Edinburgh Research Explorer

Genomic variation in macrophage-cultured European porcine reproductive and respiratory syndrome virus Olot/91 revealed using ultra-deep next generation sequencing

VIROLOGY JOURNAL

SHORT REPORT   **Open Access**

Genomic variation in macrophage-cultured European porcine reproductive and respiratory syndrome virus Olot/91 revealed using ultra-deep next generation sequencing

Zen H Lu[1], Alexander Brown[1], Alison D Wilson[1], Jay G Calvert[2], Monica Balasch[3], Pablo Fuentes-Utrilla[4], Julia Loecherbach[4], Frances Turner[4], Richard Talbot[4], Alan L Archibald[1] and Tahar Ait-Ali[1*]

Abstract

Background: Porcine Reproductive and Respiratory Syndrome (PRRS) is a disease of major economic impact worldwide. The etiologic agent of this disease is the PRRS virus (PRRSV). Increasing evidence suggest that microevolution within a coexisting quasispecies population can give rise to high sequence heterogeneity in PRRSV.

Findings: We developed a pipeline based on the ultra-deep next generation sequencing approach to first construct the complete genome of a European PRRSV, strain Olot/9, cultured on macrophages and then capture the rare variants representative of the mixed quasispecies population. Olot/91 differs from the reference Lelystad strain by about 5% and a total of 88 variants, with frequencies as low as 1%, were detected in the mixed population. These variants included 16 non-synonymous variants concentrated in the genes encoding structural and nonstructural proteins; including Glycoprotein 2a and 5.

Conclusion: Using an ultra-deep sequencing methodology, the complete genome of Olot/91 was constructed without any prior knowledge of the sequence. Rare variants that constitute minor fractions of the heterogeneous PRRSV population could successfully be detected to allow further exploration of microevolutionary events.

Keywords: PRRSV, Microevolution, Variant spectra, Ultra-deep next generation sequencing

---

Findings

Porcine Reproductive and Respiratory Syndrome virus (PRRSV) is the causative agent of a significant disease of the domestic pig (*Sus scrofa*) with global consequences. The severity of PRRSV infection ranges from subclinical to lethal and it affects pigs in both growing and reproductive stages. The virus has a positive-sense 15 kb RNA genome and its genetic diversity has been well characterised within and between European and North American strains [1]. Extensive viral genetic heterogeneity may have contributed towards the observed variations between PRRSV isolates and clones in term of virulence, interactions with the immune system, and antigenic properties of

viral proteins. Such a broad diversity indeed poses serious challenges to diagnostics and control measures.

Most previous studies of PRRSV genetic diversity have been restricted to the ORF5 and ORF7 sequences of type 2 "North American-like" viruses that also include the Asian variants. Only 14 of the 303 completed PRRSV genomes in Genbank belong to genotype 1. Furthermore, studies have shown that PRRSV mutates rapidly and multiple intra-strain variants can coexist in individually infected pigs [2]. The extensive genetic diversity displayed by PRRSV and other RNA viruses such as HIV and influenza reflects the error prone nature of RNA polymerases, which lack a proofreading function [3,4].

To identify PRRSV quasispecies, previous studies have employed conventional methodologies including reverse-transcription, PCR, cloning and Sanger sequencing of a subset of PRRSV structural and non-structural proteins

