Genetic, geographical and temporal variation of porcine reproductive and respiratory syndrome virus in Illinois

Tony L. Goldberg, Edwin C. Hahn, Ronald M. Weigel and Gail Scherba

University of Illinois, Department of Veterinary Pathobiology, 2001 South Lincoln Avenue, Urbana, IL 61820, USA

Porcine reproductive and respiratory syndrome virus (PRRSV) ORF5 gene sequences were generated by RT–PCR from 55 field isolates collected in Illinois and eastern Iowa. Spatial and temporal patterns of genetic variation in the virus were examined on a local geographical scale in order to test the hypothesis that the genetic similarity of PRRSV isolates (measured as their percentage pairwise ORF5 nucleotide similarity) was positively correlated with their geographical proximity. Levels of genetic variability in the Illinois/eastern Iowa PRRSV sample were similar to levels of variability seen across broader geographical regions within North America. The genetic similarity of isolates did not correlate with their geographical distance. These results imply that the movement of PRRSV onto farms does not generally occur via distance-limited processes such as wind or wildlife vectors, but more typically occurs via the long-distance transport of animals or semen. Genetic distances between PRRSV isolates collected from the same farms at different times increased as the time separating the collection events increased. This result implies rapid movement of new genetic types of PRRSV into and out of farms. PRRSV ORF5 displayed a pattern of third-codon-position diversity bias that was not evident in a geographically comparable sample of pseudorabies virus (a swine alphaherpesvirus) gC gene sequences. This result provides evidence that PRRSV ORF5 is experiencing stabilizing selection against structural novelty. Despite high genetic variability at all geographical levels, PRRSV ORF5 nevertheless contained potentially antigenic regions that were invariant at the amino acid level. These regions should make effective vaccine targets if they prove to be immunogenic.

Introduction

Porcine reproductive and respiratory syndrome is a highly prevalent and economically important viral disease of swine (Done *et al.*, 1996; Zimmerman *et al.*, 1997*a*). The causative agent, porcine reproductive and respiratory syndrome virus (PRRSV), is a member of the *Arteriviridae* (Conzelmann *et al.*, 1993; Meulenberg *et al.*, 1993; Plagemann & Moennig, 1992), which, along with the *Toroviridae* and *Coronaviridae*, make up the order Nidovirales (de Vries *et al.*, 1997). The single-stranded, positive-sense RNA genome of PRRSV is approximately 15 kb in length, containing seven ORFs (Meulenberg *et al.*, 1993).

Genetic studies of PRRSV field isolates have documented that the virus is highly diverse biologically. Despite similar clinical manifestations, PRRSV isolates from Europe and North America are immunologically distinct (Drew *et al.*, 1995; Wensvoort *et al.*, 1992) and genetically divergent (Mardassi *et al.*, 1994; Murtaugh *et al.*, 1995; Nelsen *et al.*, 1999). Within North America, isolates also show marked diversity, both antigenically (Yoon *et al.*, 1997) and genetically (Kapur *et al.*, 1996; Murtaugh *et al.*, 1998). However, to date, the population-genetic structure of PRRSV has not been studied on a local geographical scale. It is currently unclear whether local variants of PRRSV exist, or whether PRRSV genetic variation is distributed evenly within North America or Europe.

Examining the genetic structure of PRRSV on a finer geographical scale than has previously been attempted is important for two reasons. Firstly, the existence of local variants of PRRSV would have immediate implications for control of the disease. For example, it would indicate that vaccines developed against type strains of the virus may not be

widely effective. It would also imply that integrated disease control strategies may need to be 'custom tailored' to specific areas. Secondly, geographically focused studies of PRRSV genetics have the potential to elucidate patterns of spread of the virus across the landscape. PRRSV is persistently shed in semen (Christopher-Hennings et al., 1995; Swenson et al., 1994). The shipment of semen for artificial insemination may therefore be an important mode of transmission of PRRSV between farms (Swenson et al., 1995; Yeager et al., 1993). Alternatively, PRRSV may travel between farms even in the absence of human intervention. Zimmerman et al. (1993, 1997 b) have shown that PRRSV-infected waterfowl carry and shed live infectious virus, implying that PRRSV may travel between farms in animal vectors. It also has been suggested that airborne transmission is important for the spread of PRRSV between nearby farms (De Jong et al., 1991; Komjin et al., 1991; Le Poitier et al., 1995). A positive association between the genetic similarity of PRRSV isolates and their geographical proximity would support the hypothesis that distance-limited processes (e.g. vectors, wind) are important for the spread of PRRSV. The absence of such an association would support the hypothesis that the spread of PRRSV results primarily from human intervention.

This report describes the first investigation of PRRSV genetic variability on a local geographical level. Specifically, it examines the population-genetic structure of 55 PRRSV isolates collected from Illinois and eastern Iowa. It first describes the pattern of overall genetic diversity in this local sample and compares it with that of an expanded sample containing isolates from diverse geographical regions within North America. It next describes associations between the genetic, geographical and temporal similarity of isolates, inferring from them patterns of spread of the virus across the local landscape over space and time.

Methods

■ PRRSV isolates ($n = 62$) were obtained from the Illinois Animal Disease Laboratories in Galesburg, IL, USA. Samples had been submitted by veterinarians for diagnostic testing between 15 April 1997 and 1 July 1998. Clinical information, including the farm of origin and date of submission, accompanied each sample. All submitting veterinarians were contacted by telephone to confirm the locations of farms. Locations of farms (exact longitudes and latitudes) were obtained from street addresses by using maps and geographical databases at the Illinois Geological Survey. Locations of farms were plotted by using the computer program MapInfo (MapInfo Corp., Troy, NY, USA). Distances between farms were calculated directly from longitudes and latitudes by using the computer program The R Package (Legendre & Vaudour, 1991).

