Mol. Cells *27*, 119-123
DOI/10.1007/s10059-009-0013-3

Communication

Molecules and Cells

Identification and Molecular Characterization of PERV Gamma1 Long Terminal Repeats

Jae-Won Huh[1,2], Dae-Soo Kim[1,3], Hong-Seok Ha[1], Kung Ahn[1], Kyu-Tae Chang[2], Byung-Wook Cho[4], and Heui-Soo Kim[1,*]

Porcine endogenous retroviruses (PERVs) gamma1 in the pig genome have the potential to act as harmful factors in xenotransplantation (pig-to-human). Long terminal repeats (LTRs) are known to be strong promoter elements that could control the transcription activity of PERV elements and the adjacent functional genes. To investigate the transcribed PERV gamma1 LTR elements in pig tissues, bioinformatic and experimental approaches were conducted. Using RT-PCR amplification and sequencing approaches, 69 different transcribed LTR elements were identified. And 69 LTR elements could be divided into six groups (15 subgroups) by internal variation including tandem repeated sequences, insertion and deletion (INDEL). Remarkably, all internal variations were indentified in U3 region of LTR elements. Taken together, the identification and characterization of various PERV LTR transcripts allow us to extend our knowledge of PERV and its transcriptional study.

INTRODUCTION

The identification of porcine endogenous retrovirus (PERV) has been a major risk factor of zoonotic infection in pig-to-human xenotransplantation. PERVs have been identified in various different pig strains, including wild pigs (Le Tissier et al., 1997; Moalic et al., 2006). Approximately 50 copies of replication-competent PERVs have been known to exist in the pig genome, and these viruses are transmitted vertically (Akiyoshi et al., 1998; Patience et al., 1997; 2001). Intact proviruses of PERV harbor the structural genes of *gag*, *pol,* and *env,* and are flanked by long terminal repeats (LTRs) at both ends. Recently, PERV gamma1 and 2 sequences were identified (Klymiuk et al., 2006). Three types of PERV-A, -B, and -C belonging to PERV gamma1 have been well-categorized by their envelope protein sequences (Patience et al., 2001; Takeuchi et al., 1998). The PERV-A and -B are capable of infecting human and porcine cell lines (Martin et al., 1998; Wilson et al., 2003). The PERV-C elements do not replicate their progeny in human cells, but can infect porcine cell lines (Takeuchi et al., 1998).

After the discovery of infectious PERV, various efforts have been conducted to control the PERV elements in xenotransplantation. Antiretroviral drugs, RNA interference, and complement control have been used for the control of PERVs (Dieckhoff et al., 2007; Karlas et al., 2004; Takefman et al., 2002). Indeed, some investigators have attempted to find means by which to performed safe xenotransplantation, including the large white pig, Westran pig, and wild pig, which contain smaller amounts of replication-competent PERV sequences (Lee et al., 2002; Mang et al., 2001; Niebert et al., 2002). However, for the achievement of a safe xenotransplantation procedure considering the potential risk derived from PERVs, more investigations on the distribution and structural differences in the pig genome and expression in different pig tissues are necessary.

The long terminal repeat (LTR) is composed of the U3, R, and U5 regions, which are important for the regulation of the transcription of PERV elements (Krach et al., 2001). Thus, the mutations, including deletion, insertion, and nucleotide substitution, could change the transcription regulation activity of a PERV element. And PERV LTR elements were identified to harbor potential hormone-responsive sequences (Quinn and Langford, 2001). In addition, a cellular environment including trans-acting elements (hormones and transcription factors) could be important in terms of control elements for the regulation of PERV. Thirty-nine distinct bp repeats (18 bp and 21 bp) harboring the transcription factor binding sites of NF-Y were also identified as a major key element to control the transcription activity (Denner et al., 2003; Scheef et al., 2001; 2002; Wilson et al., 2003). From the *in vitro* study, the difference in the number of 39 bp repeats indicated the different transcription activities (Scheef et al., 2001). The expression activity of the PERV LTR element could be altered by its methylation status (Huh et al., 2007a).

