

Mapping Full-Length Porcine Endogenous Retroviruses in a Large White Pig

C. HERRING,¹ G. QUINN,¹ R. BOWER,¹ N. PARSONS,¹ N. A. LOGAN,² A. BRAWLEY,² K. ELSOME,¹
A. WHITTAM,¹ X. M. FERNANDEZ-SUAREZ,¹ D. CUNNINGHAM,¹
D. ONIONS,² G. LANGFORD,¹ AND L. SCOBIE^{2*}

*Imutran Ltd. (a Novartis Pharma AG Company), Cambridge CB2 2YP,¹ and Department of Veterinary Pathology,
University of Glasgow, Glasgow G61 1QH,² United Kingdom*

Received 23 July 2001/Accepted 26 September 2001

Xenotransplantation may bridge the widening gap between the shortage of donor organs and the increasing number of patients waiting for transplantation. However, a major safety issue is the potential cross-species transmission of porcine endogenous retroviruses (PERV). This problem could be resolved if it is possible to produce pigs that do not contain replication-competent copies of this virus. In order to determine the feasibility of this, we have determined the number of potentially replication-competent full-length PERV proviruses and obtained data on their integration sites within the porcine genome. We have screened genomic DNA libraries from a Large White pig for potentially intact proviruses. We identified six unique PERV B proviruses that were apparently intact in all three genes, while the majority of isolated proviruses were defective in one or more genes. No intact PERV A proviruses were found in this pig, despite the identification of multiple defective A proviruses. Genotyping of 30 unrelated pigs for these unique proviruses showed a heterogeneous distribution. Two proviruses were uncommon, present in 7 of 30 and 3 of 30 pigs, while three were each present in 24 of 30 pigs, and one was present in 30 of 30 animals examined. Our data indicate that few PERV proviruses in Large White pigs are capable of productive infection and suggest that many could be removed by selective breeding. Further studies are required to determine if all potentially functional proviruses could be removed by breeding or whether gene knockout techniques will be required to remove the residuum.

Xenotransplantation offers the chance to bridge the growing gap between the numbers of patients requiring solid organ transplantation and the number of organs available. In addition, a number of clinical trials have been proposed or are ongoing using xenogeneic material for indications such as diabetes and hepatic support during acute liver failure (12, 34). While the potential benefits are considerable, there has been much public and scientific debate over the possible risk of zoonosis in xenotransplantation. The prolonged presence of an animal organ, together with systemic microchimerism in the recipient, presents an opportunity for the transmission of zoonotic agents (22, 28, 43, 49). In addition, the use of antirejection strategies, such as immunosuppression or complement inhibition, may lower the body's natural defenses against such pathogens (44). While it is possible to derive specific-or qualified pathogen-free pigs to exclude known exogenous pathogens with the potential for zoonotic spread from donor pig herds, there remain risks both from unknown agents and agents that are transmitted in the germ line. Concerns have been focused on porcine endogenous retrovirus (PERV), a gammaretrovirus with replication-competent human-tropic variants (33, 50), which is found in the germ line of all pig breeds examined to date (17). Other classes of PERV have been identified, including betaretroviruses (13, 32), and while more may remain to be found, to date, all of these appear to be defective.

Domestic pigs have up to 50 copies of PERV as measured using a *pro* probe on Southern blots (22). Three subgroups of PERV have been identified to date, PERV A, B, and C (1, 33). Although these subgroups were defined initially by molecular analysis, infection and pseudotyping experiments have shown that PERV A, B, and C use different, uncharacterized, cell receptors, with receptor usage being the method of formally classifying subgroups (45, 51). The *gag* and *pol* genes of all these subgroups are highly homologous, and they differ principally in the long terminal repeats (LTRs) and *env* hypervariable regions, the latter encoding putative cell receptor binding domains. PERV A and B are ubiquitous in domestic pigs; however, Large White pigs that do not have PERV C have been identified (6, 17). PERV A and B productively infect a wide range of cell types from both humans and other mammals, while PERV C has a more restricted tropism among the human cells examined, only infecting the HT1080 cell line nonproductively (45). PERVs released from pig tissues, such as peripheral blood mononuclear cells (PBMCs), endothelial cells, and pancreatic islets, have been shown to infect human cell lines, including 293 and endothelial cells (26, 45, 48, 50, 51).

Studies of humans treated with living pig tissue have not found any evidence of PERV infection (8, 11, 12, 14, 15, 23, 30, 31, 34, 38). These studies have included patients with a range of exposure routes to pig tissues, including skin grafts, transplantation with porcine pancreatic islets, and extracorporeal liver and splenic perfusion. Examination of multiple tissues in nonhuman primates transplanted with pig endothelial cells has also shown absence of *in vivo* infection with PERV (24, 25), although these primate cells are permissive for infection in

* Corresponding author. Mailing address: Department of Veterinary Pathology, University of Glasgow, Switchback Rd., Garscube Estate, Glasgow G61 1QH, United Kingdom. Phone: 44 141 330 2283. Fax: 44 141 330 5602. E-mail: l.scobie@vet.gla.ac.uk.

vitro (5). However, two recent studies using SCID mouse models have shown evidence of PERV infection in multiple tissue compartments (10, 48). van der Laan et al. (48) showed infection of mouse cells in a SCID mouse transplanted with porcine pancreatic islets. The reason for the discrepancy between the results in humans and SCID mice is not clear, although it may reflect the highly compromised immune status of the SCID mouse or the fact that only PBMCs have been used to evaluate PERV status in human patients.

Despite the lack of evidence of PERV infection in humans, it would be desirable to use organ-source pigs that present the lowest possible zoonotic risk. Bosch et al. (6) found that most expressed PERV proviruses were defective, in common with endogenous retroviruses in other species, such as humans, mice, and cats (46). Although some pig-to-pig variation in proviral distribution has been found (6, 17), it has been speculated that some proviruses are likely to be common to all pigs. The functional status of these common proviruses will be the significant factor in determining safety. Provirus expression patterns vary between different cell types (1, 6) and may be altered due to cryptic stimuli during organ rejection. Cloning of pigs by nuclear transfer is being actively explored as a method to generate genetically modified pigs for xenotransplantation (4, 29, 35). It has been shown in other species that gene methylation status can alter during the nuclear reprogramming steps of this procedure (18–20), hence proviruses previously silenced by methylation may be reexpressed in this situation.

A systematic study is required to determine which proviruses are intact and what degree of pig-to-pig variance they display. We have taken a genomic approach to this problem and screened libraries for proviruses that have both *gag* and *env* sequences. These proviruses have then been subjected to further analysis to confirm that they have the potential to result in productive infection. Since Large White pigs negative for PERV C have been identified and our own studies have indicated a lack of full-length C proviruses in some Large White pigs analyzed to date (6) (D. Hart, personal communication), we have concentrated our analysis on PERV A and B. The study reported here indicates that although pigs carry approximately 50 proviruses, the number of potentially functional proviruses in the Large White pig is low and there is sufficient variation within one herd to remove most proviruses by breeding.

MATERIALS AND METHODS

Library preparation and screening. DNA was extracted from porcine PBMCs and partially digested with *Sau3AI* according to the vector manufacturer's protocols. Libraries were constructed in *Bam*HI-digested lambda Fix II or pSuper-Cos 1 vectors according to the manufacturer's instructions (Stratagene). The cosmid library was screened for PERV A, while the lambda library was screened for PERV B as described below.

For the cosmid library, three to fourfold genome equivalents were plated on LB agar plates supplemented with 50 µg of ampicillin/ml. Colonies were grown overnight at 30°C and lifted onto a Hybond N⁺ filter (Amersham Pharmacia Biotech) to create the master filter. The master filter was then replicated onto three other filters (sets A, B, and C) which were processed as described by in Sambrook et al. (37) and UV cross-linked on a Stratalinker cross-linker (120,000 µJ; Stratagene).

The triplicate replica filters (sets A, B, and C) were screened with probes labelled with [α -³²P]dCTP, using the Megaprime random-primed labelling kit (Amersham Pharmacia Biotech) as follows. Filter set A was screened with a

probe against PERV *gag* (sequence AF038600; bases 585 to 2159), while sets B and C were screened with a probe against PERV A *env* (sequence Y12238; bases 745 to 1101). Unincorporated nucleotide was removed by spin column chromatography on Sephadex G50 columns. The filters were prehybridized for 3 h at 65°C in hybridization buffer containing 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 20 mM NaH₂PO₄, 5× Denhardt's solution, 0.4% sodium dodecyl sulfate (SDS), and 500 µg of denatured salmon sperm DNA/ml. After addition of the probe, the filters were hybridized overnight. Filters were washed to a final stringency of 0.1× SSC–0.1% SDS at 65°C and either exposed for 24 to 48 h to X-ray film at –70°C with intensifying screens or for 1 to 5 h to a storage phosphor screen, followed by phosphorimaging on a Molecular Dynamics Storm phosphorimager. Phosphorimaged blots were printed onto acetate filters prior to alignment on a light box with other relevant filters.

For the lambda Fix II library, a total of 1.5 × 10⁶ PFU was plated on NZY agar plates according to the manufacturer's instructions (Stratagene). Duplicate filters were lifted using Hybond NX (Amersham Pharmacia Biotech) and denatured and neutralized as described previously (37). Filters were UV cross-linked as described above and prehybridized for 30 min in Quickhyb solution (Stratagene). Probes were labelled with [α -³²P]dCTP using Prime-it II (Stratagene). The *gag* probe (AF038600; bases 1334 to 1813) was added to one filter, and the *env* probe (Y12239; bases 1387 to 1735) was added to the duplicate filter. The lambda library filters were hybridized and washed as described above for the cosmid library filters. The filters were exposed to X-ray film at –70°C for up to a week and developed using a Compact X4 automatic film processor (X-ograph Imaging Systems).

Clones that replicated on both the *gag* and *env* filters were then purified to homogeneity through multiple rounds of rescreeing. Cosmid DNA was prepared using the Qiagen QIAwell system (Qiagen). Lambda DNA was prepared from liquid lysates as described by Ausubel et al. (2).

Primary analysis. Cosmid and lambda clones were analyzed differently, as described below. This was a reflection of the lower yields of lambda DNA relative to cosmid DNA. Primary analysis of cosmid clones was performed by restriction fragment analysis, using *Sma*I and *Pvu*II digests, to identify clones with identical restriction fragment fingerprints. Representative clones of each restriction fragment length polymorphism (RFLP) class were further analyzed by PCR of open reading frames (ORFs) as described below. Lambda clones were analyzed by restriction digestion with *Nor*I, *Eco*RI, *M*siI, *Sma*I/*Not*I, or *Avr*II/*M*siI, followed by Southern blotting (37) onto Hybond NX (Amersham Pharmacia Biotech) and hybridization with either the ³²P-labelled *gag* or PERV B *env* probes described above.

Secondary analysis. The presence of intact viral genes in lambda and cosmid isolates was examined by PCR across the *gag*, *pol*, and *env* ORF or by PCR of the ORF followed by protein truncation testing (PTT) to analyze the coding potential of the *gag* and *env* ORFs. Primers were designed to conserved sequences at the start and end of each ORF, based on published sequence data. PCR conditions were optimized on laboratory-generated plasmid clones of PERV C *gag* and PERV A, B, and C *env* which had been fully sequenced and which encoded full-length ORFs with >98% homology to the published prototype sequences. The *pol* PCR was optimized on DNA from PK15 cells.