Each PRRSV isolate was propagated on MARC-145 cells (Kim et al., 1993). RT–PCR was then performed directly on the virus isolates without additional steps to extract RNA. For the few samples that did not show a visible product by using this method ($n = 6$), total RNA was isolated by using TRIZOL reagent (Gibco BRL) and RT–PCR was subsequently performed to obtain a DNA product.

The ORF5 gene was selected for this study because of its documented high variability among field isolates within the United States (Kapur et al., 1996). Two primers were designed on the basis of published PRRSV gene sequences. Primers 3F (5′ GAGACCATGAGGTGGGCAAC 3′) and 5R (5′ CGCCAAAAGCACCTTTTGT 3′) anneal to the genome at locations approximately 40 bases up- and downstream of the ORF5 gene, respectively. Primer 5R was used for RT and both primers were used for the subsequent PCR and sequencing. RT was performed by using primer 5R and SuperScript II reverse transcriptase (Life Technologies) according to the recommendations of the manufacturer. PCR was performed on 2 μl of the RT reaction mixture in a buffer containing 10 mM Tris–HCl, pH 8·3, 50 mM KCl, 1·5 mM MgCl$_2$, 0·2 mM dNTPs, 0·2 pM each of primer 3F and 5R and 2·5 U AmpliTaq Gold DNA polymerase (Perkin Elmer) in a volume of 50 μl. After an initial incubation at 95 °C for 8 min, the reactions were subjected to 35 cycles of PCR as follows: 95 °C for 1 min, 59 °C for 1 min and 72 °C for 1 min followed by a terminal 6 min extension at 72 °C and an indefinite soak at 4 °C. The products were subjected to electrophoresis in 1% SeaPlaque low-melting-point agarose gels (FMC BioProducts). The 720 bp PCR product was purified from the gel by using QiaQuick gel extraction kits (Qiagen) and was then submitted for automated fluorescent sequencing at the University of Illinois Biotechnology Center. All products were sequenced in both directions with primers 3F and 5R. Ambiguous bases were resolved by repeated sequencing of both DNA strands.

All newly generated sequences were aligned by hand with published North American PRRSV sequences. The North American sequences were aligned with the European PRRSV reference sequence by using the CLUSTAL W computer program (Thompson et al., 1994). All alignments were confirmed manually and by comparison with deduced amino acid sequences.

Seventeen pseudorabies virus (PrV) partial gC gene sequences from Illinois were also included in this study for comparative purposes. These PrV isolates are described elsewhere (Scherba et al., 1999). PCR of a 788 bp portion of the 3′ end of PrV gC was conducted with primers 1L (5′ GAAGGGCTCACCGAAGAGGAC 3′) and 2U (5′ GTTTCCTG-ATTCACGCCCACGC 3′) in a buffer containing 10 mM Tris–HCl, pH 8·3, 50 mM KCl, 1·5 mM MgCl$_2$, 0·2 mM dNTPs, 0·4 pM each of primer 1L and 2U, 10% glycerol and 0·25 U AmpliTaq Gold DNA polymerase in a volume of 50 μl. After an initial incubation at 95 °C for 6 min, the reactions were subjected to 35 cycles of PCR as follows: 95 °C for 1 min, 68 °C for 1 min and 72 °C for 1 min followed by a terminal 6 min extension at 72 °C and an indefinite soak at 4 °C. PCR products were purified and sequenced by using primers 1L and 2U as described above.

To detect quantitative associations between genetic and geographical distance, genetic distances (proportions of differing nucleotide positions) were compared with geographical distances (km) between all pairs of sequences in the Illinois/eastern Iowa sample. Also included in these analyses were all pairwise temporal distances between sequences (days separating the submission of isolates), with the expectation that time of isolation may be a confounding variable. Bivariate and partial Mantel tests of matrix correlation (Mantel, 1967) were conducted for this purpose by using the computer program The R Package (Legendre & Vaudour, 1991). The methodology of Smouse et al. (1986) was used, with a standardized form of the Mantel Z statistic (r) and probabilities computed from 10 000 random permutations of the matrices (Hope, 1968). Spatial autocorrelation analysis was performed by using the computer program AIDA (Bertorelle & Barbujani, 1995) with great-circle geographical distances and with 95% confidence intervals calculated by using 1000 random permutations of the data. Phylogenetic and bootstrap analyses were performed by using the computer programs PHYLIP (Felsenstein, 1990) and MacClade (Maddison & Maddison, 1992). All statistical results were considered significant at the $P = 0·05$ level.

Results

A total of 55 PRRSV ORF5 sequences were generated, representing 48 farms in 28 Illinois counties and two locations in eastern Iowa (Fig. 1). Information on the dates of sample collection and the locations of farms was obtained for all 55 isolates sequenced and for an additional seven isolates that were not sequenced. Within the sample, ORF5 displayed considerable genetic variation; 227/603 (37·7%) of the nucleic acid positions were polymorphic, as were 77/201 (38·3%) of the deduced amino acid positions.

Table 1 presents codon-position-specific genetic diversity measures for the 55 ORF5 sequences from Illinois and eastern Iowa. For purposes of comparison, analogous genetic diversity measures for the 17 wild-type PrV gC gene partial DNA sequences from Illinois were included. PrV is an alpha-herpesvirus of swine and the gC gene, like the ORF5 gene of PRRSV, encodes an immunologically important N-glyco-sylated transmembrane protein (Ben-Porat *et al.*, 1986; Zucker-mann *et al.*, 1990).