However, the lack of information on the entire genome of the pig is a big obstacle to investigating the various LTR elements in different pig genomes. Indeed, a majority of LTR studies has been focused on the distinct 39 bp repeats in the U3 region (Denner et al., 2003; Scheef et al., 2001; 2002; Wilson et al., 2003) and specific LTR sequences flanked by the replication-competent PERVs identified by *in vitro* transfection analysis

[1]Division of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Korea, [2]National Primate Research Center, Korea Research Institute of Bioscience and Biotechnology, Chungbuk 363-883, Korea, [3]Korean BioInformation Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-806, Korea, [4]Department of Animal Science, College of Life Sciences, Pusan National University, Miryang 627-706, Korea
*Correspondence: khs307@pusan.ac.kr

Received August 20, 2008; revised November 6, 2008; accepted November 11, 2008; published online February 5, 2009

Keywords: bioinformatics, LTR element, PERV, Pig, xenotransplantation

(Czauderna et al., 2000; Harrison et al., 2004; Krach et al., 2001; Takefman et al., 2002). Although, well-established LTR elements have been investigated in previous studies, we were not able to assess all of the features of the original LTR elements that are transcribed in the pig genome (Denner et al., 2003; Quinn and Langford, 2001; Scheef et al., 2001; 2002; Wilson et al., 2003).

In the present study, we describe the characterization of transcribed LTR elements in three different pigs (hybrid, wild, and domestic pigs) using random selection and sequencing procedures. From the structural analysis, six groups, including 15 subgroups of LTR elements, were identified. Thus, this study provides the wide spectrum of transcribed LTR elements, and could allow the functional analysis and characterization of different subtype of LTR elements.

MATERIALS AND METHODS

Identification of the PERV LTR element and selection of amplified regions for LTR analysis in the pig genome

To identify the various PERV LTR elements, the PERV LTR consensus sequence was retrieved from the Repbase Update (http://www.girinst.org/). Using the consensus sequence of the PERV LTR element, numerous PERV LTR sequences were identified in the high-throughput genomic sequences (HTGS) database of the pig genome by basic local alignment and a search tool (BLAST) program (Altschul et al., 1997; Huh et al., 2007b). From the LTR sequence analysis, highly conserved regions were identified in the U3 and U5 regions. Those regions were used for RT-PCR amplification.

Total RNA isolation and RT-PCR analysis

Total RNA of liver, lung, and brain tissues from hybrid of Korean domestic pig and Yorkshire, Korean domestic pig, and wild pig (*Sus scorfa coreanus*) were isolated using Trizol reagent (Invitrogen). Turbo DNA-free™ kit (Ambion) was used to eliminate DNA contamination from total RNA. To confirm the elimination of DNA contamination from total RNA, We performed PCR amplification of DNase-treated total RNA samples without the reverse-transcription procedure (No-RT experiment). No amplification of PCR products indicated the perfect elimination of DNA contamination of the DNase-treated total RNA samples did not contain the genomic DNA (Fig. 1A) (Sellars et al., 2007). Then, M-MLV (Moloney-Murine Leukemia Virus) reverse transcriptase with an annealing temperature of 42°C was used for the reverse transcription reaction with RNase inhibitor (Promega). The transcript of the PERV LTR element was analyzed by RT-PCR. RT-PCR products were amplified by the primer pairs of PERV-LS (5′- GAT GAA AAT GCA ACC TAA CCC -3′) and PERV-LAS (5′- CCC CAA ATC ACT CAC GAG AA -3′) (Fig. 1B). The primer sequences were designed by the alignment of different PERV LTR elements. As a standard control, ACTB (actin, beta) gene was amplified by the BA-S (5′- GAC TAC CTC ATG AAG ATC-3′, bases 632-649) and BS-AS (5′- GAT CCA CAT CTG CTG GAA -3′, bases 1127-1144) primers from human ACTB (GenBank accession no. NM_001101).

Molecular cloning and sequencing

RT-PCR products were separated on a 1.5% agarose gel, purified with the QIAEX II gel extraction kit (QIAGEN), and cloned into the pGEM-T-easy vector (Promega). The cloned DNA was isolated by the alkali lysis method using the High Pure plasmid isolation kit (Roche). Sequencing services were provided by Macrogen company (Korea). To confirm the sequence accuracy, inaccurate "N" sequences were resequenced. All confirmed nucleotide sequences reported in this paper were deposited in the DDBJ nucleotide sequence databases (supplementary Table S1).

