity

To characterize the distribution of genetic diversity, we used two measures, namely the average pairwise genetic distance (π) and the maximum pairwise genetic distance (D_{max}). Both measures were calculated as uncorrected distances in units of substitutions per nucleotide site. $\pi = [2/n(n+1)] \sum_{i<j} \pi_{ij}$, where i and j are index numbers of isolates, n is the total number of isolates, π_{ij} is the proportion of different nucleotides between sequence i and j (Nei, 1987), and $D_{max} = \max(\pi_{ij})$. These measures were derived for both ORF5 and ORF7 using

sequences from Europe and North America for which the complete ORFs were available.

It would be preferable to correct for sampling variance when comparing statistics describing pairwise genetic distances. However, due to the nonrandom and clustered sampling of the isolates used in this study it was not possible to perform such a correction. Especially the average pairwise genetic distance between isolates is expected to be sensitive to biases in the sampling scheme and should be compared with caution when sampling variance is not corrected. The maximum pairwise genetic distance only describes the genetic extremes of the populations, but it has a clearly defined biological meaning and is expected to quickly reach an asymptotic level as the sampling size increases, even when samples are clustered geographically (Saunders *et al.,* 2001). The maximum pairwise genetic distance should therefore be less sensitive to sampling variance between populations.

Phylogenetic analysis

Phylogenetic analysis was performed for ORF5 and 7 separately, using all available sequences including partial ORF sequences available in GenBank. With the PAUP* software (Swofford, 2000) we constructed neighbor-joining trees using HKY (Hasegawa *et al.,* 1985) corrected distances and also using parsimony genealogies. Phylogenetic support of the branching in the genealogies was investigated by bootstrapping the data 1000 times.

Molecular clock test

A genealogy-based likelihood-ratio test for a molecular clock was performed for the 38 European type sequences with known precise isolation times (Table 3). This was done using the TipDate software (Rambaut, 2000).

Analysis for chimeric sequences

Based on conflicting result from the ORF5 and the ORF7 genealogies, a potential sign of recombination, isolates were selected and the intervening ORF6 sequences were determined. The resulting new ORF5–7 sequences were analyzed for potential sites of recombination using a sliding windows approach as implemented in SimPlot (Lole *et al.,* 1999).

ACKNOWLEDGMENTS

The technical assistance of Preben Norman and Frederikke Lihme is acknowledged. The authors also thank the participating laboratories in the EC Concerted Action AIR3-CT92-0399 for contributing virus isolates, in particular, Prof. Dr. M. Pensaert, Dr. E. Albina, Prof. Dr. H.-J. Thiel, Dr. G. Wensvoort, Dr. R. R. Bellsola, and Dr. D. J. Paton. This study was partially supported by a grant (BIOT99-2) from the Danish Ministry of Food, Agriculture, and Fishery given to Torben Storgaard. Roald Forsberg was supported by Grant 51-00-0392 from the Danish Natural Science Research Council.

REFERENCES

Allende, R., Kutish, G. F., Laegreid, W., Lu, Z., Lewis, T. L., Rock, D. L., Friesen, J., Galeota, J. A., Doster, A. R., and 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.

Allende, R., Lewis, T. L., Lu, Z., Rock, D. L., Kutish, G. F., Ali, A., Doster, A. R., and Osorio, F. A. (1999). North American and European porcine reproductive and respiratory syndrome viruses differ in non-structural protein coding regions. *J. Gen. Virol.* **80,** 307–315.

Andreyev, V. G., Wesley, R. D., Mengeling, W. L., Vorwald, A. C., and Lager, K. M. (1997). Genetic variation and phylogenetic relationships of 22 porcine reproductive and respiratory syndrome virus (PRRSV) field strains based on sequence analysis of open reading frame 5. *Arch. Virol.* **142,** 993–1001.

Bøtner, A., Nielsen, J., and Bille-Hansen, V. (1994). Isolation of porcine reproductive and respiratory syndrome (PRRS) virus in a Danish swine herd and experimental infection of pregnant gilts with the virus. *Vet. Microbiol.* **40,** 351–360.

Bøtner, A., Strandbygaard, B., Sørensen, K. J., Have, P., Madsen, K. G., Madsen, E. S., and Alexandersen, S. (1997). Appearance of acute PRRS-like symptoms in sow herds after vaccination with a modified live PRRS vaccine. *Vet. Rec.* **141,** 497–499.

Cavanagh, D. (1997). Nidovirales: A new order comprising Coronaviridae and Arteriviridae. *Arch. Virol.* **142,** 629–633.

Christopher-Hennings, J., Holler, L. D., Benfield, D. A., and Nelson, E. A. (2001). Detection and duration of porcine reproductive and respiratory syndrome virus in semen, serum, peripheral blood mononuclear cells, and tissues from Yorkshire, Hampshire, and Landrace boars. *J. Vet. Diagn. Invest.* **13,** 133–142.