PRRSV ORF5 in Illinois was approximately seven times more variable than PrV gC in Illinois at the nucleic acid level (Table 1). This difference was also reflected at the structural level; PRRSV ORF5 was approximately six times more variable than PrV gC in its amino acid sequence.
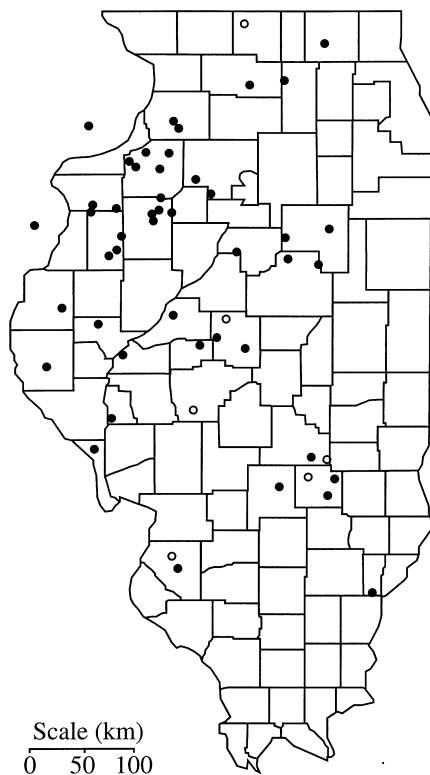

Fig. 1. Map of Illinois showing the locations of farms from which PRRSV isolates were obtained. Filled circles represent farms from which both PRRSV ORF5 RNA sequences and dates of collection were obtained. Open circles represent farms for which dates of sample collection, but not RNA sequences, were obtained.

Table 1. Comparative genetic diversity of PRRSV and PrV

The data consisted of 55 PRRSV ORF5 sequences, including 53 from Illinois and two from eastern Iowa, generated in the present study, and 17 newly generated PrV partial gC gene sequences from Illinois. Pairwise genetic differences (% nucleotide diversity), given as means ± SEM for each estimate, were calculated according to Nei (1987).

Residues analysed	n	Variable positions	Genetic diversity (%)
PRRSV ORF5			
Codon position 1	201	51	4·10 ± 0·63
Codon position 2	201	54	4·96 ± 0·73
Codon position 3	201	122	10·62 ± 1·42
All nucleotides	603	227	6·51 ± 0·78
Amino acids	201	76	7·82 ± 1·08
PrV gC			
Codon position 1	262	9	0·74 ± 0·90
Codon position 2	263	13	0·93 ± 0·99
Codon position 3	263	11	0·81 ± 0·93
All nucleotides	788	34	0·91 ± 0·56
Amino acids	262	17	1·39 ± 1·20

The patterns of codon-specific genetic diversity also differed between PRRSV ORF5 and PrV gC. PRRSV ORF5 showed a significant diversity bias at the third codon position, whereas PrV gC did not (Table 1).

Seventeen additional ORF5 sequences obtained from the literature (Collins, 1998; Kapur *et al.*, 1996; Meng *et al.*, 1995; Meulenberg *et al.*, 1993) were analysed to compare the genetic structure of PRRSV ORF5 in the Illinois/eastern Iowa sample with that from a wider geographical area. These published sequences represent virus isolates collected from locations as far south as Kansas and as far north as Québec. The nucleotide diversity of the 16 North American isolates in this sample (all codon positions included) was 8·41 ± 2·02%. This was not significantly different from the nucleotide diversity of the 55 samples from Illinois and eastern Iowa (6·51 ± 0·78%; Table 1).

Fig. 2 shows a phylogenetic tree constructed from the 55 Illinois/eastern Iowa sequences and the 17 published ORF5 sequences. In this tree, all North American sequences clustered apart from the European type strain (Lelystad). Within the North American isolates, two major clades appeared. The first clade (from the 'Prime Pac vaccine' sequence to isolate 'ISU55' on Fig. 2) was genetically diverse and consisted of sequences from Illinois, other midwestern states and Canada and the Prime Pac vaccine (Schering–Plough) sequence. The second clade (from isolate 'PPSKS1A' to the bottom of the tree) was less diverse and consisted primarily of the Illinois isolates, interspersed with sequences from elsewhere, and the RespPRRS vaccine (NOBL Laboratories Inc.) sequence. This latter grouping was supported by a bootstrap value of 78%,
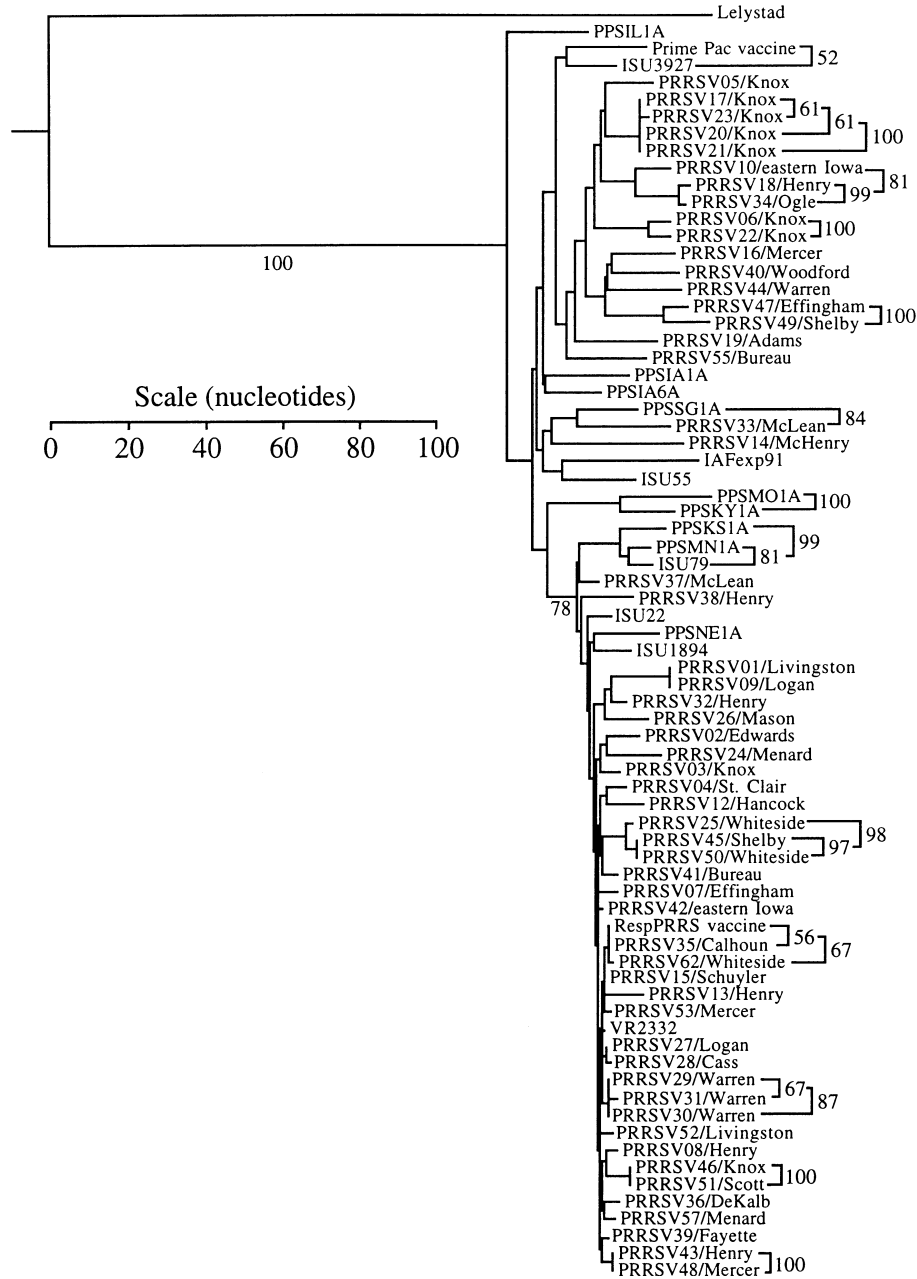
Fig. 2. Phylogenetic tree of PRRSV isolates. The tree was constructed by using the neighbour-joining algorithm of Saitou & Nei (1987), a maximum-likelihood distance correction and a transition–transversion ratio of 3·4:1, estimated from the data. Names of sequences generated in the present study begin with 'PRRSV' and are followed by the names of their county of origin within Illinois. Branch lengths are proportional to genetic distance. Numbers below branches and to the right of indicated groupings are bootstrap values (%) obtained from 100 bootstrap replicates of the data by using the computer package PHYLIP (Felsenstein, 1985). Only bootstrap values $\geqslant 50\%$ are shown.