For *gag* ORF PCR, the following primers were used: *gag* F (5'-GGATCCTAATACGACTACTATAGGAACAGACCACCATGGGACAGACAGTGACT ACC-3') and *gag* R (5'-CCCTCCACCTTCAAAGTTAC-3') (GenBank accession no. AF038600). Fifty-microliter reaction mixtures contained 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 150 nM each primer, 200 nM each deoxynucleoside triphosphate (dNTP), and 2.5 U of Ampliqaq (PE Biosystems) with approximately 1 ng of cosmid or lambda template. Reaction mixtures were then cycled as follows: 95°C for 3 min followed by 30 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min 50 s, followed by 10 min at 72°C. PCR was performed in an ABI2400 or ABI9600 thermocycler (PE Biosystems).

For *pol* ORF PCR, the primers used were: *pol* F (5'-GGATCCTAATACGACTACTATAGGAACAGACCACCATGGGTGCCACAGGGCAAC-3') and *pol* R (5'-GACCATTGTCTGACCCGATTA-3') (accession no. AF038600). Fifty-microliter reaction mixtures contained 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 150 nM each primer, 200 nM each dNTP, and 2.5 U of Ampliqaq (PE Biosystems) with approximately 1 ng of cosmid or lambda template. Reaction mixtures were then cycled as follows: 95°C for 3 min, followed by 30 cycles of 95°C for 40 s, 58°C for 30 s, 72°C for 3 min, and a final soak at 72°C for 10 min.

For *env* ORF PCR, the primers used were: *env* F (5'-GGATCCTAATACGACTACTATAGGAACAGACCACCATGCATCCACGTTAAGCCG-3') and *env* R (5'-CGCTCTAGACTAAGCGTAGTCTGGGACGTCGTATGGGTAG AACTGGGAAGGGTAGAGGTCAGT-3') (GenBank accession no. Y12238). Fifty-microliter reaction mixtures were prepared under the same conditions as described for *gag* (above) and then cycled as follows: 95°C for 3 min and then 30

cycles of 94°C for 1 min, 62°C for 1 min, 72°C for 2 min 10 s, followed by 10 min at 72°C.

The forward primer sequences shown for *gag* and *env* were modified so that they contained a T7 RNA polymerase promoter sequence, such that PCR products from clones giving approximately the predicted product sizes for *gag* and *env* could be further analyzed by PTT. Additionally, the stop codon in the *env* reverse primer was replaced with an in-frame hemagglutinin (HA) epitope tag, enabling detection of full-length products by Western blotting against the HA tag.

PTT (36) for ORFs was carried out using coupled transcription-translation in rabbit reticulocyte lysate (TNT T7 Rapid for PCR; Promega), with nonradioactive detection using biotinylated lysine tRNA. Each 25- μ l reaction mixture contained 2.5 μ l of unpurified PCR product, 1 μ l of Transcend tRNA (Promega), and 20 μ l of TNT T7 PCR Quick master mix (Promega) supplemented with 20 μ M methionine. The reaction mixtures were incubated at 30°C for 90 min, and 2 μ l was subjected to SDS-polyacrylamide gel electrophoresis on a 4 to 12% NuPAGE Bis-Tris gel (Novex/Invitrogen, Paisley, United Kingdom). Gels were semi-dry blotted to a Hybond enhanced chemiluminescence (ECL) membrane (Amersham Pharmacia Biotech). Biotinylated translation products were detected by blotting them with streptavidin-horseradish peroxidase (HRP) conjugate (1/3,750 dilution; Roche), followed by ECL detection (Amersham Pharmacia Biotech). Product sizes were assessed by reference to a translated *gag* or *env* plasmid clone of known sequence and a biotinylated protein molecular weight marker (Roche). HA epitope-tagged products were detected by blotting with high-affinity rat anti-HA antiserum (1/500; Roche) and a secondary sheep anti-rat HRP conjugate (1/10,000; Sigma), followed by ECL detection.

Sequence analysis of cosmid and lambda clones. Clones identified as having full-length *gag*, *pol*, and *env* genes together with full-length *gag* and *env* translation products were then sequenced by a combination of methods. Cosmid DNA and PCR products were directly sequenced using ABI PRISM Dye Terminator Cycle Sequencing on an ABI373 sequencer (Applied Biosystems), while some PCR products were subcloned into pCR2.1 (Invitrogen) prior to sequencing. The lambda clones were directly sequenced on a LICOR 4000 DNA sequencer (MWG Biotech) as follows: 800 ng of lambda DNA was prepared using a Thermo Sequenase fluorescence-labelled sequencing kit (Amersham), according to the manufacturer's instructions.

Analysis of PCR-negative cosmid clones. Most clones that gave no PCR products for *gag*, *pol*, or *env* ORFs were subjected to Southern blotting to confirm that they were defective and/or contained deleted material. Southern blots of *Sma*I-digested cosmid DNA were probed with biotin-labelled oligonucleotides homologous to the region around the primer binding site (PBS) 5'-biotin, GCC TTTCATTGGTGGCGTTGGCCGGGAAATCCTCGCGACCACCCCTTACA C-3' (nucleotide [nt] 661 to 711; accession no. AJ133817), and the 3'-end of *env* 5'-biotin, ACTGACCTCTAGCCTTCCAGTCTAAGATTAGAACTATTA ACAAGACAA-3' (nt 8121 to 8171; accession no. AJ133817). These probes were external to the probe sequences used in the initial library screening and to the *gag* and *env* PCR primer sets. Southern-blotted DNA was prehybridized at 40°C for 6 h in hybridization buffer (as described above) and then hybridized overnight at 40°C in hybridization buffer containing 27 nM biotin-labelled probe oligonucleotide. Blots were then washed in 1 \times SSC-0.1% SDS three times for 30 min at 42°C and then developed with streptavidin-HRP conjugate and ECL detection as described for the PTT above. It was assumed that, because of the sequence conservation seen about these regions in all known PERV variants, failure to bind these oligonucleotides indicated a provirus containing a deletion (deleted provirus).

Proviral flanking sequence analysis. Proviral flanking sequence was obtained for all clones that were shown to contain apparently intact proviruses. Flanking sequence was generated by linker-mediated PCR between a primer in the PERV LTR and a primer in a linker ligated onto blunt-end-cut DNA (40). One-half microgram of *Pvu*II-digested cosmid or lambda DNA was ligated to 60 pmol of double-stranded adaptor by use of 10 U of T4 DNA ligase (Invitrogen) in a 7- μ l reaction mixture for 16 to 48 h at 15°C. The adaptor consisted of annealed upper-strand oligonucleotide (5'-GTAATACGACTCACTATAGGGCAGCGG TGGTCGACGCCCCGGGCTGGT-3') and a 3'-amine-blocked lower-strand oligonucleotide (5'-ACCAGCCCC-NH₂-3'). One microliter of a 1:100 dilution of this reaction was used as template in a 25- μ l PCR reaction mix containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 50 nM LTR-specific primer, 50 nM adaptor primer (5'-GTAATACGACTCACTATAGGGC-3'), 200 nM each dNTP, and 2.5 U of Ampliqaq Gold (PE Biosystems). Reactions were cycled as follows: 94°C for 9 min, followed by 30 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 4 min, and 72°C for 10 min. The LTR-specific primers used were in the U5 region, which is conserved among all known PERV variants; primer LTR-5' (5'-GTGAACCCATAAAAGCTGTC-3') was used to isolate 5'-flanking sequence, while primer LTR-3' (5'-GACAGCTTTTATGGGGTTCAC-3') was

used to isolate 3'-flanking sequence (GenBank accession no. AF038600 and AF038601).

Two PCR products were usually produced for each reaction, representing an LTR-internal proviral sequence and an LTR-flanking DNA sequence. Both products were either direct sequenced or gel purified with the Qiaquick Gel Extraction kit (Qiagen), cloned into pCR2.1-TOPO (Invitrogen), and then sequenced.

Analysis of provirus prevalence within a transgenic Large White pig herd. Animals from Imutran Ltd's herd of pigs transgenic for the human complement regulator decay-accelerating factor (hDAF) were used to determine the incidence of the different proviruses. DNA was isolated from pig tissue by either a standard proteinase K-phenol-chloroform method (37) or by using Qiagen Genomic tips (Qiagen) according to the manufacturer's protocol. Based on sequence generated from the flanking PCR above, primer pairs were designed that would amplify between the PERV LTR and the adjacent flanking DNA, yielding a PCR product that would uniquely identify each provirus. Positive control reactions contained the lambda or cosmid clone the provirus was originally identified from. β -Globin (GenBank accession no. X86791) PCR was performed on each sample to confirm that the DNA was suitable for amplification, using ~35 ng of template DNA in a 50- μ l reaction mixture containing 150 nM each primer β -globin F1 (5'-GCAGATTCCTCAAAACCTTCGAGAG-3') and β -globin R1 (5'-TCTGCCCAAGTCCTAAATGTGCGT-3'), 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 200 nM each dNTP, and 1.25 U of Ampliqaq Gold (PE Biosystems). Reactions were cycled as follows: 95°C for 9 min, 30 cycles of 95°C for 30 s, 58°C for 30 s and 72°C for 30 s, followed by 7 min at 72°C.

For the cosmid, 78 provirus, each reaction used ~60 ng of template DNA in a 50- μ l reaction mixture containing: 300 nM each primer 78ltrs (5'-GAACCC ATAAAAGCTGTCC-3') and 78ltrs (5'-GATCCTATGTTGGGTGCATTT-3'), 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 200 nM each dNTP, and 2 U of Ampliqaq Gold (PE Biosystems). Reactions were cycled as follows: 94°C for 9 min, followed by 35 cycles of 94°C for 30 s, 54°C for 30 s, and 72°C for 30 s, and a final 10 min at 72°C. The predicted product size was 323 bp based on sequence analysis.

Reaction conditions for all remaining primer sets described below were as follows: ~20 ng of template DNA in a 50- μ l reaction mixture containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 150 nM each primer, 200 nM each dNTP, and 1.25 U of Ampliqaq Gold (PE Biosystems).

Primer sets and cycle parameters used were as follows: Provirus 35121i, primers FLPr3 (5'-CATGACGGCAACTCTGAAG-3') and revltr1 (5'-GTGAACCCATAAAAGCTGTCC-3') under the following cycle parameters: 95°C for 9 min, and then 35 cycles of 94°C for 1 min, 58°C for 1 min, and 72°C for 2 min, followed by 10 min at 72°C. The predicted product size was 455 bp based on sequence analysis.

Provirus 21321, primers 21f2 (5'-GGTCTGTGACCTACTCCATA-3') and ltr1 (5'-GACAGCTTTTATGGGGTTCAC-3') under the following cycle parameters: 95°C for 9 min, and then 35 cycles of 94°C for 1 min, 58°C for 1 min, and 72°C for 2 min, followed by 10 min at 72°C. The predicted product size was 518 bp based on sequence analysis.

Provirus 73414i, primers 73-4r-3 (5'-GTCATGCCGTTTCATTTGGGA-3') and revltr1 under the following cycle parameters: 95°C for 9 min, and then 40 cycles of 94°C for 1 min, 56°C for 1 min, and 72°C for 2 min, followed by 10 min at 72°C. The predicted product size was 388 bp based on sequence analysis.