Conzelmann, K. K., Visser, N., Van Woensel, P., and Thiel, H. J. (1993). Molecular characterization of porcine reproductive and respiratory syndrome virus, a member of the arterivirus group. *Virology* **193,** 329–339.

Drew, T. W., Lowings, J. P., and Yapp, F. (1997). Variation in open reading frames 3, 4 and 7 among porcine reproductive and respiratory syndrome virus isolates in the UK. *Vet. Microbiol.* **55,** 209–221.

Drew, T. W., Meulenberg, J. J., Sands, J. J., and Paton, D. J. (1995). Production, characterization and reactivity of monoclonal antibodies to porcine reproductive and respiratory syndrome virus. *J. Gen. Virol.* **76,** 1361–1369.

Forsberg, R., Oleksiewicz, M. B., Petersen, A.-M. K., Hein, J., Bøtner, A., and Storgaard, T. (2001). A molecular clock dates the common ancestor of European-type porcine reproductive and respiratory syndrome virus at more than 10 years before the emergence of disease. *Virology* **289,** 174–179.

Gagnon, C. A., and Dea, S. (1998). Differentiation between porcine reproductive and respiratory syndrome virus isolates by restriction fragment length polymorphism of their ORFs 6 and 7 genes. *Can. J. Vet. Res.* **62,** 110–116.

Goldberg, T. L., Hahn, E. C., Weigel, R. M., and Scherba, G. (2000). Genetic, geographical and temporal variation of porcine reproductive and respiratory syndrome virus in Illinois. *J. Gen Virol.* **81,** 171–179.

Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol Evol.* **22,** 160–174.

Indik, S., Valicek, L., Klein, D., and Klanova, J. (2000). Variations in the major envelope glycoprotein GP5 of Czech strains of porcine reproductive and respiratory syndrome virus. *J. Gen. Virol.* **81,** 2497–2502.

Kapur, V., Elam, M. R., Pawlovich, T. M., and Murtaugh, M. P. (1996). Genetic variation in porcine reproductive and respiratory syndrome virus isolates in the midwestern United States. *J. Gen. Virol.* **77,** 1271–1276.

Keffaber, K. K. (1989). Reproductive failure of unknown etiology. *Am. Assoc. Swine Practitioners* **1,** 1–10.

Le Gall, A., Legeay, O., Bourhy, H., Arnauld, C., Albina, E., and Jestin, A. (1998). Molecular variation in the nucleoprotein gene (ORF7) of the porcine reproductive and respiratory syndrome virus (PRRSV). *Virus Res.* **54,** 9–21.

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W., and Ray, S. C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* **73,** 152–160.

Mardassi, H., Mounir, S., and Dea, S. (1994). Identification of major differences in the nucleocapsid protein genes of a Quebec strain and European strains of porcine reproductive and respiratory syndrome virus. *J. Gen. Virol.* **75,** 681–685.

Meng, X. J. (2000). Heterogeneity of porcine reproductive and respiratory syndrome virus: Implications for current vaccine efficacy and future vaccine development. *Vet. Microbiol.* **74,** 309–329.

Meng, X. J., Paul, P. S., Halbur, P. G., and Lum, M. A. (1995a). Phylogenetic analyses of the putative M (ORF 6) and N (ORF 7) genes of porcine reproductive and respiratory syndrome virus (PRRSV): Implication for the existence of two genotypes of PRRSV in the U.S.A. and Europe. *Arch. Virol.* **140,** 745–755.

Meng, X. J., Paul, P. S., Halbur, P. G., and Morozov, I. (1995b). Sequence comparison of open reading frames 2 to 5 of low and high virulence United States isolates of porcine reproductive and respiratory syndrome virus. *J. Gen. Virol.* **76,** 3181–3188.

Mengeling, W. L., Vorwald, A. C., Lager, K. M., Clouser, D. F., and Wesley, R. D. (1999). Identification and clinical assessment of suspected vaccine-related field strains of porcine reproductive and respiratory syndrome virus. *Am. J. Vet. Res.* **60,** 334–340.

Morozov, I., Meng, X. J., and Paul, P. S. (1995). Sequence analysis of open reading frames (ORFs) 2 to 4 of a U.S. isolate of porcine reproductive and respiratory syndrome virus. *Arch. Virol.* **140,** 1313–1319.

Murtaugh, M. P., Faaberg, K. S., Laber, J., Elam, M., and Kapur, V. (1998). Genetic variation in the PRRS virus. *Adv. Exp. Med. Biol.* **440,** 787–794.

Nei, M. (1987). "Molecular Evolutionary Genetics." Columbia Univ. Press, New York.