---

* Correspondence: tahar.aitali@roslin.ed.ac.uk
[1]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Edinburgh EH25 9RG, UK
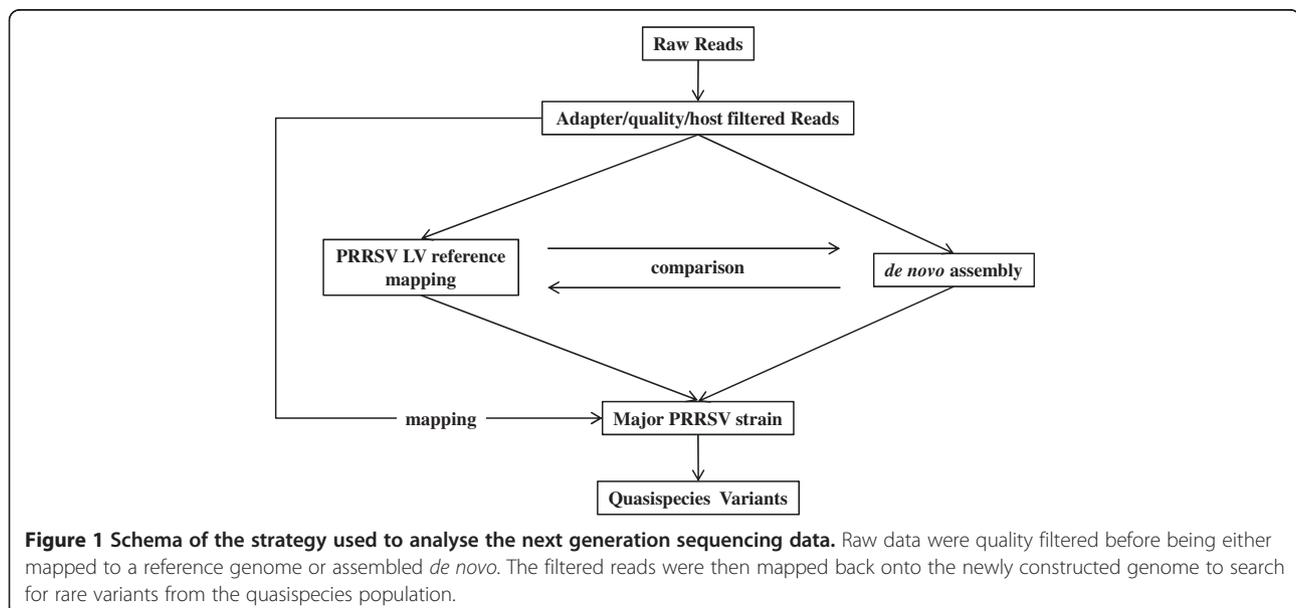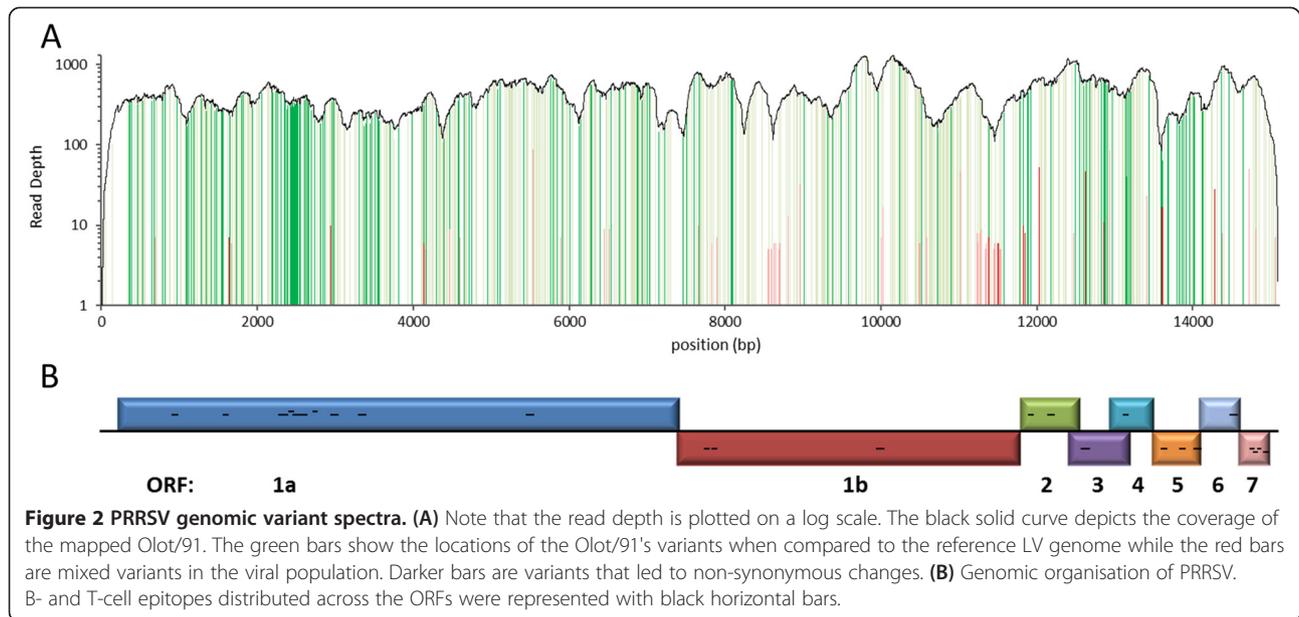Full list of author information is available at the end of the article