indicating that this clade was reasonably well supported statistically.

Fig. 3 shows the results of the Mantel tests of matrix correlation, used for testing associations between genetic, geographical and temporal similarity within the Illinois/eastern Iowa sample. No significant relationship existed between genetic distance and either geographical distance or temporal distance for the isolates in this sample. A marginally significant association did exist, however, between geographical distance and temporal distance. In other words, isolates collected for diagnostic testing at approximately the same times tended to be geographically close to one another, irrespective of their genetic similarity.

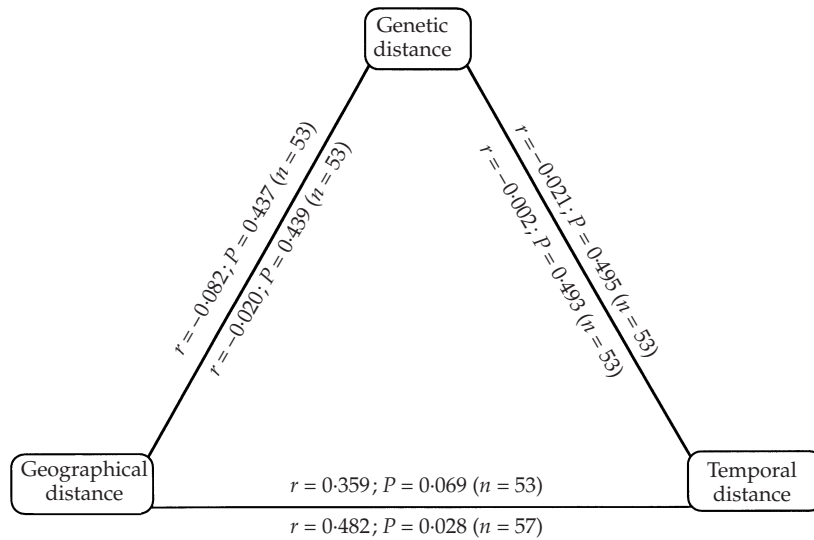To detect the presence of any non-linear association

Fig. 3. Relationships between genetic distance, geographical distance and temporal distance for Illinois and eastern Iowa PRRSV isolates. Genetically identical isolates from the same farm ($n = 2$) were collapsed to avoid pseudoreplication. Bivariate Mantel tests were conducted for each pair of variables and are shown on the outside of the triangle. Values on the inside of the triangle represent partial Mantel correlations between two matrices, conditional on the value of the third. Sample sizes vary due to missing data.
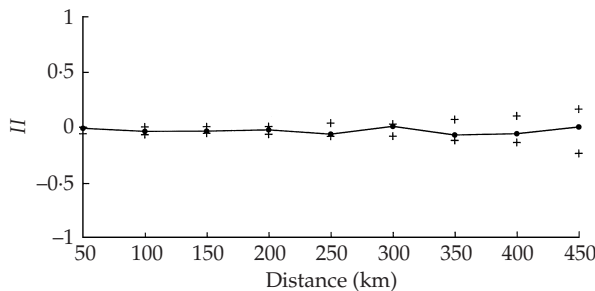