Provirus 73414ii, primers 73-3f-3 (5'-AGCAACAGAATTGAAGTCAG-3') and revltr1 under the same cycle parameters as described above for 73414i. The predicted product size was 438 bp based on sequence analysis.

Provirus 310518-5', the primers 31f8 (5'-GGATGGAGACTATGCTCAGC-3') and ltr1 under the following cycle parameters: 95°C for 9 min, and then 35 cycles of 94°C for 1 min, 56°C for 1 min, and 72°C for 2 min, followed by 10 min at 72°C. The predicted product size was 1,641 bp based on sequence analysis.

Provirus 310518-3', the primers 31f2 (5'-AACTCATGATAAATGAAACAC C-3') and revltr1 under the same cycle parameters as described above for 310518-5'. The predicted product size was 815 bp based on sequence analysis.

RESULTS

Isolation of proviruses containing both *gag* and *env* sequences. The screening strategy used was designed to rapidly exclude defective proviruses. To this end, only clones that gave signals for both *gag* and *env* probes were analyzed further. Proviruses that were missing either the entire *gag* gene or the

VRA and VRB portions of *env* were thus excluded from further analysis. Clones with more subtle deletions or rearrangements were eliminated at later stages as described below.

The cosmid library was screened for PERV A proviruses by use of a PERV A *env* probe, and the lambda library was screened for PERV B proviruses by use of a PERV B *env* probe. Both libraries were also screened with PERV *gag* probes as described above. Both libraries yielded substantially more *gag*-positive colonies than *env* positives, partly reflecting the non-PERV subtype selectivity of the *gag* probes and in the case of the cosmid library the larger size of the *gag* probe used. Eighty-four clones were identified from the cosmid library as *gag* and *env* A positive, of which 78 were recovered. Of the 78 clones which were recovered, 74 were purified to homogeneity through further rounds of screening, whereas sequence analysis showed that the other four clones did not contain an insert. From the lambda library, 40 plaques were initially isolated, of which four were found to be either *gag* or *env* B negative upon the second round of screening and were discarded. The remaining 36 positive plaques were purified to homogeneity. Multiple clones derived from each initial plaque were analyzed to ensure that any instability of the insert would not affect the results. The lambda library was also screened for the presence of PERV A; however, no PERV A *env* positives were associated with *gag* positives, suggesting an absence of full-length intact PERV A (data not shown). As a result, this was not pursued any further.

The doubly positive cosmid clones were then subjected to restriction fragment analysis to identify duplicate clones as shown in Table 1. Of the 74 clones analyzed, 54 were unique and 20 were duplicates of other clones. Where there were two or more cosmid clones with identical RFLP patterns, two or more representatives of each were analyzed to confirm results.

As the yield of DNA from the lambda preparations was much lower than that from the cosmid clones, lambda clones were subjected to Southern blot analysis using the *gag* and *env* probes described in Materials and Methods. The results of the analysis of the lambda clones are shown in Table 2. Southern blotting showed that the majority of PERV B lambda clones possessed a variant RFLP with respect to the published PERV B sequence, suggesting they were potentially defective or deleted.

Secondary analysis. Clones that gave both *gag* and *env* signals upon primary library screening were then analyzed by PCR or PCR followed by PTT, which identified clones with intact ORFs. The result of the PCR was regarded as positive if the clone gave a PCR product of approximately the size expected for the intact ORF based on published PERV sequences. If a clone was defective for either *gag*, *pol*, or *env*, analysis of the clone was terminated, because the provirus was deemed to be defective. In some cases, this means that results are not available for *gag*, *pol*, and *env* PCR. However, the results in Tables 1 and 2 demonstrate that sufficient data were obtained to conclude that such clones did contain defective proviruses.

***gag*.** The *gag* PCR showed major differences between the PERV A-screened cosmid library and the PERV B-screened lambda library. While 32 of 36 lambda clones gave full-length *gag* PCR products, only 13 of 54 unique cosmid clones yielded full-length products (Tables 1 and 2). Sequencing from the

cosmid multiple cloning site (MCS) showed that in a number of cases this was a result of genomic DNA being cloned into the vector within the proviral sequence (termed cloning deletions in Table 1). However, for the majority of clones that gave no *gag* PCR product, sequence adjacent to the MCS was not PERV sequence, indicating that no cloning deletion had occurred. No *gag* products that were smaller than expected, which could have represented proviruses with internal deletions, were detected from either library.

PTT of apparently full-length *gag* genes from cosmid clones showed that 8 of 13 clones had intact *gag* ORFs and 1 of 13 had a truncated ORF, while 4 of 13 clones were not tested as they had already been excluded from analysis on the basis of a truncated *env* gene (Table 1).

In contrast to the cosmid clone results, only 11 of 32 lambda clones had an intact *gag* ORF and 20 of 34 had a truncated *gag* ORF. One clone of 32 was not analyzed by PTT as it had previously been found to be defective in *gag* (Table 2 and Fig. 1A).

***pol*.** The *pol* PCR primers were designed to amplify an ~2.7-kb product, based on sequence alignments with the prototype PERV sequence. Thirty-seven of 54 unique cosmid clones gave *pol* PCR products, while 16 of 54 clones gave no *pol* PCR products; 1 of 54 clones was not analyzed as it had been truncated in cloning (Table 1). A smaller ~720-bp product was amplified from some cosmids and was the sole product from cosmid 5 (Fig. 1B). Sequencing of this product showed >96% homology over three regions to the PERV *pol* gene, with two internal deletions. There was homology from the *pol* ATG codon at base 2307 to base 2496 of PERV (accession no. AF038600), then a deletion, and homology from bases 3721 to 4051, then a deletion, and homology from base 4872 to the end of the PCR product at base 5070, giving a sequenced product with a size of 718 bp. This small product did not encode a functional Pol protein, and clones giving such PCR products were regarded as defective. Some clones also gave a larger than predicted 3.8-kb product. The 3.8-kb product could potentially encode a functional Pol protein, but this was not investigated further as work on *gag* and *env* showed clones giving this product were defective in at least one other gene. As shown in Fig. 1B, a few cosmid clones gave both the predicted full-length product and the 718-bp deleted product. In these cases, the larger fragment was assumed to be associated with the provirus under study due to the possibility that there were two or more proviruses or proviral fragments on the cosmid inserts.

Of the 18 of 36 lambda clones subjected to *pol* PCR, only one did not give the expected 2.7-kb product (Table 2). The remaining 18 lambda clones had previously been identified as being defective in Gag or Env and were not subjected to *pol* PCR.

***env*.** For the *env* PCR, 18 of 54 cosmid clones gave a PCR product of the predicted size of the ORF at ~2 kb, 35 of 54 gave no PCR product, and one was not tested (Table 1 and Fig. 1C). No smaller PCR products representing *env* internal deletions were identified.

PTT was performed on the *env* PCR products shown in Fig. 1C. Only 2 of 18 cosmid clones gave full-length translation products for Env, which was confirmed by carboxy-terminal epitope tagging PTT as described in Materials and Methods, and is illustrated for clone 78 (Fig. 1C). Sequence analysis of

TABLE 1. Analysis of *gag* and *env* double-positive clones obtained from a *PERV A*-screened cosmid library from the Large White pig

Clone no. ^a	RFLP pattern ^b	Presence of indicated probe by indicated test						Defect and probe
		<i>gag</i> PCR	<i>gag</i> PTT	<i>pol</i> PCR ^c	<i>env</i> PCR	<i>env</i> PTT	<i>env</i> HA PTT	
1	+1	CD ^g		+1	+	TP ^e	-	Cloning deletion
3	U	-		-	-			Yes
4	U	+	FL ^d	+	-			Yes, <i>env</i>
5	U	-		+s	-			Yes, <i>gag</i> , <i>pol</i> , and <i>env</i>
6	U	-		+	+	TP		Yes, <i>env</i>
7	+11	-		+	+	TP		Yes, <i>gag</i> and <i>env</i>
9	U	-		+	+	TP		Yes, <i>gag</i> and <i>env</i>
10	U	-		+1	-			Yes, <i>gag</i> and <i>env</i>
12	U	-		-	-			Yes
13	+6	-		+1	-			Yes, <i>gag</i> and <i>env</i>
15	U	-		+	+	TP		Yes, <i>gag</i> and <i>env</i>
18	U	+	ND ^f	+	+	TP		Yes, <i>env</i>
19	U	-		-	-			Yes
21	U	+	ND	-	-			Yes, <i>pol</i> and <i>env</i>
22	U	-		-	-			Yes
24	U	CD		+1	-			Cloning deletion
25	U	+		+	+	TP		Yes, <i>env</i>
27	U	-		+	-			Yes, <i>gag</i> and <i>env</i>
29	U	-		+1	-			Yes, <i>gag</i> and <i>env</i>
30	U	-		-	-			Yes, <i>gag</i> and <i>env</i>
31	U	-		+1	-			Yes, <i>gag</i> and <i>env</i>
34	U	-		+1	-			Yes, <i>gag</i> and <i>env</i>
35	U	-		+	+	TP		Yes, <i>gag</i> and <i>env</i>
36	U	CD		ND	ND			Cloning deletion
38	U	-		-	-			Yes, <i>gag</i> and <i>env</i>
39	U	-		-	-			Yes, <i>gag</i> and <i>env</i>
41	U	-		-	-			Yes, <i>gag</i> and <i>env</i>
43	U	-		-	-			Yes, gross deletion
45	U	-		+1	-			Yes, <i>gag</i> and <i>env</i>
46	U	+	TP	+	+	FL	+	Yes, <i>gag</i>
47	U	-		-	-			No insert
48	U	+	FL	+	+	TP	-	Yes, <i>env</i>
49	U	+	FL	+	+	TP	-	Yes, <i>env</i>
53	U	-		-	-			Yes, <i>gag</i> and <i>env</i>
54	+1	-		+	+	TP		Yes, <i>gag</i> and <i>env</i>
56	U	CD ^h		-	-			Cloning deletion
57	U	-		+	+	TP		Yes, <i>env</i>
58	U	+	FL	+	+	TP	-	Yes, <i>env</i>
59	U	+	ND	+1	-			Yes, <i>env</i>
60	U	+	FL	+	+	TP	-	Yes, <i>env</i>
61	U	+	FL	+	+	TP	-	Yes, <i>env</i>
62	U	-		-	-			Yes, <i>env</i>
64	U	CD		+1	-			Cloning deletion
66	U	-		+1	-			Yes, <i>gag</i> and <i>env</i>
67	U	CD		+1	-			Cloning deletion
69	U	-		+	-			Yes, <i>gag</i> and <i>env</i>
70	U	-		-	-			Yes, <i>gag</i> and <i>env</i>
74	U	-		+1	-			Yes
77	U	-		+1	-			Yes, <i>gag</i> and <i>env</i>
78	U	+	FL	+	+	FL	+	No intact
79	+1	-		-	-			Yes, <i>gag</i> and <i>env</i>
82	U	-		+	-			Yes, <i>gag</i> and <i>env</i>
83	U	+	FL	+	+	TP	-	Yes, <i>env</i>
84	U	-		+	-			Yes, <i>gag</i> and <i>env</i>

^a Clones were numbered sequentially on initial picking.