[2,5]. More recently next-generation sequencing (NGS) of fragments generated by long range RT-PCR has been used to characterise multiple PRRSV genomes [6]. However, this approach relies upon prior knowledge of the target sequence and the assumption that the PCR primer binding sites are non-variable. Here we describe an approach which requires no prior knowledge of the target sequences and which should enable the detection of low-frequency nucleotide variants and hence provides a snapshot of the microevolution in the entire viral population.

We analysed the intra-strain sequence diversity of low passage PRRSV Olot/91 strain, passaged exclusively on primary porcine alveolar macrophages (PAM). First reported in Spain in 1991, Olot/91 is the parent strain of the commercial Suvaxyn PRRSV inactivated vaccine which is used in Spain and Portugal. Only a partial sequence of 3,383 nt [GenBank:X92942], that covers ORFs 2-7 and the 3'-UTR of this strain has previously been published [7].

To reduce potential complications with sub-genomic RNA, the virus was grown in porcine alveolar macrophages for 5 days and viral particles were pelleted from cell culture supernatant through a 30% sucrose cushion. Viral RNA was then prepared using BioSprint 96 DNA Blood kit (Qiagen) and sequencing libraries produced using the Illumina TruSeq RNA sequencing library preparation kit, with a modification eliminating the initial mRNA isolation steps. A total of 7.56 million 225 bp reads were generated on a Illumina MiSeq machine. They were analysed following a 2-step strategy whereby the consensus sequence of the major Olot/91 strain was first obtained and then used to call rare variants present in the evolving PRRSV viral population (Figure 1). To account for potential contamination with host DNA, the raw reads were initially filtered against the pig genome (Sscrofa10.2 [8]) using the program Best Match Tagger (v1.1.0) [9] which removed about 88% of the reads. The remaining reads then underwent a stringent quality filtering whereby the first 5 bp of the 5' ends were trimmed to remove any remnant adapter sequences; reads with ambiguous bases and Phred scores lower than 20 were discarded using Sickle (v1.2) [10] and the 3' ends were further scanned for run-through adapter sequences using Scythe (v0.991b) [11]. The final high quality reads (4.3%) were next either mapped against the reference Lelystad (LV) genome [Genbank:M96262] [12] using BWA (v0.6.2) [13]/ Samtools (v0.1.18) [14] or assembled *de novo* with Velvet (v1.2.9) [15]. Variants used to call the 15,111 nt consensus sequence in the mapped major Olot/91 strain were determined as those having allele frequencies higher than 0.7 by the program Lofreq (v0.6.0) [16]. A total of 742 single nucleotide variants were identified that differed from the reference LV strain (Figure 2, Additional file 1: Table S1). Of these, 180 resulted in non-synonymous changes in coding sequences. No insertions or deletions were found. For the *de novo* assembly, contigs from two kmer values of 161 and 195 were merged with Lasergene's SeqMan Pro (v10.1) to yield a single contig of 15,082 nt. Both the mapped consensus sequence and the *de novo* assembled contig were aligned against each other to ensure the accuracy of the genome. These sequences were exactly identical to each other and the full length Olot/91 genome has now been deposited in Genbank with the accession number KF203132. This full length Olot/91 contains 8 SNPs and a 3 nt insertion when compared with the published partial Olot/91 sequence. However, this PAM-passaged virus only shares 96.8% similarity with the MARC-145 cell-adapted PRRSV Olot/91 strain