Nelsen, C. J., Murtaugh, M. P., and Faaberg, K. S. (1999). Porcine reproductive and respiratory syndrome virus comparison: Divergent evolution on two continents. *J. Virol.* **73,** 270–280.

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

Nielsen, J., Bøtner, A., Bille-Hansen, V., Oleksiewicz, M. B., and Storgaard, T. (2002). Experimental inoculation of late term pregnant sows with a field isolate of porcine reproductive and respiratory syndrome vaccine-derived virus. *Vet. Microbiol.* **84,** 1–13.

Oleksiewicz, M., Bøtner, A., Madsen, K. G., and 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.

Oleksiewicz, M. B., Bøtner, A., Toft, P., Grubbe, T., Nielsen, J., Kamstrup, S., and Storgaard, T. (2000). Emergence of porcine reproductive and respiratory syndrome virus deletion mutants: Correlation with the porcine antibody response to a hypervariable site in the ORF 3 structural glycoprotein. *Virology* **267,** 135–140.

Pirzadeh, B., Gagnon, C. A., and Dea, S. (1998). Genomic and antigenic variations of porcine reproductive and respiratory syndrome virus major envelope GP5 glycoprotein. *Can. J. Vet. Res.* **62,** 170–177.

Plana Duran, J., Climent, I., Sarraseca, J., Urniza, A., Cortes, E., Vela, C., and Casal, J. I. (1997). Baculovirus expression of proteins of porcine reproductive and respiratory syndrome virus strain Olot/91. Involvement of ORF3 and ORF5 proteins in protection. *Virus Genes.* **14,** 19–29.

Rambaut, A. (2000). Estimating the rate of molecular evolution: Incorporating non-contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics* **16,** 395–399.

Saunders, I. W., Taveré, S., and Watterson, G. A. (2001). On the genealogy of nested subsamples from a haploid population. *Adv. Appl. Probab.* **16,** 471–491.

Schierup, M. H., and Hein, J. (2000). Recombination and the molecular clock. *Mol. Biol. Evol.* **17,** 1578–1579.

Snijder, E. J., and Meulenberg, J. J. (1998). The molecular biology of arteriviruses. *J. Gen. Virol.* **79,** 961–979.

Storgaard, T., Oleksiewicz, M., and Bøtner, A. (1999). Examination of the selective pressures on a live PRRS vaccine virus. *Arch. Virol.* **144,** 2389–2401.

Suárez, P., Zardoya, R., Martin, M. J., Prieto, C., Dopazo, J., Solana, A., and Castro, J. M. (1996). Phylogenetic relationships of European strains of porcine reproductive and respiratory syndrome virus (PRRSV) inferred from DNA sequences of putative ORF-5 and ORF-7 genes. *Virus Res.* **42,** 159–165.

Suárez, P., Zardoya, R., Prieto, C., Solana, A., Tabarés, E., Bautista, J. M., and Castro, J. M. (1994). Direct detection of the porcine reproductive and respiratory syndrome (PRRS) virus by reverse polymerase chain reaction (RT-PCR). *Arch. Virol.* **135,** 89–99.

Swofford, D. L. (2000). PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, MA.

van Vugt, J. J. F. A., Storgaard, T., Oleksiewicz, M. B., and 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.

Wensvoort, G., Terpstra, C., Pol, J. M. A., ter Laak, E. A., Bloemraad, M., de Kluyver, E. P., Kragten, C., van Buiten, L., den Besten, A., Wagenaar, F., Broeckhuijsen, J. M., Moonen, P. L. J. M., Zetstra, T., de Boer, E. A., Tibben, H. J., de Jong, M. F., van't Veld, P., Groenland, G. J. R., van Gennep, J. A., Voets, M., Verheijden, J. H. M., and Bramskamp, J. (1991). Mystery swine disease in the Netherlands: The isolation of Lelystad virus. *Vet. Q.* **13,** 121–130.

Wesley, R. D., Mengeling, W. L., Lager, K. M., Clouser, D. F., Landgraf, J. G., and Frey, M. L. (1998). Differentiation of a porcine reproductive and respiratory syndrome virus vaccine strain from North American field strains by restriction fragment length polymorphism analysis of ORF 5. *J. Vet. Diag. Invest.* **10,** 140–144.

Wiuf, C., Christensen, T., and Hein, J. (2001). A simulation study of recombination detection methods. *Mol. Biol. Evol.* **18,** 1929–1939.

Wrights, S. (1931). Evolution in Mendelian populations. *Genetics* **16,** 97–159.

Yuan, S., Nelsen, C. J., Murtaugh, M. P., Schmitt, B. J., and Faaberg, K. S. (1999). Recombination between North American strains of porcine reproductive and respiratory syndrome virus. *Virus Res.* **61,** 87–98.