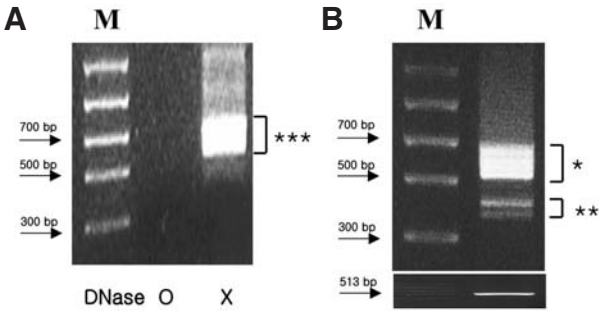

Fig. 1. No-RT experiment (A) and RT-PCR analysis of LTR elements (B). In the No-RT experiment, no-amplication product indicates the efficient elimination of DNA contamination. *** indicates the DNA contamination of DNase-untreated total RNA samples. M indicate the size marker. * indicates the major band and ** indicates the minor band.

Data analyses

For the comparison of different LTR elements, multiple alignment analysis was conducted (http://prodes.toulouse.inra.fr/multalin/multalin.html) and manually corrected using the BioEdit program. Tandem repeat sequences in the LTR element were analyzed by the tandem repeat finder (http://tandem.bu.edu/trf/trf.html). The PERV LTR structures were reconstructed by the analysis of multiple alignments and tandem repeat sequences (Fig. 2).

RESULTS AND DISCUSSION

Amplification and identification of actively transcribed PERV LTR elements

Using a computational approach, we were able to detect 124 different gamma1 LTR sequences in the pig genome. However, 65 copies of the LTR elements were detected in the EST database. Therefore, we attempted to identify the expressed LTR elements in pig tissues derived from a hybrid of the Korean domestic pig and Yorkshire, Korean domestic pig, and wild pig. Using the specific RT-PCR primers and mRNA samples from brain, liver, lung tissues, various PERV LTR products were amplified, and those products were then randomly inserted into the pGEM-T-easy vector for sequencing analysis. Finally, 69 members of the LTR elements transcribed in pig tissues were identified. All of this information is shown in supplementary Table S1.

Structure analysis of PERV LTR elements

In order to draw the structures of the LTR elements containing several deletions, insertions, and tandem duplication sequences, we used the BioEdit program and performed comparative analysis between those LTR sequences (Fig. 2). All structural differences among the several groups were detected in the U3 region of the LTR elements. A group (A1, A2, A3) showed 98% sequence similarity with that of PERV-B (43) LTR, which could generate infectious particles after transfection into human 293 cells (Scheef et al., 2001). The B, C, and D groups contained 24 bp inserted sequences of "TGA AGA GCC CTA ACT CCA GCT TCC" in the 5′ region of the LTR elements. To