Fig. 4. Spatial autocorrelogram based on 55 PRRSV RNA sequences. An autocorrelation index for DNA analysis (*II*) was calculated for each of nine distance classes. Crosses represent the upper and lower 95% confidence intervals around the null distribution of *II* at each distance class, calculated by using 1000 random permutations of the data. Significant spatial autocorrelation was not observed at any distance class.
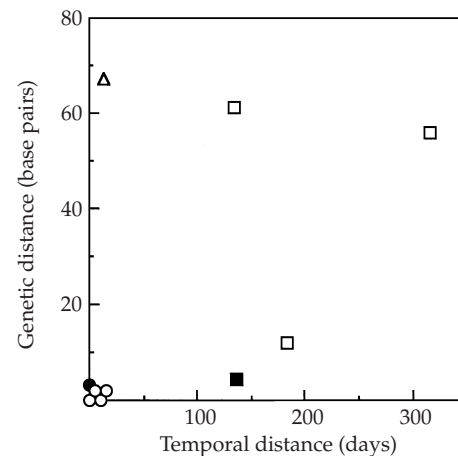

Fig. 5. Pairwise genetic distance versus pairwise temporal distance for 12 pairs of PRRSV isolates collected from the same farms at different times. Farms were designated as farms 1 ($\bigcirc$), 2 ($\bullet$), 3 ($\square$), 4 ($\blacksquare$) and 5 ($\triangle$).

between genetic and geographical distance of ORF5 sequences, a spatial autocorrelogram was constructed (Fig. 4). No significant spatial autocorrelation existed at any distance class. Geographical proximity does not therefore significantly predict how genetically similar two PRRSV ORF5 sequences are likely to be, regardless of their geographical distance.

A retrospective statistical-power analysis was performed to determine the sensitivity of the correlation analyses described above for detecting any association between geographical proximity and genetic similarity of PRRSV isolates. Assuming an alpha level of 0·05 (one-tailed), a statistical power of 0·8 and a sample size of 55, it was calculated that the analyses described above were sensitive enough to detect an effect of geographical proximity on genetic similarity of 10·4% or greater.

Within the major clades on the phylogenetic tree shown in Fig. 2, some county-level clustering of PRRSV isolates was observed (note, for example, the clustering of isolates PRRSV 17, 20, 21 and 23, all from Knox county). This implies that local evolutionary diversification may in fact be occurring despite a

general pattern of geographical panmixis. To investigate this possibility further, the tree in Fig. 2, based on genetic data, was compared with an analogous tree (not shown) constructed directly from a matrix of pairwise geographical distances separating PRRSV isolates. Genetic changes (nucleotide positions) were then mapped onto each tree by using the computer program MacClade (Maddison & Maddison, 1992). The length of each tree was calculated as the total number of reconstructed nucleotide position changes along its branches.

The overall length of the phylogenetic tree constructed from genetic distances was 565 nucleotide changes. This was substantially less than the length of the tree constructed from geographical distances (1120 nucleotide changes). By the criterion of parsimony (Swofford & Olsen, 1990), the genetic tree provided a markedly better explanation of the sequence data than did the tree in which geographical clustering was
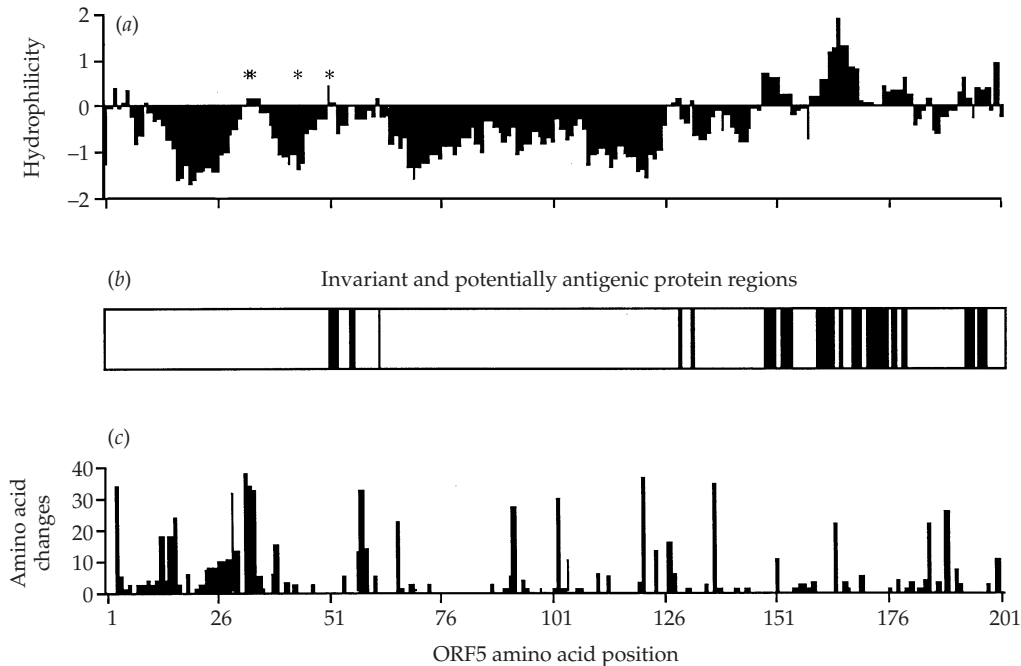
Fig. 6. Genetic variability of the ORF5 protein. (*a*) The hydrophilicity score of each amino acid position in the ORF5 consensus sequence is displayed and was calculated by using the algorithm of Hopp & Woods (1981) with an averaging group length of six amino acid residues. Hydrophilic regions deviate in the positive direction. Asterisks (*) indicate the locations of putative *N*-linked glycosylation sites. (*b*) Vertical bars indicate regions that are both hydrophilic (and thus potentially antigenic) and invariant at the amino acid level. (*c*) The amount of amino acid variation along ORF5 is displayed for the 71 United States field isolates analysed in this study (see Fig. 2). Bars represent the number of sequences differing from the consensus at each position.

presumed. This phylogeny-based test therefore confirmed that geographical sorting of PRRSV genotypes did not occur to an appreciable degree in this dataset.