Figure 1 Schema of the strategy used to analyse the next generation sequencing data. Raw data were quality filtered before being either mapped to a reference genome or assembled *de novo*. The filtered reads were then mapped back onto the newly constructed genome to search for rare variants from the quasispecies population.

Figure 2 PRRSV genomic variant spectra. (A) Note that the read depth is plotted on a log scale. The black solid curve depicts the coverage of the mapped Olot/91. The green bars show the locations of the Olot/91's variants when compared to the reference LV genome while the red bars are mixed variants in the viral population. Darker bars are variants that led to non-synonymous changes. **(B)** Genomic organisation of PRRSV. B- and T-cell epitopes distributed across the ORFs were represented with black horizontal bars.
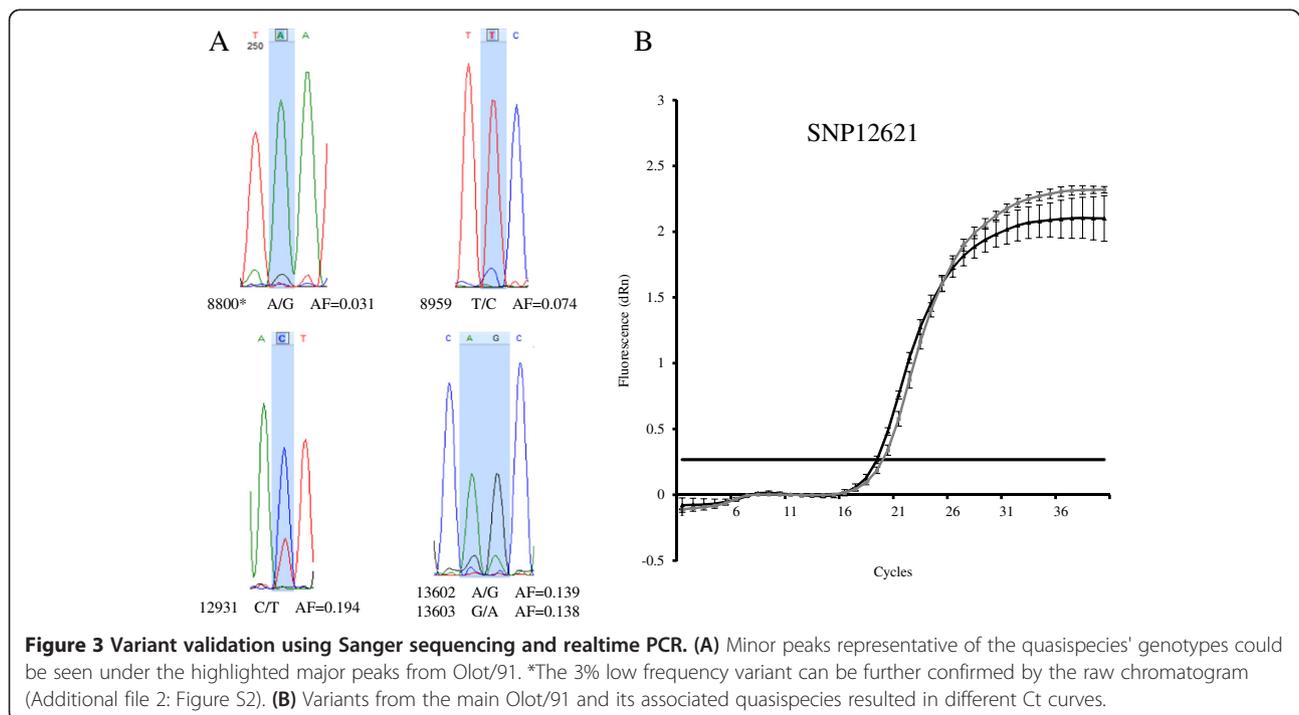
(Genbank:KC862570) [6]. The differences may represent the molecular basis for the adaption to growth on MARC-145 cells which are of African Green Monkey origin [17].