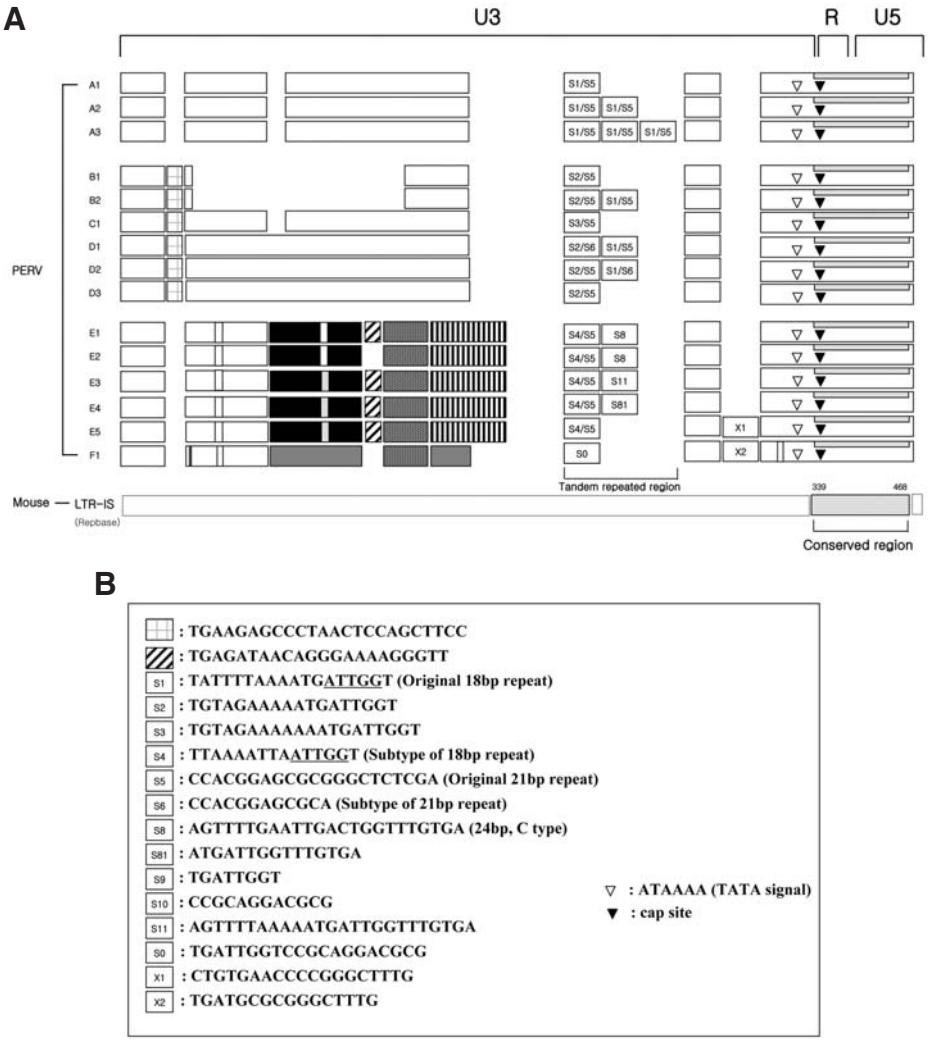

Fig. 2. Schematic representation of identified gamma1 PERV LTR elements (A) and consensus sequences of figured boxes (B). Different subtypes of PERV LTR elements were identified by specific consensus sequences and different number of tandem repeats. In the tandem repeated region, different box numbers indicates the different numbers of specific tandem repeated sequences. In the case of the mouse LTR-IS element, only the conserved region (open gray box) has high homology with the PERV LTR elements. Open and closed boxes indicate the conserved regions, and figured boxes indicate the specific sequences. In the consensus sequences, the underlined nucleotide sequences indicate the binding site for the transcription factor, NF-Y.

date, those specific 24 bp sequences have not yet been identified in the GenBank database. In detail, these three groups show different structures due to insertion and deletion (INDEL) events. The LTR elements belonging to the B group show a major deletion of 235 bp sequences derived by the homologous recombination of "AAA AGC CCT AG" compared to that of the D group. The LTR elements belonging to the C group had common deletion sequences of 50 bp as shown in group A. However, these 50 bp sequences were already identified on PERV-A7 locus (Huh et al., 2007b). We thus concluded that 50 bp differences were derived by transcription of different genomic template. Internal sequences of groups E and F were totally different from the A, B, C, and D group. The LTR elements belonging to the E group had 81% sequence similarity with replication competent PK15-PERV-A (58) and Bac-PERV-A (151B10) LTRs. The LTR elements between the E group (E1, E2, E3, E4, E5) and the F group (F1) shared common sequences in the 5′ region of the LTRs. However, the F group sequences could not be assigned as previously investigated LTRs by the BLAST search program. All LTR groups had highly conserved TATA signals and cap sites. All of those six groups of LTR elements were sub-divided into 15 subgroups according to their short INDEL sequences and the number of tandem repeated sequences (Fig. 2).

Evolutionary analysis of transcribed PERV LTR elements

Fundamentally, replication competent PERVs are initiated by the R-U5 sequences at its 5′end. However the greater part of transcribed PERV LTRs elements harbor the U3-R-U5 structure when we analyzed PERV elements by Blast searching method in the database of GenBank non-mouse and non-human EST entries (data not shown). Those phenomena could be explained by two alternative reasons. One is the inaccurate prediction of PERV LTR structure. The other is that LTR sequences might be transcribed by the upstream promoter elements, not by their own transcriptional activity. However, that sequences could show us the more diverse LTR elements for the PERV LTR investigation, we selected the full length PERV LTRs for our investigation.