^b Unique (U) clones are indicated. Where duplicate clones were identified, the number of such duplicates is noted.

^c +, a product of the predicted size was obtained; + with an l or s, a respectively larger or smaller product was observed in the absence of the predicted product.

^d FL, full-length protein was translated.

^e TP, truncated protein was produced.

^f ND, not done.

^g CD, *gag* sequence was adjacent to vector sequence as a result of cloning deletions. Such clones are recorded as defective only if negative results were obtained for the *env* gene.

^h In the case of clone 56, the clone was truncated but negative PCR results were obtained for *pol* and *env*. This is recorded as a cloning deletion as it was not analyzed further.

TABLE 2. Analysis of *gag* and *env* double-positive clones obtained from a *PERV B*-screened lambda library from the Large White pig

Original plaque identification (ID)	Clone ID ^a	Southern blot data ^b	Presence of indicated probe by indicated test					Defect and probe
			<i>gag</i> PCR	<i>gag</i> PTT	<i>pol</i> PCR	<i>env</i> PCR	<i>env</i> PTT	
2.1	21321	PERV B Var	+	FL ^c	+	+	FL	No intact
2.2	22332	ND	+	TP ^d	ND ^e	+	FL	Yes, <i>gag</i>
2.4	2441	PERV B Var	+	FL	+	+	FL	No, intact
2.5	25321	PERV B Var	+	TP	+	+	TP	Yes, <i>gag</i> and <i>env</i>
2.6	2622	PERV B Var	+	TP	ND	+	ND	Yes, <i>gag</i>
2.7	27244	PERV B Var	+	TP	+	+	FL	Yes, <i>gag</i>
2.8	28321	PERV B Var	+	TP	ND	+	ND	Yes, <i>gag</i>
3.1	31211	PERV B Var	+	TP	+	+	FL	Yes, <i>gag</i>
3.2	3211	PERV B Var	+	TP	+	+	TP	Yes, <i>gag</i> and <i>env</i>
3.3	3321	PERV B	+	TP	ND	+	FL	Yes, <i>gag</i>
3.4	34321	PERV B Var	+	TP	ND	+	FL	Yes, <i>gag</i>
3.5	35121	PERV B	+	FL	+	+	FL	No intact
3.6	3611	PERV B Var	+	TP	ND	+	ND	Yes, <i>gag</i>
3.7	3711	PERV B	-		-	-		Yes
3.8	3821	PERV B	+	TP	ND	+	ND	Yes, <i>gag</i>
3.9	39111	PERV B	+	FL	ND	+	TP	Yes, <i>env</i>
3.10	310518	PERV B	+	FL	+	+	FL	No, intact
3.12	312146	PERV B Var	+	TP	ND	+	ND	Yes, <i>gag</i>
4.1	411	PERV B Var	+	TP	ND	+	ND	Yes, <i>gag</i>
4.3	436	PERV B Var	+	ND	+	-		Yes, <i>env</i>
4.4	44111	PERV B Var	+	FL	ND	+	TP	Yes, <i>env</i>
4.5	45411	PERV B	+	FL	+	+	TP	Yes, <i>env</i>
5.1	5112	PERV B Var	+	FL	+	+	TP	Yes, <i>env</i>
5.2	52111	PERV B	+	TP	ND	+	TP	Yes, <i>gag</i> and <i>env</i>
6.1	61121	PERV B Var	+	TP	ND	+	ND	Yes, <i>gag</i>
6.2	62321	PERV B	+	TP	+	+	FL	Yes, <i>gag</i>
6.4	6411	PERV B	+	TP	ND	+	ND	Yes, <i>gag</i>
6.5	65321	PERV B Var	+	FL	+	+	FL	No, intact
7.1	7111	PERV B Var	-		ND	-		Yes, <i>gag</i> and <i>env</i>
7.2	7211	PERV B Var	-		ND	-		Yes, <i>gag</i> and <i>env</i>
7.3	73414	PERV B Var	+	FL	+	+	FL	No, intact
7.4	74411	PERV B Var	+	FL	+	+	TP	Yes, <i>env</i>
7.5	75321	PERV B Var	-		ND	-		Yes, <i>gag</i> and <i>env</i>
8.1	81111	PERV B Var	+	TP	+	+	TP	Yes, <i>gag</i> and <i>env</i>
8.2	82222	PERV B Var	+	TP	+	+	TP	Yes, <i>gag</i> and <i>env</i>
8.3	8311	PERV B Var	+	TP	ND	+	TP	Yes, <i>gag</i> and <i>env</i>

^a From each original plaque picked, multiple plaques were analyzed, the clone ID is the number of the clone that was analyzed in the greatest detail and was the least defective.

^b PERV B, the restriction pattern obtained was identical to that of the published PERV B provirus Y17013; PERV B Var, a variant RFLP pattern was observed, indicating that the proviral sequence was either variant from the published data or had a deletion somewhere in the proviral sequence.

^c FL, full-length protein was translated.

^d TP, truncated protein was produced.

^e ND, not done.

cosmid clone 78, which also gave a full-length *gag* product, revealed that this was in fact a PERV B clones (see Fig. 3).

Subsequent sequencing of the 3' end of the *env* PCR product for clone 25 showed a CGA to TGA, Arg-to-stop mutation at amino acid 578, consistent with the ~64-kDa product seen in Fig. 1C. The product for clone 15 contained multiple base insertions throughout the sequence, causing multiple frame-shifts, and no significant ORF was present. The *env* genes for cosmid clones 6 and 50 were not sequenced.

Sixteen of 18 lambda PCR products gave truncated and therefore nonfunctional Env translation products. Sequence analysis of a number of cosmid isolates, including cosmid clones 15 and 25 (Fig. 1C), demonstrated that they were PERV A as expected.

In contrast to the results obtained from the cosmid library, 31 of 34 clones from the lambda library gave full-length PCR *env* products and only 3 of 34 yielded no product (Table 2). Thirteen of 31 PCR-positive clones gave full-length Env trans-

lation products on PTT, 10 of 31 gave truncated products, and 8 were not analyzed, as they had already been observed to be Gag defective.

To summarize, where the clones were amplifiable by one or more sets of primers but gave no product for one or more of the other ORFs, it was presumed that the remaining proviral sequences were defective and analysis was not continued further. By these criteria, no intact PERV A clones were identified from the cosmid library: of the 54 unique clones, only one gave full-length PTT products for Gag and Env, together with an appropriately sized *pol* PCR product. Sequence analysis of this clone, cosmid 78, showed that it was indeed PERV B. From the lambda library, six PERV B clones were identified as being potentially full length, clones 21321, 2441, 35121, 310518, 65321, and 73414. Analysis of these seven clones was continued further.

Analysis of the clones that did not give PCR products. A number of the *gag*- and *env*-positive cosmid clones isolated did

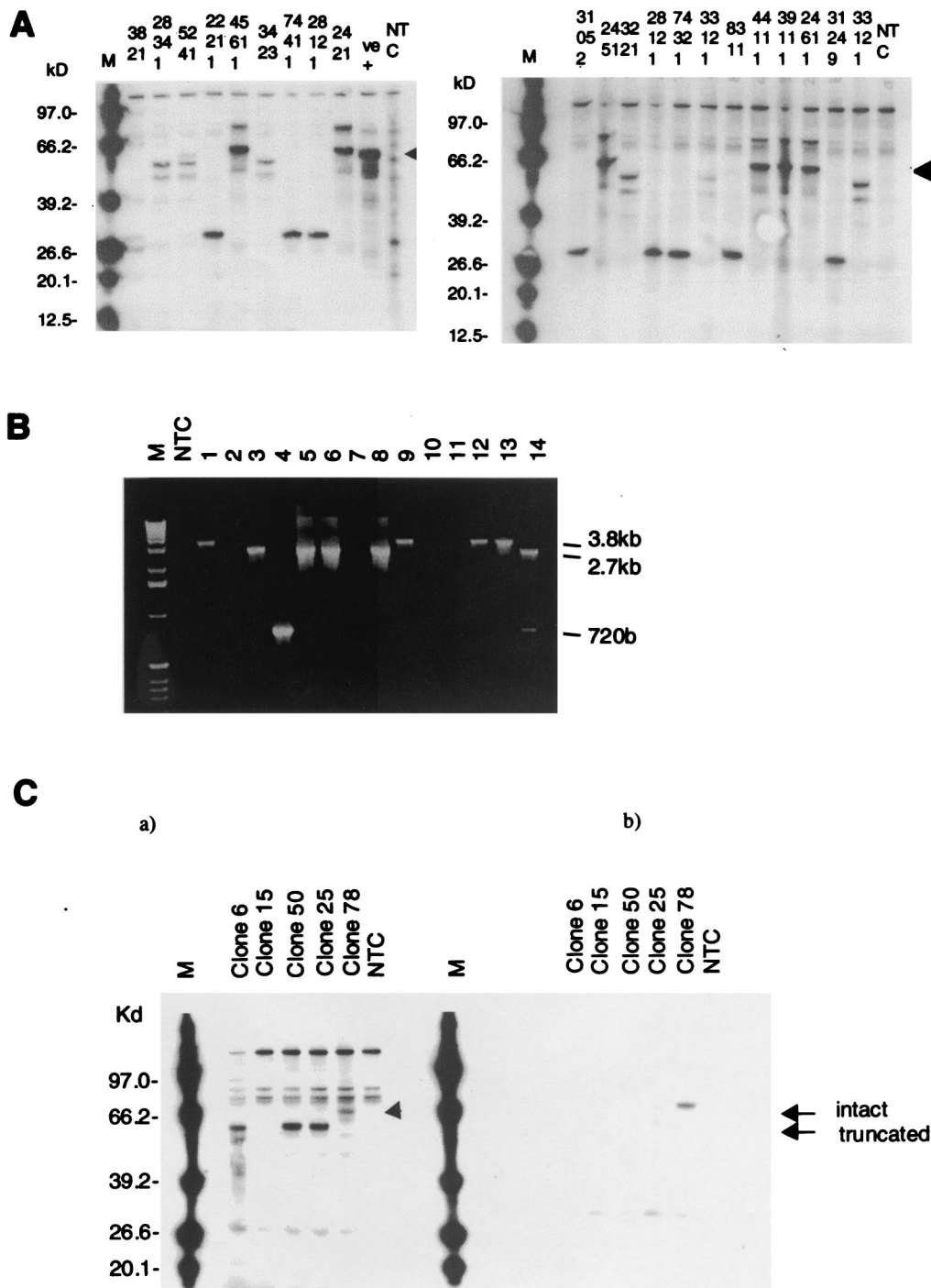


FIG. 1. Sample results for the PCR and PTT assays for *gag*, *pol*, and *env* ORFs. PCR and PTT were carried out as described in Materials and Methods for *gag*, *pol*, and *env*. Selected results are shown. Full results are shown in Tables 1 and 2. (A) *gag* analysis. The size of the full-length product (arrowhead), at ~60 kDa, is as predicted from translation of the sequence (59.5 kDa) and is the same size as the positive control. Full-length products were obtained from lambda clones 45611, 2421, 2451, 44111, 39111, and 24611. Products obtained from other clones were smaller than this and were recorded as truncated proteins; the provirus would therefore be defective. M, biotinylated protein marker (Roche); +ve, control sequenced full-length PERV C *gag* plasmid clone; NTC, no template control. (B) *pol* analysis. *pol* PCR on cosmid clones gave the following amplification products: negative (lanes 2, 10, and 11); 2.7-kb products as predicted for the *pol* ORF (lanes 3, 5, 6, and 8); ~3.8-kb products, larger than the predicted product (lanes 1, 9, 12, and 13); ~720-bp product, smaller than the predicted product (lane 5); and both 2.7-kb and ~720-bp products (lane 14). (C) *env* analysis. Results are shown for five cosmids. (a) Conventional PTT shows that cosmid 78 yielded an ~74-kDa product as predicted for the 73-kDa full-length PERV B *env*, with the ~1-kDa epitope tag (arrow, intact). Clones 6, 25, and 50 gave truncated products at ~63 kDa (arrow, truncated), whereas clone 15 gave a product of <20 kDa that is not visible on this gel. (b) Epitope-tagged PTT showed a strong signal only from clone 78, at ~73 kDa, confirming the presence of the full-length protein identified in Fig. 1C, panel a. Signal is not present from products generated from the other clones, confirming the presence of a premature stop codon.