The Illinois/eastern Iowa sample contained 12 pairs of isolates collected from the same farms at different times. This enabled an examination of the relationship between genetic distance and temporal distance for PRRSV ORF5 on these farms. The results of this examination (Fig. 5) indicate that diverse genetic types can exist simultaneously on a farm, as shown by the large genetic difference between isolates collected on farm 5. Nevertheless, samples from some farms (e.g. farm 4) seemed comparatively homogeneous. Overall, as time increases, farms are increasingly likely to harbour genetic variants of PRRSV different from those they harboured previously. This relationship was statistically significant (Spearman's $r = 0.61$; $P = 0.043$), although the validity of this analysis is questionable since points represent pairwise comparisons and are not independent.

Finally, the distribution of genetic diversity across the ORF5 protein was investigated for all 71 North American sequences included in the previous phylogenetic analysis. Together, these sequences contained 282/603 (46·1%) polymorphic nucleic acid positions and 94/201 (46·8%) polymorphic amino acid positions and had an overall nucleotide diversity of $7.21 \pm 0.78\%$. Fig. 6 demonstrates that, despite the high intrinsic variability of ORF5 within North American

isolates, there are, nevertheless, potentially antigenic protein regions that are also genetically invariant at the amino acid level. Statistically, these predicted antigenic regions were no more genetically variable than were non-antigenic regions ($t = 0.84$; $P = 0.403$).

Discussion

This study documents a high level of genetic variability in the ORF5 gene of PRRSV isolates collected from Illinois and eastern Iowa. The amount of genetic variability in this local sample was not significantly different from that of a sample of North American isolates from a much wider geographical area. Phylogenetic analyses indicated that PRRSV isolates do not generally sort into location-specific clades. In support of this observation, no significant association was documented between the geographical proximity of PRRSV isolates and the similarity of their ORF5 genes. Statistical-power analysis indicated that distance-limited transmission and local evolution, if they occur at all, could account for at most only 10·4% of the variability in PRRSV genetic diversity observed in this dataset.

These observations suggest that PRRSV does not typically move between farms via distance-limited processes such as wind or wildlife vectors. Despite the biological plausibility of these modes of inter-farm transmission (Albina, 1997; Zimmer-

man *et al.*, 1997*b*), the movement of PRRSV isolates directly from farm to nearby farm does not explain the geographical pattern of genetic variability in this sample. Rather, the absence of statistical association between geographical and genetic distance suggests that PRRSV typically moves from farm to farm via long-distance processes, such as the transport of animals or semen. This conclusion is supported by the observations of Weigel *et al.* (1998) that the purchase of semen for artificial insemination significantly increases the risk that a farm in Illinois will test positive serologically for PRRSV.

Some local evolution of PRRSV genotypes may, in fact, be occurring, as suggested by the phylogenetic clustering of certain PRRSV isolates at the county level on the tree in Fig. 2. However, local evolution is an unparsimonious explanation of the current distribution of PRRSV genetic types across the landscape. A phylogenetic tree constructed from geographical distances provided a substantially less parsimonious explanation of the nucleotide sequence data than did a tree constructed from genetic distances. This lack of evidence for strong effects of local ORF5 evolution may indicate either a lack of sufficient evolutionary time for such effects to manifest themselves or a 'swamping' of local evolution by long-distance mixing, or both.

Genetic similarity between pairs of PRRSV isolates, while not correlated with geographical proximity, correlated positively with temporal proximity. Given the previous analyses, this observation cannot be explained as a result of concentrically spreading 'outbreaks' of PRRSV across the landscape. If such outbreaks were occurring, the relationships between genetics and time and between genetics and geography would also have been positive and significant. The documented positive relationship between geographical and temporal distance could reflect the fact that veterinarians from specific regions may diagnose PRRS or submit samples for diagnostic testing more often at certain times than at others. This, in turn, could result from local environmental factors that precipitate simultaneous outbreaks.

This study documents that genetic distance increases with temporal distance for pairs of PRRSV isolates collected from the same farms at different times. This relationship could potentially result from the *in situ* mutation and evolution of PRRSV isolates on individual farms. Alternatively, time may simply increase the likelihood that new variants of PRRSV will arrive on a farm. Regardless of mechanism, rates of genetic change appear to differ among farms, as do levels of initial PRRSV genetic variability. PRRSV appears to be genetically homogeneous over time on some farms (e.g. farm 4 in Fig. 5), while being diverse on others (e.g. farms 3 and 5 in Fig. 5). The reasons for these differences are unclear, but may reflect differences in management practices on these farms. Accurate quantification of this phenomenon will require longitudinal sampling of PRRSV isolates collected from a larger number of pre-selected farms.

Despite the high variability of PRRSV isolates within regions and over time on individual farms, no evidence was found that structural novelty in PRRSV ORF5 is under positive selection. Antigenic regions of ORF5 were no more diverse than non-antigenic regions. Third-codon-position changes predominated in ORF5, as would be expected under stabilizing selection (Li & Graur, 1991). This is in contrast to PrV gC, which, despite lower overall variability, was as variable in its first and second codon positions as in its third codon positions. The overall difference in diversity between PRRSV and PrV may be attributed to the fact that the polymerase enzymes of DNA viruses such as PrV possess higher intrinsic fidelity than those of RNA viruses such as PRRSV, leading to comparatively higher rates of mutation and evolution in RNA viruses (Holland *et al.*, 1982). Nevertheless, PrV gC, and not PRRSV ORF5, appears to tolerate structural variation. Preliminary analyses in our laboratory indicate that structural novelty within the PrV gC glycoprotein may, in fact, be positively selected for as a response to adaptive host immune pressure.