The same set of cleaned reads was then remapped as previously onto the newly generated Olot/91 genome to an average depth of approximately 530×. Eighty-eight low-frequency single-nucleotide variants representing the genotypic makeup of the heterogeneous PRRSV population

were identified, of which 16 represented non-synonymous changes in coding sequences (Figure 2, Additional file 1: Table S1). Among the 22 low frequency variants that overlap with the major Olot/91 versus reference LV strain variants, 17 have evolved into genotypes identical to the reference strain. The frequency of intra-strain variants detectable in this study ranged from 1 to 20% with each supported by a minimum of 3 reads. A subset of these low-frequency variants was validated using two alternative

Figure 3 Variant validation using Sanger sequencing and realtime PCR. (A) Minor peaks representative of the quasispecies' genotypes could be seen under the highlighted major peaks from Olot/91. *The 3% low frequency variant can be further confirmed by the raw chromatogram (Additional file 2: Figure S2). **(B)** Variants from the main Olot/91 and its associated quasispecies resulted in different Ct curves.

methods; Sanger sequencing and real-time PCR. Reverse transcription was performed as described previously [18]. Briefly, one microgram of the Olot/91 RNA was reverse transcribed using a TaqMan kit (Applied Biosystems, Foster City, CA). For Sanger sequencing, PCR products encompassing the variants were sequenced in both directions using BigDye terminator v3.1 Cycle Sequencing kit (Applied Biosystems). Variants with frequencies as low as 3% could be successfully identified by the overlapping peaks (Figure 3A and Additional file 2: Figure S1). Major and minor allele specific real-time PCRs employing Platinum SYBR Green PCR SuperMix UDG (Invitrogen, Paisley, UK) were run on a Stratagene MX3000P (Stratagene). Samples were run in triplicate, beta-actin was utilized as the housekeeping gene and results were calculated as described previously [18]. Again, the main Olot/91 sequence and the variants present in the related mixed population could both be successfully detected (Figure 3B).

While variants in the major Olot/91 strain do congregate on known highly hypervariable B- and T-cell epitopes of type I PRRSV (Figure 2 and Additional file 3: Figure S2) [19-23], most of the variants representative of the probable quasispecies population map to the *nsp12 and gp2a* genes of ORF1b and ORF2 respectively (Figure 2); suggesting a higher microevolutionary rate at this region under the investigated culture conditions. In addition, the predicted short signal peptide of GP2a was found to harbour a high concentration of non-synonymous variants and a potential N-glycosylation site was created from a residue change (S37N) on GP5. Further analysis, like molecular modelling, may be necessary to decipher the potential impacts of these evolving variants have on such functions as interactions between these proteins and other viral or host cellular proteins.

Using ultra-deep NGS we have constructed the complete sequence of the PRRSV Olot/91 genome cultured in their natural host - porcine alveolar macrophages - and identified single-nucleotide variants present in the associated viral population. Further investigation using this methodology will help to establish if a link exists between the microevolutionary dynamics and pathogenesis of PRRSV viral strains. However, it is also important to note that the exact nature of the pathogenesis cannot be truly identified until the viral haplotypes within the quasispecies population can be reconstructed with confidence [24].