TABLE 3. Analysis by Southern blot of cosmid library clones that failed to give PCR products for *gag*, *pol*, or *env*^a

Clone	PBS band size (kb)	<i>env</i> band size (kb)	Comments
12	0.5, 0.65	—	No <i>env</i> signal: defective
13		—	No signal for PBS <i>gag</i> or <i>env</i> LTR: defective
19	1.2, 2.5	1.2	<i>env</i> and PBS signal on the same fragment; therefore, deleted
30		—	No signal for PBS <i>gag</i> or <i>env</i> LTR: defective
43	2.3	+	<i>env</i> and PBS signal on the same fragment; therefore, deleted
53		—	No signal for PBS <i>gag</i> or <i>env</i> LTR: defective
59	>6.0 ^b	—	No <i>env</i> signal: defective
62	>6.0 ^b	—	No <i>env</i> signal: defective
66	>6.0 ^b	—	No <i>env</i> signal: defective
69		—	No signal for PBS <i>gag</i> or <i>env</i> LTR: defective
70		—	No signal for PBS <i>gag</i> or <i>env</i> LTR: defective
84	1.3	+	<i>env</i> and PBS signal on the same fragment; therefore, deleted

^a Cosmid DNA was digested with *Sma*I, electrophoresed, and subjected to Southern blotting. Blots were probed with a biotin-labelled probe against either PBS *gag* or *env* LTR, followed by detection of probe with streptavidin-HRP conjugate and ECL detection as described in Materials and Methods.

^b Product was too large to determine size accurately.

not amplify the *gag*, *pol*, or *env* primer sets. To exclude the possibility that these represented major sequence variants on the published PERV proviruses and hence were not amplified by the PCR primers used, these were further analyzed in one of two ways. (i) Sequencing from the cloning sites into the insert was used to determine if the provirus had been truncated during the cloning process. Six of 54 cosmid clones had internal proviral sequence adjacent to the cosmid cloning site due to such deletions (Table 1). No definitive analysis of these clones could therefore be completed. (ii) Southern blots of *Sma*I-digested cosmid DNA were probed with biotin-labelled oligonucleotides homologous to the region around the PBS and *env* LTR junction. It was assumed that, due to the sequence conservation seen within these regions in all known PERV variants, failure to bind these oligonucleotides would indicate a deleted provirus. Most of these clones showed no signal from either of these probes, indicating that they had deletions at both ends (Table 3). The remainder had deletions either at one end or the other. Cosmid clone 74 showed signals for both PBS and *env* LTR; however, sequence data showed a deletion of the 3' end of the *env* gene (not shown).

Analysis of flanking sequence and proviral prevalence. Sequence data was obtained for the LTR-flanking DNA junction for each unique, intact clone identified from the cosmid and lambda libraries by use of the linker-mediated PCR method. For most clones 3'-flanking sequence was obtained (cosmid 78 and lambda clones 65321, 2441, 73414, and 35121). For lambda clone 21321, 5'-flanking sequence was obtained and for 310518 both 3'- and 5'-flanking sequences were identified. Sequence comparisons of a portion of the LTR-flanking sequences obtained by this method are shown in Fig. 2.

Lambda clones 65321, 2441, and 310518 contained a common flanking sequence, confirming that they represented a

single duplicated provirus. Lambda clones 73414 and 35121 both gave two different LTR-3'-flanking sequences (termed 73414i and 73414ii and 35121i and 35121ii, respectively); 73414ii and 35121ii were identical, while the other sequences showed no similarity (Fig. 2).

As the method used to generate the data was direct PCR or sequencing, the presence of two proviruses on a single clone means that we cannot distinguish between 73414 and 35121. These clones may represent a total of three infectious proviruses, complementary defective proviruses, or a combination of these possibilities. We treated them as intact proviruses and performed further analysis on them to determine their prevalence.

The 3'-flanking sequence of cosmid 78 was not related to any of the flanking sequences identified from the lambda clones.

From the 5'-flanking sequence obtained, clones 21321 and 310518 were different, but because 5'-flanking sequence was not obtained for the other clones, it was not possible to determine if 21321 was the same provirus as the 73414-35121 group or cosmid 78. However, although the patterns of prevalence in pigs were identical for 35121 and 21321 (Table 4), there are a number of nucleotide and amino acid differences in Env that confirm that they represent different proviruses (Fig. 3).

Primer pairs were designed for each of these unique LTR-flanking DNA junction sequences to analyze 30 unrelated pigs for the presence of each of the proviruses. The herd from which the test animals were selected was Imutran Ltd's closed genetic herd of pigs transgenic for the human complement regulator hDAF. These pigs have previously been used in Imutran Ltd's xenotransplantation work (7). Thirty animals were selected to exclude siblings and closely related animals and to include as many ancestors (different parents and grandparents) as possible. This selection of animals on pedigree increased the number of variable genotypes tested compared to a random selection of animals, because commonality by descent was minimized. Suitability of the DNA for PCR amplification was confirmed by using porcine β -globin primers, and the results are summarized in Table 4.

The PERV cosmid 78 provirus was present in 7 of 30 pigs; 23 of 30 pigs were negative for this provirus.

Provirus 73414i was present in 3 of 30 pigs, while provirus 73414ii was in 30 of 30 pigs. Provirus 35121i was present in 24 of 30 pigs. A smaller ~200-bp fragment was produced from lambda clone 35121, the positive control, which was also visible in some of the pig PCRs. As this did not interfere with the assay, it was disregarded and was not sequenced to determine its origin.

For provirus 310518, independent PCRs were performed for the 5' and 3' flanks. These gave concordant results in 29 of 30 pigs; 23 of 30 pigs were positive for the provirus and 6 of 30 were negative. Pig H539 was positive for the 3'-LTR-flanking junction and negative for the 5'-LTR-flanking junction, thus giving an ambiguous result. An attempt to use the primers in the flanking DNA to PCR amplify the preintegration site failed, giving an irrelevant PCR product which was sequenced to reveal an amplified porcine repeat sequence (data not shown). Provirus 21321 was present in 24 of 30 pigs.

***env* sequence analysis.** Sequencing of *env* genes on all clones containing proviruses that appeared intact as shown in Tables 1 and 2 showed that each contained a full-length Env ORF,

a)

```

151                                     200
*****3'-LTR*****
65321 cgctcttcc gagcccgac gaggggatt gttcttatac tggcctttca
2441  cgctcttcc gagcccgac gaggggatt gttcttatac tggcctttca
310518 cgctcttcc gagcccgac gaggggatt gttcttatac tggcctttca
73141ii cgctcttcc gagcccgac gaggggatt gttcttttac tggcctttca
35121ii cgctcttcc gagcccgac gaggggatt gttcttttac tggcctttca
73414i  cgctcttcc gagcccgac gaggggatt gttcttttac tggcctttca
35121i  cgctcttcc gagcccgac gaggggatt gttcttttac tggcctttca
78      cgctcttcc gagcccgac gaggggatt gttcttttac tggcctttca

201                                     250
65321 tgacaatcca tagtaaaagt aaaataatcc cattgctttt tcttaatgaa
2441  tgacaatcca tagtaaaagt aaaataatcc cattgctttt tcttaatgaa
310518 tgacaatcca tagtaaaagt aaaataatcc cattgctttt tcttaatgaa
73141ii gtagtgtgtg catgtgcacg tgtgtgcatg tgaagaataa gacaaattcc
35121ii ctagtgtgtg catgtgcacg tgtgtgcatg tgaagaataa gacaaattcc
73414i  ccctcagagc tttgtttaga ctgctgttgc ccttgcattg tagtctggg
35121i  gaactcctgg aagcattggt tttcttccta gtgcacctgg ctataatctt
78      ctatcagaca aaatagacac gaaaataaag actgttaca gagaaaaaga

251                                     300
65321 taaaagtttt gagcatgaat aacattttga gaaaaaatgt actttttctt
2441  taaaagtttt gagcatgaat aacattttga gaaaaaatgt actttttctt
310518 taaaagtttt gagcatgaat aacattttga gaaaaaatgt actttttctt
73141ii atcatcagca agtcacttaa tttctctgga cttcattttc actccataa
35121ii atcatcagca agtcacttaa tttctctgga cttcattttc actccataa
73414i  tatccagtc ccattagtag ggttggcgag caacctgct tcccctagag
35121i  ggacaagtta gttgattttt ctaagtttta gaattgtgt taggtaata
78      agggcactac atcatgatca agggctcaat ccaagattta acaatttgaa

301                                     350
65321 ttgggaaacg gctgttttca attgaattgt tattaacctt ttcattataa
2441  ttgggaaacg gctgttttca attgaattgt tattaacctt ttcattataa
310518 ttgggaaacg gctgttttca attgaattgt tattaacctt ttcattataa
73141ii acagttattt cttaaagtat tgtaattgatt agagaaaatt tactttaaat
35121ii acagttattt cttaaagtat tgtaattgatt agagaaaatt tactttaaat
73414i  ctttctgca gactgatttc ctttcagtct cttgctcaat atggtctcta
35121i  tgaattatta aatgtatttc ccctaattta tatagtact atcaaattaa
78      ataaaaatgc acccaacata ggatcacctc aacatgtaag gcaaatgta

b)

101                                     150
*****5'-LTR*****
21321 acctatttac tggcagccag cagggtctgg aattcggggg catttagagc
310518 acctatttac tggcagccag cagggtctgg aattcggggg catttagagc

151                                     200
*****
21321 tttttattaa cttcctgggt tctgggaggg ttaggttgca tttcatcct
310518 tttttattaa cttcctgggt tctgggaggg ttaggttgca tttcatcct

201                                     250
*****
21321 ttcagttcct gctgtggaat ggtgggttaa ggatccagtg ttgctacagc
310518 ttcagtcaaa aattaaagtg cttcaaaata gcatatgta gagaaatgt

251
21321 tatggagtag gtcacagacc agccggggcg tgcaccagc gtg
310518 atatggaggg gatgcttata cattgctgct taaaataaa att

```

FIG. 2. Analysis of proviral flanking sequences. Flanking sequence was obtained for all clones by the linker-mediated PCR method. For most clones 3'-flanking sequence was obtained (cosmid 78 and lambda clones 65321, 2441, 73414, and 35121) (GenBank accession no. AY056030 to -32 and AY056034 to -35), for lambda clone 21321 5'-flanking sequence (GenBank accession no. AY056033) was obtained, and for 310518 both 3' and 5'-flanking sequences were obtained (GenBank accession no. AY056029). Sequence alignments were performed with the Wisconsin GCG Sequence analysis package, version 9 Pileup program. Pileups of a portion of either the 3'-LTR (a) or the 5'-LTR (b) flanking junctions are shown above. Asterisks above the sequence denote the U3(a) or U5 (b) region; sequence beyond this is flanking sequence.