From the massive cloning and sequencing procedure, we could identify the 69 different actively transcribed LTR clones from three different Korean pigs including six groups and 15 sub-groups (Fig. 2). Those transcribed LTR elements have not yet been identified in any previous report. Structure analysis of the LTR elements indicates that the 69 LTR sequences seem to have originated from one type of exogenous retrovirus infection, and then followed a series of evolutionary events (Figs. 2 and 3). The infection event of the exogenous retrovirus might have occurred on the common ancestor genome of the current

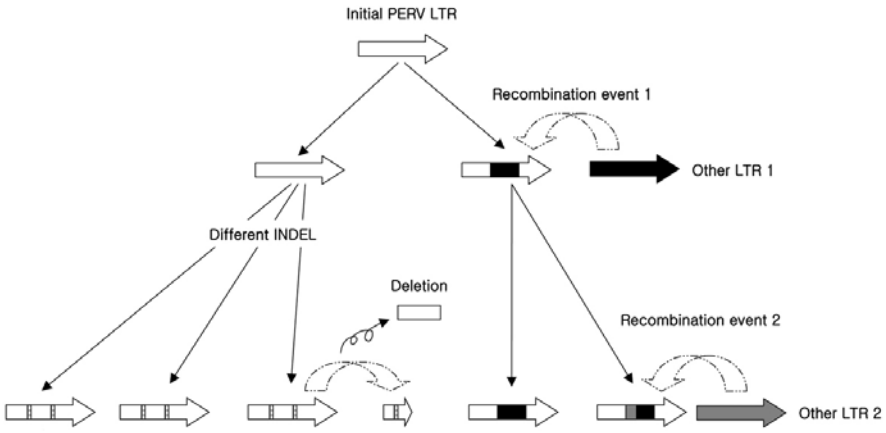

Fig. 3. Schematic representation of the occurrence of six groups of LTR elements. The arrow indicates the hypothetical evolutionary pathway of the LTR transcripts. Figured boxes indicate the repeated sequences of "AAA AGC CCT AG".

pig strains approximately 7.6×10^6 years ago (Nibert et al., 2005; Tonjes et al., 2003). Then, most of the PERV LTR sequences were selected as endogenous retroviruses (Herring et al., 2001). However, few PERVs have sustained the capacity to replicate their progeny and produce the infectious virus particles in human and porcine cell lines (Czauderna et al., 2000; Preuss et al., 2006; Woods et al., 1973). Thus, they could be major concerns for xenotranplantation.

From the sequence analysis of the six current groups of LTR elements, we assume that several evolutionary events occurred in the U3 region during pig evolution (Fig. 2). The group B LTR might have occurred through a homologous recombination excision event of short repeated sequences of the group D LTR. All types of LTR sequences except for the B type LTR sequences were identified in the genomic sequences using the BLAST search tools. Thus, we might conclude that B type sequences have emerging through the mechanism of transcription. However, current pig genome sequences did not cover the entire genome of pig. For more confidence, experimental validation might be needed. Intriguingly, group E and F show the totally different internal sequences (Fig. 2A). From the conserved marginal sequences and heterogeneous integral sequences, we could propose two separate recombination event (Fig. 4). Another individual INDEL event might also have occurred in the primitive PERV LTR sequences for the current group A, C, and D sequences. Specific INDEL events of the LTR element have occasionally been discovered by the LTR study of other ERV LTR elements (Huh et al., 2003; Ling et al., 2002). We could not trace the exact evolutionary pathway of the LTR elements, but current structure of the LTR elements may be good clues for the investigation of its evolutionary pathway. We thus assumed that two separated recombination events might have occurred in PERV LTR elements with other retrovirus LTRs (Fig. 3).

As a result, we identified structurally different six groups (15 subgroups) of LTR elements and investigated the evolutionary relationship of them by comparing the conserved and INDEL regions. Due to the experimental limitation, all kinds of transcribed LTR elements could not be investigated. However, our genome-wide analysis of transcribed LTR elements on pig genome could inform the basic information about transcribed PERV-LTR elements for the understanding of PERVs related with xenotransplantation.

Note: Supplementary information is available on the Molecules and Cells website (www.molcells.org).

ACKNOWLEDGEMENT

This work was supported by a grant (Code 20070101034012) from BioGreen 21 Program, Rural Development Administration, Korea.

REFERENCES

Akiyoshi, D.E., Denaro, M., Zhu, H., Greenstein, J.L., Banerjee, P.T., and Fishman, J.A. (1998). Identification of a full-length cDNA for an endogenous retrovirus of miniature swine. J. Virol. *72*, 4503-4507.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, W., Miller, W., and Lipman, D.J. (1997). Gapped Blast and PSI-BLAST: a new generation of protein database search program. Nucleic Acids Res. *25*, 3389-3402.