Additional files

Additional file 1: Table S1. Annotation of Single-nucleotide variants detected in the mixed Olot/91 population. (A) SNPs detected in comparison to the reference LV strain. (B) Rare variants detected in comparison to the major Olot/91 strain.

Additional file 2: Figure S1. Unprocessed raw chromatogram depicting the unique low-frequency variant in the mixed Olot/91 population. The unique black G peak (circled) stands up among the noise shoulders of A (green) and T (red).

Additional file 3: Figure S2. Sequence Entropy of type I European PRRSV Strains. Multiple sequence alignment of the genomes [Olot/91-GenBank:KF203132, GenBank:M96262, GenBank:GU737264, GenBank:A26843, GenBank:GQ461593, GenBank:FJ349261, GenBank:DQ489311, GenBank:JF802085, GenBank:GU047344, GenBank:GU047345, GenBank:AY588319, GenBank:AY366525, GenBank:GU067771, GenBank:EU076704, GenBank:DQ864705] were generated with MUSCLE and entropy analysed with Hyphy [25]. Higher entropies denote regions with higher variability.

Authors' contributions
ZHL carried out all the bioinformatics analyses and drafted the manuscript. AB participated in the discussion of data analysis. JGC and MB provided the samples. ADW cultured the virus, purified the viral RNA and performed the real time PCR. PFU constructed the RNA libraries and ran the next-generation sequencing. JL and FT prepared the raw Illumina data. RT designed the sequencing run and collaborated in drafting the manuscript. ALA conceived the study and collaborated in drafting the manuscript. TAA conceived/ designed/coordinated the study and drafted the manuscript. All authors read and approved the final manuscript.

Author details
[1]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Edinburgh EH25 9RG, UK. [2]Zoetis Inc.,Global Biologics Research, 333 Portage St, 49007 Kalamazoo, MI, USA. [3]VMRD Olot Zoetis, Ctra. EU Regional Vaccines Group, Camprodon s/n Finca "La Riba", 17813 Vall de Bianya, Girona, Spain. [4]Edinburgh Genomics, University of Edinburgh, Easter Bush, Edinburgh EH25 9RG, UK.

References
1. Murtaugh MP, Stadejek T, Abrahante JE, Lam TT, Leung FC: **The ever-expanding diversity of porcine reproductive and respiratory syndrome virus.** *Virus Res* 2010, **154**:18–30.
2. Goldberg TL, Lowe JF, Milburn SM, Firkins LD: **Quasispecies variation of porcine reproductive and respiratory syndrome virus during natural infection.** *Virology* 2003, **317**:197–207.
3. Lauring AS, Andino R: **Quasispecies theory and the behavior of RNA viruses.** *PLoS Pathogens* 2010, **6**:e1001005.
4. Domingo E, Sheldon J, Perales C: **Viral quasispecies evolution.** *Microbiol Mol Biol Rev* 2012, **76**:159–216.
5. Schommer SK, Kleiboeker SB: **Use of a PRRSV infectious clone to evaluate in vitro quasispecies evolution.** *Adv in Exp Med and Biol* 2006, **581**:435–438.
6. Kvisgaard LK, Hjulsager CK, Fahnøe U, Breum SØ, Ait-Ali T, Larsen LE: **A fast and robust method for full genome sequencing of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) Type 1 and Type 2.** *J Virol Methods* 2013, **193**:697–705.
7. Urniza A, Climent I, Duran J, Corte´s E, Sarraseca J, Casal JI, Vela C: **Baculovirus expression of proteins of porcine reproductive and respiratory syndrome virus strain olot/91. Involvement of ORF3 and ORF5 proteins in protection.** *Virus Genes* 1997, **14**:19–29.
8. Groenen MA, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens HJ, Li S, Larkin DM, Kim H, Frantz

LA, Caccamo M, Ahn H, Aken BL, Anselmo A, Anthon C, Auvil L, Badaoui B, Beattie CW, Bendixen C, Berman D, Blecha F, Blomberg J, Bolund L, Bosse M, Botti S, Bujie Z, et al: **Analyses of pig genomes provide insight into porcine demography and evolution.** Nature 2012, **491**:393–398.