TABLE 4. Prevalence of six potentially full-length PERV B proviruses in 30 unrelated Large White pigs^a

Pig no.	Provirus tested						
	78	35121i	310518-3'	310518-5'	21321	73414i	73414ii
H370	-	++	++	++	++	-	++
H225	-	++	++	++	++	-	++
G838	-	++	++	++	++	-	++
H768	-	-	++	++	-	-	++
G759	++	++	++	++	++	-	++
G597	-	++	++	++	++	-	++
G585	-	++	++	++	++	-	++
G473	++	++	++	++	++	-	++
G116	-	++	++	++	++	-	++
F934	-	++	++	++	++	-	++
F894	-	-	-	-	-	-	++
F800	-	++	++	++	++	++	++
F748	-	++	++	++	++	-	++
E765	-	-	++	++	-	-	++
D681	-	++	++	++	++	-	++
H744	++	++	++	++	++	-	++
H800	-	++	++	++	++	-	++
H679	-	-	++	++	-	++	++
H338	++	-	-	-	-	-	++
H863	-	++	++	++	++	-	++
H522	++	++	++	++	++	-	++
H888	-	-	++	++	-	-	++
H874	-	++	-	-	++	-	++
H539	-	++	++ ^b	-	++	-	++
H791	-	++	-	-	++	-	++
H765	-	++	-	-	++	++	++
H858	-	++	-	-	++	-	++
H440	++	++	++	++	++	-	++
H588	-	++	++	++	++	-	++
G768	++	++	++	++	++	-	++
Total	7/30	24/30	24/30	23/30	24/30	3/30	30/30

^a PCR screening for the presence of each provirus was carried out as described in Materials and Methods. ++, provirus was present in the pig, -, provirus was not present.

^b This result showed a discrepancy between the 3'- and 5'-flanking LTR for provirus 310518. It is possible that this provirus may exist as two alleles, one with a partial deletion at the 5' LTR.

confirming the PTT results. Comparison of the proviral *env* with published PERV sequences showed a high degree of homology (>99%) with the published PERV B sequence Y12239 at both the nucleotide and amino acid levels (Fig. 3). The single sequence labelled 2441 is the consensus sequence obtained from clones 65321, 2441, and 310518 *env* sequences. Sequencing of *env* from lambda clones 73414 and 310518 gave only a single *env* consensus sequence from each, despite the two different 3' LTRs that were present in these clones.

The receptor-interacting VRA and VRB regions in the middle of the Env protein showed no variation from the prototype sequence, although there were a few polymorphisms between the *env* genes that confirmed that they represented different proviruses (Fig. 3).

In summary, all clones were found to be intact in all three ORFs, as confirmed by PTT and sequencing. Sequence analysis of the clones with published PERV variants showed high levels of homology between them, as expected.

DISCUSSION

While extensive analysis of patients treated with living porcine tissue has shown that none of these patients has been

infected by PERV, the future of porcine xenotransplantation may be dependent on using pigs with the smallest possible load of PERV proviruses. As a first step toward the objective of selecting such pigs, we have identified a number of potentially full-length, functional PERV proviruses and isolated their flanking sequences to enable their chromosomal locations to be identified. In addition, we have analyzed their prevalence in 30 Large White pigs transgenic for the human complement regulator hDAF.

Our strategy was designed to identify full-length proviral inserts encoding potentially infectious PERV A and B viral subtypes from a cosmid and a lambda library, respectively. Both libraries were derived from DNA of the same pig. These clones were coscreened by using both *gag* and *env* probes to increase the probability of isolating full-length proviruses and to rapidly eliminate grossly defective proviruses. Doubly positive clones were then screened using either restriction fragment analysis alone (cosmid library) or restriction fragment analysis with Southern blotting (lambda library) to identify unique and duplicate clones and to confirm the presence of proviral sequences. Subsequently, clones were screened by PCR and PTT to identify clones containing full-length Gag and Env ORFs, and each provirus was sequenced to confirm they were intact. Sequence flanking the proviral integration site of apparently intact proviruses was isolated and used to design primer pairs to screen pigs for the presence of each provirus. The data are summarized in Tables 1 and 2.

From the original cosmid-screening filters, only 10 to 12% of clones that gave *gag*-positive signals were also *env* positive (data not shown). This is due mainly to the nonvariant specificity of the *gag* probe, which would have detected all PERV variants, compared to the *env* probes, which were variant specific. We attempted to differentially screen for PERV A and B, using probes in the VRA and VRB regions that showed the most difference between the two. Although the PERV A and B probes showed specificity on Southern blots, they cross-hybridized to some extent under the conditions employed for cosmid screening. Consequently, it is likely that we would have detected proviruses with less homology to the published PERV A and B sequences. Such sequence variants have been identified by other groups (27, 51). Additionally, we detected a PERV B provirus in the PERV A-screened cosmid library, which presumably reflects some degree of probe cross-reactivity.

Most of the proviruses had *pol* sequences that were amplifiable by PCR (Fig. 1 and 2). We identified some polymorphism in this product, obtaining fragments of three different sizes. Sequencing revealed the smallest product to be a *pol*-related fragment with multiple deletions, proviruses giving this fragment would be defective. One product was of the size predicted for published PERV *pol* sequences, while the other fragment was larger than the published ones. Sequencing the ends of both of these products showed that they both had >96% *pol* homology and could both potentially encode functional *pol*, although the larger products were not fully sequenced since proviruses with this *pol* sequence were defective in either *gag* or *env*. Clone 78, the only provirus in the cosmid library that gave both full length *gag* and *env* ORFs by PTT, gave a *pol* product of the predicted size from published *pol* genes.

On sequencing the *env* gene of the provirus on cosmid 78, it

```

Y12239      LTWLIIDPDTGVTNSTRGVAPRGTWPELHFCRLRLINPAVKSTPPNLVRSYGFYCCPGT
78          LTWLIIDPDTGVTNSTRGVAPRGTWPELHFCRLRLINPAVKSTPPNLVRSYGFYCCPGT
2441       LTWLIIDPDTGVTNSTRGVAPRGTWPELHFCRLRLINPAVKSTPPNLVRSYGFYCCPGT
73414      LTWLIIDPDTGVTNSTRGVAPRGTWPELHFCRLRLINPAVKSTPPNLVRSYGFYCCPGT
35121      LTWLIIDPDTGVTNSTRGVAPRGTWPELHFCRLRLINPAVKSTPPNLVRSYGFYCCPGT
21321      LTWLIIDPDTGVTNSTRGVAPRGTWPELHFCRLRLINPAVKSTPPNLVRSYGFYCCPGT
*****

Y12239      EKEKYCGSGESFCRRWSCVTSNDGDWKWPI SLQDRVKFSFVNSGPGKYKVMKLYKDKSC
78          EKEKYCGSGESFCRRWSCVTSNDGDWKWPI SLQDRVKFSFVNSGPGKYKVMKLYKDKSC
2441       EKEKYCGSGESFCRRWSCVTSNDGDWKWPI SLQDRVKFSFVNSGPGKYKVMKLYKDKSC
73414      EKEKYCGSGESFCRRWSCVTSNDGDWKWPI SLQDRVKFSFVNSGPGKYKVMKLYKDKSC
35121      EKEKYCGSGESFCRRWSCVTSNDGDWKWPI SLQDRVKFSFVNSGPGKYKVMKLYKDKSC
21321      EKEKYCGSGESFCRRWSCVTSNDGDWKWPI SLQDPVKFSFVNSGPGKYKVMKLYKDMSC
*****

Y12239      SPDDL DYLKISFTEK GKQENIQKWINGMSWGI VFYKYGGGAGSTLTIRLRIETGTEPPVA
78          SPDDL DYLKISFTEK GKQENIQKWINGMSWGI VFYKYGGGAGSTLTIRLRIETGTEPPVA
2441       SPDDL DYLKISFTEK GKQENIQKWINGMSWGI VFYKYGGGAGSTLTIRLRIETGTEPPVA
73414      SPDDL DYLKISFTEK GKQENIQKWINGMSWGI VFYKYGGGAGSTLTIRLRIETGTEPPVA
35121      SPDDL DYLKISFTEK GKQENIQKWINGMSWGI VFYKYGGGAGSTLTIRLRIETGTEPPVA
21321      SPDDL DYLKISFTEK GKQENIQKWINGMSWGI VFYKYGGGAGSTLTIRLRIETGTEPPVA
*****

Y12239      VGPDKVLAEQGPPALEPPHNLFPVQLTSLRPDITQPPSNGTTGLIPTNTPRNSPGVPVKT
78          VGPDKVLAEQGPPALEPPHNLFPVQLTSLRPDITQPPSNGTTGLIPTNTPRNSPGVPVKT
2441       VGPDKVLAEQGPPALEPPHNLFPVQLTSLRPDITQPPSNGTTGLIPTNTPRNSPGVPVKT
73414      VGPDKVLAEQGPPALEPPHNLFPVQLTSLRPDITQPPSNGTTGLIPTNTPRNSPGVPVKT
35121      VGPDKVLAEQGPPALEPPHNLFPVQLTSLRPDITQPPSNGTTGLIPTNTPRNSPGVPVKT
21321      VGPDKVLAEQGPPALEPPHNLFPVQLTSLRPDITQPPSNGTTGLIPTNTPRNSPGVPVKT
*****

Y12239      GQRLFSLIQGAFAINSTDPDATSSCWLC LSSGPPYYEGMAKEGKFNVTKEHRNQCTWGS
78          GQRLFSLIQGAFAINSTDPDATSSCWLC LSSGPPYYEGMAKEGKFNVTKEHRNQCTWGS
2441       GQRLFSLIQGAFAINSTDPDATSSCWLC LSSGPPYYEGMAKEGKFNVTKEHRNQCTWGS
73414      GQRLFSLIQGAFAINSTDPDATSSCWLC LSSGPPYYEGMAKEGKFNVTKEHRNQCTWGS
35121      GQRLFSLIQGAFAINSTDPDATSSCWLC LSSGPPYYEGMAKEGKFNVTKEHRNQCTWGS
21321      GQRLFSLIQGAFAINSTDPDATSSCWLC LSSGPPYYEGMAKEGKFNVTKEHRNQCTWGS
*****

Y12239      RNKLTLETVSGKGTICIGKAPPSHQHLCYSTVVYEQASENQYL VPGYNRWACNTGLTPCV
78          RNKLTLETVSGKGTICIGKAPPSHQHLCYSTVVYEQASENQYL VPGYNRWACNTGLTPCV
2441       RNKLTLETVSGKGTICIGKAPPSHQHLCYSTVVYEQASENQYL VPGYNRWACNTGLTPCV
73414      RNKLTLETVSGKGTICIGKAPPSHQHLCYSTVVYEQASENQYL VPGYNRWACNTGLTPCV
35121      RNKLTLETVSGKGTICIGKAPPSHQHLCYSTVVYEQASENQYL VPGYNRWACNTGLTPCV
21321      RNKLTLETVSGKGTICIGKAPPSHQHLCYSTVVYEQASENQYL VPGYNRWACNTGLTPCV
*****

Y12239      STSVFNQSKDFCFVMVQIVPRVYHPEEVVLD EYD YRNRPKREPVSLTLAVMLGLGTAVG
78          STSVFNQSKDFCFVMVQIVPRVYHPEEVVLD EYD YRNRPKREPVSLTLAVMLGLGTAVG
2441       STSVFNQSKDFCFVMVQIVPRVYHPEEVVLD EYD YRNRPKREPVSLTLAVMLGLGTAVG
73414      STSVFNQSKDFCFVMVQIVPRVYHPEEVVLD EYD YRNRPKREPVSLTLAVMLGLGTAVG
35121      STSVFNQSKDFCFVMVQIVPRVYHPEEVVLD EYD YRNRPKREPVSLTLAVMLGLGTAVG
21321      STSVFNQSKDFCFVMVQIVPRVYHPEEVVLD EYD YRNRPKREPVSLTLAVMLGLGTAVG
*****

Y12239      VGTGTAALITGPQOLEKGLGELHAAMTE DLRALEESVSNLEESLTSLSEVVLQNRRLDLD
78          VGTGTAALITGPQOLEKGLGELHAAMTE DLRALEESVSNLEESLTSLSEVVLQNRRLDLD
2441       VGTGTAALITGPQOLEKGLGELHAAMTE DLRALEESVSNLEESLTSLSEVVLQNRRLDLD
73414      VGTGTAALITGPQOLEKGLGELHAAMTE DLRALEESVSNLEESLTSLSEVVLQNRRLDLD
35121      VGTGTAALITGPQOLEKGLGELHAAMTE DLRALEESVSNLEESLTSLSEVVLQNRRLDLD
21321      VGTGTAALITGPQOLEKGLGELHAAMTE DLRALEESVSNLEESLTSLSEVVLQNRRLDLD
:****:*****

Y12239      LFLREGGLCAALKEECCFYVDHSGAIRDSMSKLRERLERRRREREADQGWFEWFNRSFW
78          LFLREGGLCAALKEECCFYVDHSGAIRDSMSKLRERLERRRREREADQGWFEWFNRSFW
2441       LFLREGGLCAALKEECCFYVDHSGAIRDSMSKLRERLERRRREREADQGWFEWFNRSFW
73414      LFLREGGLCAALKEECCFYVDHSGAIRDSMSKLRERLERRRREREADQGWFEWFNRSFW
35121      LFLREGGLCAALKEECCFYVDHSGAIRDSMSKLRERLERRRREREADQGWFEWFNRSFW
21321      LFLREGGLCAALKEECCFYVDHSGAIRDSMSKLRERLERRRREREADQGWFEWFNRSFW