Variability along the ORF5 protein was also not evenly distributed. Fig. 6 indicates that both hypervariable and invariant regions of the ORF5 protein exist. Among these are regions that are both antigenic and genetically invariant. These regions are probably critical to the function of the protein and thus are logical targets for vaccine development. Future research is warranted into the immunogenicity and specific functions of these regions.

The results and interpretations of this study are dependent on the particular evolutionary dynamics of the ORF5 gene of PRRSV. While ORF5 is the most variable of the PRRSV ORFs, its usefulness for reconstructing specific phylogenetic relationships has been called into question because of intragenic recombination (Murtaugh *et al.*, 1998). Gene sequences of non-recombining segments of the PRRSV genome (e.g. ORF6; Murtaugh *et al.*, 1998) may yield topologies in phylogenetic trees more reliable than those presented here. Nevertheless, the frequency of recombination across the PRRSV genome has recently been estimated by Wesley *et al.* (1999) to be approximately 7%. If this is the case, then the frequency of recombination within ORF5, which represents approximately 4% of the PRRSV genome, should be just 0·28%, assuming that recombination is equally probable anywhere in the genome. This rate of recombination is low and suggests that fewer than one of the 55 isolates included in the present study is a recombinant.

The results of this study also depend on the sample of PRRSV isolates analysed, which is biased. Only that subset of PRRSV isolates that were submitted by veterinarians for diagnostic testing was examined. Analysis of an unbiased sample of PRRSV isolates (e.g. collected by diagnostic screening of a randomly chosen set of farms) might yield relationships different from those documented here. Similarly, analysis of PRRSV isolates collected over a longer time span may reveal hitherto undocumented temporal trends in the genetic variability of the virus.

Nevertheless, the present sample represents a clinically important subset of the PRRSV population. To the extent that the conclusions of this study can be generalized from, they indicate that long-distance movement is a far stronger determinant of PRRSV ORF5 genetic diversity than is localized evolution. Similar to PRRSV across North America (Murtaugh *et al.*, 1998), PRRSV on a local geographical scale appears to represent a single, intermixing virus population. If so, vaccines and other strategies to control the spread of PRRSV need not be tailored to specific geographical regions. Control of the future spread of PRRSV between US farms should focus on the inter- and intra-state shipment of animals and animal products. Future research into the spatial, temporal and evolutionary dynamics of other PRRSV genes should help to confirm the validity and generality of these conclusions.

References

Albina, E. (1997). Epidemiology of porcine reproductive and respiratory syndrome (PRRS): an overview. *Veterinary Microbiology* **55**, 309–316.

Ben-Porat, T., DeMarchi, J. M., Lomniczi, B. & Kaplan, A. S. (1986). Role of glycoproteins of pseudorabies virus in eliciting neutralizing antibodies. *Virology* **154**, 325–334.

Bertorelle, G. & Barbujani, G. (1995). Analysis of DNA diversity by spatial autocorrelation. *Genetics* **140**, 811–819.

Christopher-Hennings, J., Nelson, E. A., Hines, R. J., Nelson, J. K., Swenson, S. L., Zimmerman, J. J., Chase, C. L., Yaeger, M. J. & Benfield, D. A. (1995). Persistence of porcine reproductive and respiratory syndrome virus in serum and semen of adult boars. *Journal of Veterinary Diagnostic Investigation* **7**, 456–464.

Collins, J. (1998). Interpreting PRRSV sequencing data. In *Proceedings of the Allen D. Leman Swine Conference*, pp. 1–4. University of Minnesota, USA.

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

De Jong, M. F., Cromwijk, W. & Van't Veld, P. (1991). The new pig disease: epidemiology and production losses in the Netherlands. In *Report of a Seminar on the New Pig Disease* (*PRRS*), pp. 9–19. Brussels, Belgium.

de Vries, A. A. F., Horzinek, M. C., Rottier, P. J. M. & de Groot, R. J. (1997). The genome organization of the Nidovirales: similarities and differences between arteri-, toro-, and coronaviruses. *Seminars in Virology* **8**, 33–47.

Done, S. H., Paton, D. J. & White, M. E. C. (1996). Porcine reproductive and respiratory syndrome (PRRS): a review, with emphasis on pathological, virological and diagnostic aspects. *British Veterinary Journal* **152**, 153–174.

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

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Felsenstein, J. (1990). PHYLIP: phylogeny inference package, version 3.5c. Department of Genetics, University of Washington, Seattle, WA, USA.

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. & VandePol, S. (1982). Rapid evolution of RNA genomes. *Science* **215**, 1577–1585.

Hope, A. C. A. (1968). A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society, Series B* **30**, 582–598.

Hopp, T. P. & Woods, K. R. (1981). Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Sciences, USA* **78**, 3824–3828.

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

Kim, H. S., Kwang, J., Yoon, I. J., Joo, H. S. & Frey, M. L. (1993). Enhanced replication of porcine reproductive and respiratory syndrome (PRRS) virus in a homogeneous subpopulation of MA-104 cell line. *Archives of Virology* **133**, 477–483.

Komjin, R. E., Van Klink, E. G. M. & Van Der Sande, W. J. H. (1991). The possible effect of weather conditions on the spread of the 'new' pig disease in the Netherlands. In *Report of a Seminar on the New Pig Disease* (*PRRS*), pp. 28–31. Brussels, Belgium.

Legendre, P. & Vaudour, A. (1991). *The R Package: Multidimensional Analysis, Spatial Analysis*. Montréal: Université de Montréal Département de sciences biologiques.

Le Poitier, M. F., Blanquefort, P., Morvan, E. & Albina, E. (1995). Results of a control program for PRRS in the French area 'Pays de Loire.' In *Proceedings of the 2nd International Symposium on PRRS*, p. 34. Copenhagen, Denmark.

Li, W.-H. & Graur, D. (1991). *Fundamentals of Molecular Evolution*. Sunderland, MA: Sinauer Associates.