9.   BMTagger: **Best match tagger for removing human reads from metagenomics datasets.** ftp://ftp.ncbi.nlm.nih.gov/pub/agarwala/bmtagger/.

10. **Sickle - a windowed adaptive trimming tool for FASTQ files using quality.** http://github.com/najoshi/sickle.

11. **Scythe - a bayesian adapter trimmer.** http://github.com/vsbuffalo/scythe.

12. Meulenberg JJM, Hulst MM, de Meijer EJ, Moonen PLJM, den Besten A, de Kluyver EP, Wensvoort G, Moormann RJM: **Lelystad Virus, the Causative Agent of Porcine Epidemic Abortion and Respiratory Syndrome (PEARS), Is Related to LDV and EAV.** Virology 1993, **192**:62–72.

13. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform.** Bioinformatics 2009, **25**:1754–1760.

14. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing Subgroup: **The sequence alignment/Map format and SAMtools.** Bioinformatics 2078–2079, **2009**:25.

15. Zerbino DR, Birney E: **Velvet: algorithms for de novo short read assembly using de Bruijn graphs.** Genome Res 2008, **18**:821–829.

16. Wilm A, Aw PP, Bertrand D, Yeo GH, Ong SH, Wong CH, Khor CC, Petric R, Hibberd ML, Nagarajan N: **LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets.** Nucleic Acids Res 2012, **40**:11189–11201.

17. Delrue I, Van Gorp H, Van Doorsselaere J, Delputte PL, Nauwynck HJ: **Susceptible cell lines for the production of porcine reproductive and respiratory syndrome virus by stable transfection of sialoadhesin and CD163.** BMC Biotechnol 2010, **10**:48.

18. Ait-Ali T, Wilson AD, Carre W, Westcott DG, Frossard JP, Mellencamp MA, Mouzaki D, Matika O, Waddington D, Drew TW, Bishop SC, Archibald AL: **Host inhibits replication of European porcine reproductive and respiratory syndrome virus in macrophages by altering differential regulation of type-I interferon transcriptional response.** Immunogenetics 2011, **63**:437–448.

19. Díaz I, Pujols J, Ganges L, Gimeno M, Darwich L, Domingo M, Mateu E: **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 2009, **27**:5603–5611.

20. Oleksiewicz MB, Bøtner A, Normann P: **Porcine B-cells recognize epitopes that are conserved between the structural proteins of American- and European-type porcine reproductive and respiratory syndrome virus.** J Gen Virol 2002, **83**:1407–1418.

21. Costers S, Vanhee M, Van Breedam W, Van Doorsselaere J, Geldhof M, Nauwynck HJ: **GP4-specific neutralizing antibodies might be a driving force in PRRSV evolution.** Virus Res 2010, **154**:104–113.

22. Parida R, Choi I-S, Peterson DA, Pattnaik AK, Laegreid W, Zuckermann FA, Osorio FA: **Location of T-cell epitopes in nonstructural proteins 9 and 10 of type-II porcine reproductive and respiratory syndrome virus.** Virus Res 2012, **169**:13–21.

23. Liao YC, Lin HH, Lin CH, Chung WB: **Identification of cytotoxic T lymphocyte epitopes on swine viruses: multi-epitope design for universal T cell vaccine.** PLoS One 2013, **8**:e84443.

24. Schirmer M, Sloan WT, Quince C: **Benchmarking of viral haplotype reconstruction programmes: an overview of the capacities and limitations of currently available programmes.** Brief Bioinform 2012. doi:10.1093/bib/bbs081.

25. Pond SL, Frost SD, Muse SV: **HyPhy: hypothesis testing using phylogenies.** Bioinformatics 2005, **21**:676–679.