```

Downloaded from jvi.asm.org by on July 6, 2009

FIG. 3. Sequence alignment of predicted Env proteins from full-length proviruses. Sequencing of *env* genes of apparently intact proviruses confirmed the presence of a full-length Env ORF, as predicted by the PTT results. Comparison of the proviral *env* sequence showed a high degree of homology (>99%) with the prototype PERV B sequence Y12239 at both the nucleotide and amino acid levels. The single sequence labelled 2441 is the consensus sequence obtained from clone 65321, 2441, and 310518 *env* sequences. Sequencing of *env* from lambda clones 73414 and 310518 gave only a single *env* consensus sequence, despite the two different 3' LTRs that were present in these clones, suggesting that they were either identical or that one provirus did not have an intact *env* gene (GenBank accession no. AY056024 to -28).

was apparent that it was a PERV B provirus (Fig. 2), despite the fact that it was identified using a PERV A *env* probe. The reason for this is not clear, as we did not identify the other PERV B proviruses that were found in the lambda library from the cosmid library. Conversely, the clone 78 PERV B provirus was not isolated from the lambda library. This possibly reflects the fact that the PERV A *env* probe could cross-hybridize with PERV B *env* to some extent. It is also possible that as filters were washed in bulk, the part of the library screening filter that clone 78 was isolated from was not subjected to a stringent enough wash after hybridization, hence cross-reactivity was observed. The failure to identify the cosmid 78 PERV B provirus in the lambda library is cause for more concern. It is possible due to the screening strategy used that lambda clones that contained the cosmid 78 provirus were all truncated during library construction and despite the three to four genome equivalents of library that were screened, the provirus would not be identified by this approach.

Unlike the results of the *pol* PCR assay, in the *gag* and *env* PCR screen only a relatively small fraction of clones isolated from the cosmid library yielded products (Table 1). This is unlikely to be due to small sequence changes that prevented PCR primer extension, as all published PERV A-, B-, and C-related proviruses have highly homologous sequences around the start and stop codons of all three viral genes. Additionally, we have successfully used a single primer pair to amplify *gag* and *env* genes of all currently known PERV variants under the low target complexity used here (not shown). It is likely that failure to produce a PCR product is related to gross mutations or deletions around the PBSs. This was supported by the Southern blot data shown in Table 3, where most of the clones that failed to yield PCR products failed to bind probes encompassing either the PBS or *env* LTR junction.

Analysis of the clones isolated from the lambda and cosmid library showed that there were significant differences between the clones from both libraries. Results for PERV B isolated from the lambda library showed some clear differences from the PERV A clones from the cosmid library as described above (compare Tables 1 and 2). First, most clones gave PCR products for both *gag* and *env*, indicating that more of the PERV B clones were intact at the gross sequence level. Approximately 50% of these clones gave full-length *gag* or *env* translation products as determined by PTT. Six clones gave both *gag* and *env* translation products, all of which gave an appropriately sized *pol* PCR product. Analysis of the flanking sequences from lambda clones 65321, 2441, and 310518 demonstrated that they had a common flanking sequence and represented a single provirus. Lambda clones 73414 and 35121 each gave two different LTR-3'-flanking sequences (termed 73414i and 73414ii and 35121i and 35121ii, respectively); 73414ii and 35121ii were identical, whereas 73414i and 35121i sequences showed no similarity. The 21321 5'-flanking sequence was not the same as the 310518 sequence, indicating that this represented another provirus. Thus, we identified five unique PERV B proviruses from the lambda library together with one PERV B clone from the cosmid library on the basis of their differing flanking sequences (Fig. 2). These flanking sequences were not the same as those identified by Deng et al. (9). It is also unclear if the flanking sequences isolated by Deng et al. were associated with a full-length proviral insert (9). Only one *env* se-

quence was obtained from clones 73414 and 35121, indicating that there were either two identical *env* sequences or that one provirus was defective and yielded no *env* product. The six *env* sequences showed that they have a high (>99%) degree of homology with the prototype PERV B sequence Y12239 at both the amino acid level (Fig. 3) and the DNA level (not shown).

As described above, we did not detect any intact PERV A proviruses in this pig. These differences between the results for PERV A and B cannot be explained simply by differences between clones isolated from the lambda and cosmid libraries, as the cosmid library contained multiple defective PERV A proviruses and we isolated an intact PERV B provirus, clone 78. Due to the larger insert size of the cosmid library, it is less likely that proviruses would be truncated during library construction than those used for the lambda library. Krach et al. (21) were able to isolate an infectious PERV A clone from PK15 cells, a cell type that is known to express high levels of infectious PERV A and B. There are two main possibilities to account for the absence of infectious PERV A in this pig. Data derived by coculture of pig aortic endothelial cells with permissive cells for replication suggest that PERV infectivity differs between different pigs (D. Cunningham et al. unpublished data). This implies that the number of infectious proviruses is low, as Krach et al. (21) have also suggested, and that there is significant diversity between individuals, as we have found (see below). Alternatively, full-length proviruses encoding infectious PERV A may be absent from this pig and infectious PERV A may be generated by recombination with other PERV sequences present during infection cycles. Wilson et al. (51) demonstrated that passage of their infectious virus PERV-NIH through human cells generated clones that when sequenced were found to be recombinants of PERV A and PERV C.

Retroviruses are diploid, and virions can undergo complex recombinations involving more than one internal template (switch displacement/assimilation or forced copy choice model) during reverse transcription (16). During this process, packaged defective A genomes could undergo recombination with subgroup B genomes to generate an infectious subgroup A virus. Recombinations of this form are known to be required for the generation of polytropic murine leukemia virus (42) and the subgroup B feline leukemia virus (41). It is also possible that, due to the relatively poor superinfection resistance conferred by PERV Env expression in PERV-infected cells, some of the proviruses examined by Krach et al. (21) were derived through reinfection cycles in the long-established PK15 cell line. It is possible that pig tissues or primary cell lines may differ in this respect.

The number of intact, unique proviruses identified here is broadly in agreement with other data on numbers of intact proviruses identified in Large White pigs. Bosch et al. (6) analyzed PERV RNA from Large White pigs and found 11 different transcripts, of which only 4 had intact *env* genes. No data were published on *gag* or *pol* genes. Additionally, other nonexpressed proviruses may have been present in this pig. We currently do not know if the six unique proviruses we have identified are expressed in Large White pigs and which, if any, correspond to the proviral transcripts noted by Bosch et al. (6). Analysis of the LTRs of these proviruses is currently under

way. Krach et al. (21) identified one PERV A and two PERV B infectious proviruses from their screen of a genomic library of PK15 cells. We do not currently know if the six proviruses identified in this work are infectious.

A caveat of our screening approach is that if there are two or more proviral sequences on the same clone, we cannot determine if either is intact, due to cross-complementation in both the PCR and sequencing reactions. As noted above, both clones 73414 and 35121 appear to be contiguous and may each contain two defective, but complementary, proviruses based on the results of LTR analysis. This question will be answered by further work on these two clones. Additionally, as discussed above, viral complementation and recombination in other species have been shown to lead to the emergence of infectious virus and may be the origin for infectious PERV A in some pigs. In addition to removing full-length functional proviruses, it may be necessary to examine the possibility that recombination between defective proviruses may generate infectious viruses after infection. Further studies on the relative expression rates of defective genomes and how efficiently they are copackaged is needed to address this point.

We used LTR-flanking sequence to design primer pairs to determine the prevalence of the six proviruses within 30 unrelated Large White hDAF pigs. As shown in Table 4, three patterns of prevalence were observed: 73414ii was present in all pigs examined; 35121i, 310518, and 21321 were common, but some pigs did not have them; and cosmids 78 and 73414i were uncommon. This pattern may reflect either founder effects within the herd or the fact that the more common proviruses were integrated farther back in the evolution of the domestic pig. Ancient proviruses in the founder animals will be present in all offspring, while more recently integrated proviruses may show a more sparse distribution. The degree of heterogeneity between pigs is similar to that noted for PERV C (6,17). Interestingly, sequence analysis of the LTR-flanking sequences showed that lambda clones 73414 and 35121 are contiguous, sharing LTR-flanking sequence 73414ii, while 35121i and 73414i were unique to each respective lambda clone. This indicates that there are three almost concatameric PERV proviruses at this locus, implying a preferred integration site. Preferred targets for retroviral integration have been suggested previously. Shih et al. (39) identified that within these targets, independent integration events can occur at the same nucleotide base.