Maddison, W. P. & Maddison, D. R. (1992). MacClade: Analysis of Phylogeny and Character Evolution. Sunderland, MA: Sinauer Associates.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**, 209–220.

Mardassi, H., Mounir, S. & Dea, S. (1994). Identification of major differences in the nucleocapsid protein genes of a Québec strain and European strains of porcine reproductive and respiratory syndrome virus. *Journal of General Virology* **75**, 681–685.

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

Meulenberg, J. J. M., Hulst, M. M., de Meijer, E. J., Moonen, P. L. J., den Besten, A., de Kluyver, E. P., Wensvoort, G. & Moormann, R. J. M. (1993). Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV. *Virology* **192**, 62–72.

Murtaugh, M. P., Elam, M. R. & Kakach, L. T. (1995). Comparison of the structural protein coding sequences of the VR-2332 and Lelystad virus strains of the PRRS virus. *Archives of Virology* **140**, 1451–1460.

Murtaugh, M. P., Faaberg, K. S., Laber, J., Elam, M. & Kapur, V. (1998). Genetic variation in the PRRS virus. *Advances in Experimental Medicine and Biology* **440**, 787–794.

Nei, M. (1987). *Molecular Evolutionary Genetics.* New York: Columbia University Press.

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

Plagemann, P. G. W. & Moennig, V. (1992). Lactate dehydrogenase-elevating virus, equine arteritis virus, and simian hemorrhagic fever virus: a new group of positive-strand RNA viruses. *Advances in Virus Research* **41**, 99–192.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406–425.

Scherba, G., Wiemers, J. F., Siegel, A. M., Jin, L., Austin, C. C., Bowman, L., Redeford, B., Johnston, N. A. & Weigel, R. M. (1999). Application of a quantitative algorithm to restriction endonuclease analysis of Aujeszky's disease (pseudorabies) virus from a geographically localized outbreak. *Journal of Veterinary Diagnostic Investigation* **11**, 423–431.

Smouse, P. E., Long, J. C. & Sokal, R. R. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology* **35**, 627–632.

Swenson, S. L., Hill, H. T., Zimmerman, J. J., Evans, L. E., Landgraf, J. G., Wills, R. W., Sanderson, T. P., McGinley, J. M., Brevik, A. K., Ciszewski, D. K. & Frey, M. L. (1994). Excretion of porcine reproductive and respiratory syndrome virus in semen after experimentally induced infection in boars. *Journal of the American Veterinary Medical Association* **204**, 1943–1948.

Swenson, S. L., Hill, H. T., Zimmerman, J. J., Evans, L. E., Wills, R. W., Yoon, K.-J., Schwartz, K. J., Althouse, G. C., McGinley, J. M. & Brevik, A. K. (1995). Preliminary assessment of an inactivated PRRS virus vaccine on the excretion of virus in semen. *Swine Health and Production* **3**, 244–247.

Swofford, D. L. & Olsen, G. J. (1990). Phylogeny reconstruction. In *Molecular Systematics*, pp. 411–501. Edited by D. M. Hillis & C. Moritz. Sunderland, MA: Sinauer Associates.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680.

Weigel, R., Firkins, L. & Scherba, G. (1998). Risk factors for infection of Illinois swine herds with porcine reproductive and respiratory syndrome virus (PRRSV). In *Conference of Research Workers in Animal Diseases* (*CRWAD*), p. 44. Chicago, USA.

Wensvoort, G., de Kluyver, E. P., Luijtze, E. A., den Besten, A., Harris, L., Collins, J. E., Christianson, W. T. & Chladek, D. (1992). Antigenic comparison of Lelystad virus and swine infertility and respiratory syndrome (SIRS) virus. *Journal of Veterinary Diagnostic Investigation* **4**, 134–138.

Wesley, R. D., Mengeling, W. L., Lager, K. M., Vorwald, A. C. & Roof, M. B. (1999). Evidence for divergence of restriction fragment length polymorphism patterns following *in vivo* replication of porcine reproductive and respiratory syndrome virus. *American Journal of Veterinary Research* **60**, 463–467.

Yeager, M. J., Prieve, T., Collins, J., Christopher-Hennings, J., Nelson, E. & Benfield, D. (1993). Evidence for the transmission of porcine reproductive and respiratory syndrome (PRRS) virus in boar semen. *Swine Health and Production* **1**(5), 7–9.

Yoon, K.-J., Wu, L.-L., Zimmerman, J. J. & Platt, K. B. (1997). Field isolates of porcine reproductive and respiratory syndrome virus (PRRSV) vary in their susceptibility to antibody dependent enhancement (ADE) of infection. *Veterinary Microbiology* **55**, 277–287.

Zimmerman, J., Yoon, K. Y., Pirtle, E. C., Sanderson, T. J., Hill, H. T., Wills, R. W., McGinley, M. J. & Brevik, A. (1993). Susceptibility of four avian species to PRRS virus. In *Proceedings of the Annual Meeting of the Livestock Conservation Institute*, pp. 107–108. St Louis, USA.

Zimmerman, J. J., Yoon, K.-J., Wills, R. W. & Swenson, S. L. (1997 *a*). General overview of PRRSV: a perspective from the United States. *Veterinary Microbiology* **55**, 187–196.

Zimmerman, J. J., Yoon, K.-J., Pirtle, E. C., Wills, R. W., Sanderson, T. J. & McGinley, M. J. (1997 *b*). Studies of porcine reproductive and respiratory syndrome (PRRS) virus infection in avian species. *Veterinary Microbiology* **55**, 329–336.

Zuckermann, F. A., Zsak, L., Mettenleiter, T. C. & Ben-Porat, T. (1990). Pseudorabies virus glycoprotein gIII is a major target antigen for murine and swine virus-specific cytotoxic T lymphocytes. *Journal of Virology* **64**, 802–812.