Czauderna, F., Fischer, N., Boller, K., Kurth, R., and Tonjes, R.R. (2000). Establishment and characterization of molecular clones of porcine endogenous retroviruses replicating on human cells. J. Virol. *74*, 4028-4038.

Denner, J., Specke, V., Thiesen, U., Karlas, A., and Kurth, R. (2003). Genetic alteration of the long terminal repeat of an ecotropic porcine endogenous retrovirus during passage in human cells. Virology *314*, 125-133.

Dieckhoff, B., Karlas, A., Hofmann, A., Kues, W.A., Petersen, B., Pfeifer, A., Niemann, H., Kurth, R., and Denner, J. (2007). Inhibition of porcine endogenous retroviruses (PERVs). in primary porcine cells by RNA interference using lentiviral vectors. Arch. Virol. *152*, 629-634.

Harrison, I., Takeuchi, Y., Bartosch, B., and Stoye, J.P. (2004). Determinants of high titer in recombinant porcine endogenous retroviruses. J. Virol. *78*, 13871-13879.

Herring, C., Quinn, G., Bower, R., Parsons, N., Logan, N.A., Brawley, A., Elsome, K., Whittam, A., Fernandez-Suarez, X.M., Cunningham, D., et al. (2001). Mapping full-length porcine endogenous retroviruses in a large white pig. J. Virol. *75*, 12252-12265.

Huh, J.W., Hong, K.W., Yi, J.M., Kim, T.H., Takenaka, O., Lee, W.H., and Kim, H.S. (2003). Molecular phylogeny and evolution of the human endogenous retrovirus HERV-W LTR family in hominoid primates. Mol. Cells *15*, 122-126.

Huh, J.W., Ha, H.S., Kim, Y.J., Lee, J.R., Kim, D.S., Cho, B.W., and Kim, H.S. (2007a). Methylation and promoter activity of long terminal repeats from porcine endogenous retroviruses in the Pig, *Sus scrofa*. Korean J. Genet. *29*, 171-176.

Huh, J.W., Cho, B.W., Kim, D.S., Ha, H.S., Yi, J.M., Kim, Y.J., Lee, J.R., Ahn, K., Lee, W.H., and Kim, H.S. (2007b). Long terminal repeats of porcine endogenous retroviruses in *Sus scrofa*. Arch. Virol. *52*, 2271-2276.

Karlas, A., Kurth, R., and Denner, J. (2004). Inhibition of porcine endogenous retroviruses by RNA interference: increasing the safety of xenotransplantation. Virology *325*, 18-23.

Klymiuk, N., Müller, M., Brem, G., and Aigner, B. (2006). Phylogeny, recombination and expression of porcine endogenous retrovirus gamma2 nucleotide sequences. J. Gen. Virol. *87*, 977-986.

Krach, U., Fischer, N., Czauderna, F., and Tonjes, R.R. (2001). Comparison of replication-competent molecular clones of porcine endogenous retrovirus class A and class B derived from pig and human cells. J. Virol. *75*, 5465-5472.

Kumar, S., Tamura, K., Jakobsen, I.B., and Nei, M. (2001). MEGA2: molecular evolutionary genetics analysis software. Bioinformatics *17*, 1244-1245.

Le Tissier, P., Stoye, J.P., Takeuchi, Y., Patience, C., and Weiss, R.A. (1997). Two sets of human-tropic pig retrovirus. Nature *389*, 681-682.

Lee, J.H., Webb, G.C., Allen, R.D., and Moran, C. (2002). Characterizing and mapping porcine endogenous retroviruses in Westran pigs. J. Virol. *76*, 5548-5556.

Ling, J., Pi, W., Bollag, R., Zeng, S., Keskintepe, M., Saliman, H., Krantz, S., Whitney, B., and Tuan, D. (2002). The solitary long terminal repeats of ERV-9 endogenous retrovirus are conserved during primate evolution and possess enhancer activities in embryonic and hematopoietic cells. J. Virol. *76*, 2410-2423.

Mang, R., Maas, J., Chen, X., Goudsmit, J., and van der Kuyl, A.C. (2001). Identification of a novel type C porcine endogenous retrovirus: evidence that copy number of endogenous retroviruses increases during host inbreeding. J. Gen. Virol. *82*, 1829-1834.

Martin, U., Kiessl, V., Blusch, H., Haverich, A., von der Helm, K., Herden, T., and Steinhof, G. (1998). Expression of pig endogenous retrovirus by primary porcine endotherial cells and infection of human cells. Lancet *352*, 692-694.

Moalic, Y., Blanchard, Y., Felix, H., and Jestin, A. (2006). Porcine endogenous retrovirus integration sites in the human genome: features in common with those of murine leukemia virus. J. Virol. *80*, 10980-10988.

Niebert, M., and Tonjes, R.R. (2005). Evolutionary spread and recombination of porcine endogenous retroviruses in the suiformes. J. Virol. *79*, 649-654.

Niebert, M., Rogel-Gaillard, C., Chardon, P., and Tonjes, R.R. (2002). Characterization of chromosomally assigned replication-competent gamma porcine endogenous retroviruses derived from a large white pig and expression in human cells. J. Virol. *76*, 2714-2720.

Patience, C., Takeuchi, Y., and Weiss, R.A. (1997). Infection of human cells by an endogenous retrovirus of pigs. Nat. Med. *3*, 282-286.

Patience, C., Switzer, W.M., Takeuchi, Y., Griffiths, D.J., Goward, M.E., Heneine, W., Stoye, J.P., and Weiss, R.A. (2001). Multiple groups of novel retroviral genomes in pigs and related species. J. Virol. *75*, 2771-2775.

Preuss, T., Fischer, N., Boller, K., and Tonjes, R.R. (2006). Isolation and characterization of an infectious replication-competent molecular clone of ecotropic porcine endogenous retrovirus class C. J. Virol. *80*, 10258-10261.

Quinn, G., and Langford, G. (2001). The porcine endogenous retrovirus long terminal repeat contains a single nucleotide polymorphism that confers distinct differences in estrogen receptor binding affinity between PERV A and PERV B/C subtypes. Virology *286*, 83-90.

Scheef, G., Fischer, N., Krach, U., and Tonjes, R.R. (2001). The number of a U3 repeat box acting as an enhancer in long terminal repeats of polytropic replication-competent porcine endogenous retroviruses dynamically fluctuates during serial virus passages in human cells. J. Virol. *75*, 6933-6940.

Scheef, G., Fischer, N., Flory, E., Schmitt, I., and Tonjes, R.R. (2002). Transcriptional regulation of porcine endogenous retroviruses released from porcine and infected human cells by heterotrimeric protein complex NF-Y and impact of immunosuppressive drugs. J. Virol. *76*, 12553-12563.

Sellars, M.J., Vuocolo, T., Leeton, L.A., Coman, G.J., Degnan, B.M., and Preston, N.P. (2007). Real-time RT-PCR quantification of Kuruma shrimp transcripts: a comparison of relative and absolute quantification procedures. J. Biotechnol. *129*, 391-399.

Takefman, D.M., Spear, G.T., Saifuddin, M., and Wilson, C.A. (2002). Human CD59 incorporation into porcine endogenous retrovirus particles: implications for the use of transgenic pigs for xenotransplantation. J. Virol. *76*, 1999-2002.

Takeuchi, Y., Patience, C., Magre, S., Weiss, R.A., Banerjee, P.T., Le Tissier, P., and Stoye, J.P. (1998). Host range and interference studies of three classes of pig endogenous retrovirus. J. Virol. *72*, 9986-9991.

Tönjes, R.R., and Niebert, M. (2003). Relative age of proviral porcine endogenous retrovirus sequences in Sus scrofa based on the molecular clock hypothesis. J. Virol. *77*, 12363-12368.

Wilson, C.A., Wong, S.W., Muller, J., Davidson, C.E., Rose, T.M., and Burd, P. (1998). Type C retrovirus released from porcine primary peripheral blood mononuclear cells infects human cells. J. Virol. *72*, 3082-3087.

Wilson, C.A., Laeeq, S., Ritzhaupt, A., Colon-Moran, W., and Yoshimura, F.K. (2003). Sequence analysis of porcine endogenous retrovirus long terminal repeats and identification of transcriptional regulatory regions. J. Virol. *77*, 142-149.

Woods, W.A., Papas, T.S., Hirumi, H., and Chirigos, M.A. (1973). Antigenic and biochemical characterization of the C-type particle of the stable porcine kidney cell line PK-15. J. Virol. *12*, 1184-1186.