We have demonstrated that of the approximately 50 PERV proviruses enumerated using a *pro* probe (22), only six unique intact PERV B proviruses and no intact PERV A proviruses were present in the genome of the Large White pig studied. Thus, the majority of proviruses are defective in one or more of the viral genes. No PERV C env sequences were detectable in this pig by PCR. However, there inevitably remain a large number of proviruses that have one or two ORFs, which may complement each other. It is possible that some of the PERV proteins may have functions, such as protection against reinfection with exogenous PERV viruses, perhaps by providing superinfection resistance (3) or by Fv-type restriction mechanisms (47).

Regardless of speculation, these data indicate that it would be straightforward to remove proviruses 78 and 73414i by breeding. Whether provirus 73414ii can be removed by breed-

ing will require a survey of a larger number of pigs to determine if there is variation at this locus. The recent developments in pig cloning (4, 29, 35) using nuclear transfer may also enable the genetic knockout of the more ubiquitous proviruses, whereas breeding out would necessarily be a more protracted process. The remaining proviruses show sufficient variation to indicate that a breeding strategy to remove them should be possible. In this regard 2 of the 30 pigs examined (F894 and H338) had much lower proviral numbers. However, it is recognized that selection purely on the basis of these five proviruses may fix other proviruses that are not represented in this pig but are present in other pigs. Consequently, a breeding strategy will require detailed analysis, using a similar strategy, of the proviruses present in a number of founder pigs. This will ensure that all intact proviruses present in the founder herd are identified.

ACKNOWLEDGMENTS

This work was funded by Imutran Ltd., a Novartis Pharma company.

Thanks are due to Grahame McKenzie for the cosmid and lambda libraries and Eleanor Lees-Jones for the identification of unrelated and outbred Large White pigs.

REFERENCES

- Akiyoshi, D. E., M. Denaro, H. Zhu, J. L. Greenstein, P. Banerjee, and J. A. Fishman. 1998. Identification of a full-length cDNA for an endogenous retrovirus of miniature swine. *J. Virol.* 72:4503–4507.
- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.). 1997. Current protocols in molecular biology. John Wiley & Sons, Ltd., New York, N.Y.
- Berkhauser, R. D., and S. P. Goff. 1993. Point mutations in Moloney murine leukemia virus envelope protein: effects on infectivity, virion association, and superinfection resistance. *Virology* 196:748–757.
- Bethhauser, J., E. Forsberg, M. Augenstein, L. Childs, K. Eilertsen, J. Enos, T. Forsythe, P. Golueke, G. Jurgella, R. Koppang, T. Lesmeister, K. Mallon, G. Mell, P. Misica, M. Pace, M. Pfister-Genskow, N. Strelchenko, G. Voelker, S. Watt, S. Thompson, and M. Bishop. 2000. Production of cloned pigs from in vitro systems. *Nat. Biotechnol.* 18:1055–1059.
- Blusch, J. H., C. Patience, Y. Takeuchi, C. Templin, C. Roos, H. Von Der, G. Steinhoff, and U. Martin. 2000. Infection of nonhuman primate cells by pig endogenous retrovirus. *J. Virol.* 74:7687–7690.
- Bosch, S., C. Arnauld, and A. Jestin. 2000. Study of full-length porcine endogenous retrovirus genomes with envelope gene polymorphism in a specific-pathogen-free Large White swine herd. *J. Virol.* 74:8575–8581.
- Cozzi, E., F. Bhatti, M. Schmoedel, G. Chavez, K. G. Smith, A. Zaidi, J. R. Bradley, S. Thiru, M. Goddard, C. Vial, D. Ostlie, J. Wallwork, D. J. White, and P. J. Friend. 2000. Long-term survival of nonhuman primates receiving life-supporting transgenic porcine kidney xenografts. *Transplantation* 70:15–21.
- Cunningham, D. A., C. Herring, X. M. Fernández-Suárez, A. J. Whittam, K. Paradis, and G. A. Langford. 2001. Trends in cardiology, in press.
- Deng, Y. M., J. H. Lee, C. Moran, J. H. Jin, B. E. Tuch, and W. D. Rawlinson. 2000. Mapping dispersed repetitive loci using semi-specific PCR cloning and somatic cell hybrid mapping. *Nucleic Acids Res.* 28:E103.
- Deng, Y. M., B. E. Tuch, and W. D. Rawlinson. 2000. Transmission of porcine endogenous retroviruses in severe combined immunodeficient mice xenografted with fetal porcine pancreatic cells. *Transplantation* 70:1010–1016.
- Dinsmore, J. H., C. Manhart, R. Raineri, D. B. Jacoby, and A. Moore. 2000. No evidence for infection of human cells with porcine endogenous retrovirus (PERV) after exposure to porcine fetal neuronal cells. *Transplantation* 70:1382–1389.
- Elliott, R. B., L. Escobar, O. Garkavenko, M. C. Croxson, B. A. Schroeder, M. McGregor, G. Ferguson, N. Beckman, and S. Ferguson. 2000. No evidence of infection with porcine endogenous retrovirus in recipients of encapsulated porcine islet xenografts. *Cell Transplant.* 9:895–901.
- Ericsson, T., B. Oldmixon, J. Blomberg, M. Rosa, C. Patience, and G. Andersson. 2001. Identification of novel porcine endogenous betaretrovirus sequences in miniature swine. *J. Virol.* 75:2765–2770.
- Heneine, W., A. Tibell, W. M. Switzer, P. Sandstrom, G. V. Rosales, A. Mathews, O. Korsgren, L. E. Chapman, T. M. Folks, and C. G. Groth. 1998. No evidence of infection with porcine endogenous retrovirus in recipients of porcine islet-cell xenografts. *Lancet* 352:695–699.
- Herring, C., D. A. Cunningham, A. J. Whittam, X. M. Fernandez-Suarez,

- and G. A. Langford. 2001. Monitoring xenotransplant recipients for infection by PERV. *Clin. Biochem.* **34**:23–27.
16. Hu, W. S., and H. M. Temin. 1992. Effect of gamma radiation on retroviral recombination. *J. Virol.* **66**:4457–4463.
 17. Jin, H. Y., Inoshima, and D. Wu. 2000. Expression of porcine endogenous retrovirus in peripheral blood leukocytes from ten different breeds. *Transplant Infect. Dis.* **2**:11–14.
 18. Kikyo, N., and A. P. Wolffe. 2000. Reprogramming nuclei: insights from cloning, nuclear transfer and heterokaryons. *J. Cell Sci.* **113**:11–20.
 19. Kono, T. 1997. Nuclear transfer and reprogramming. *Rev. Reprod.* **2**:74–80.
 20. Kono, T., Y. Obata, T. Yoshimizu, T. Nakahara, and J. Carroll. 1996. Epigenetic modifications during oocyte growth correlates with extended parthenogenetic development in the mouse. *Nat. Genet.* **13**:91–94.
 21. Krach, U., N. Fischer, F. Czauderna, and R. R. Tonjes. 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.
 22. Le Tissier, P., J. P. Stoye, Y. Takeuchi, C. Patience, and R. A. Weiss. 1997. Two sets of human-tropic pig retrovirus. *Nature* **389**:681–682.
 23. Levy, M. F., J. Crippin, S. Sutton, G. Netto, J. McCormack, T. Curiel, R. M. Goldstein, J. T. Newman, T. A. Gonwa, J. Bancheau, L. E. Diamond, G. Byrne, J. Logan, and G. B. Klintmalm. 2000. Liver allotransplantation after extracorporeal hepatic support with transgenic (hCD55/hCD59) porcine livers: clinical results and lack of pig-to-human transmission of the porcine endogenous retrovirus. *Transplantation* **69**:272–280.
 24. Loss, M., H. Arends, M. Winkler, M. Przemek, G. Steinhoff, S. Rensing, F. J. Kaup, H. J. Hedrich, M. E. Winkler, and U. Martin. 2001. Analysis of potential porcine endogenous retrovirus (PERV) transmission in a whole-organ xenotransplantation model without interfering microchimerism. *Transplant. Int.* **14**:31–37.
 25. Martin, U., G. Steinhoff, V. Kiessig, M. Chikobava, M. Anssar, T. Morschheuser, B. Lapin, and A. Haverich. 1998. Porcine endogenous retrovirus (PERV) was not transmitted from transplanted porcine endothelial cells to baboons in vivo. *Transplant. Int.* **11**:247–251.
 26. Martin, U., M. E. Winkler, M. Id, H. Radeke, L. Arseniev, Y. Takeuchi, A. R. Simon, C. Patience, A. Haverich, and G. Steinhoff. 2000. Productive infection of primary human endothelial cells by pig endogenous retrovirus (PERV). *Xenotransplantation* **7**:138–142.
 27. Onions, D., D. Hart, C. Mahoney, D. Galbraith, and K. Smith. 1998. Endogenous retroviruses and the safety of porcine xenotransplantation. *Trends Microbiol.* **6**:430–431.
 28. Onions, D., D. K. Cooper, T. J. Alexander, C. Brown, E. Claassen, J. E. Foweraker, D. L. Harris, B. W. Mahy, P. D. Minor, A. D. Osterhaus, P. P. Pastoret, and K. Yamanouchi. 2000. An approach to the control of disease transmission in pig-to-human xenotransplantation. *Xenotransplantation* **7**:143–155.
 29. Onishi, A., M. Iwamoto, T. Akita, S. Mikawa, K. Takeda, T. Awata, H. Hanada, and A. C. Perry. 2000. Pig cloning by microinjection of fetal fibroblast nuclei. *Science* **289**:1188–1190.
 30. Paradis, K., G. Langford, Z. Long, W. Heneine, P. Sandstrom, W. M. Switzer, L. E. Chapman, C. Locky, D. Onions, and E. Otto. 1999. Search for cross-species transmission of porcine endogenous retrovirus in patients treated with living pig tissue. *Science* **285**:1236–1241.
 31. Patience, C., G. S. Patton, Y. Takeuchi, R. A. Weiss, M. O. McClure, L. Rydberg, and M. E. Breimer. 1998. No evidence of pig DNA or retroviral infection in patients with short-term extracorporeal connection to pig kidneys. *Lancet* **352**:699–701.
 32. Patience, C., W. M. Switzer, Y. Takeuchi, D. J. Griffiths, M. E. Goward, W. Heneine, J. P. Stoye, and R. A. Weiss. 2001. Multiple groups of novel retroviral genomes in pigs and related species. *J. Virol.* **75**:2771–2775.
 33. Patience, C., Y. Takeuchi, and R. A. Weiss. 1997. Infection of human cells by an endogenous retrovirus of pigs. *Nat. Med.* **3**:282–286.
 34. Pitkin, Z., and C. Mullan. 1999. Evidence of absence of porcine endogenous retrovirus (PERV) infection in patients treated with a bioartificial liver support system. *Artif. Organs* **23**:829–833.
 35. Polejaeva, I. A., S. H. Chen, T. D. Vaught, R. L. Page, J. Mullins, S. Ball, Y. Dai, J. Boone, S. Walker, D. L. Ayares, A. Colman, and K. H. Campbell. 2000. Cloned pigs produced by nuclear transfer from adult somatic cells. *Nature* **407**:86–90.
 36. Roest, P. A., R. G. Roberts, S. Sugino, G. J. van Ommen, and J. T. den Dunnen. 1993. Protein truncation test (PTT) for rapid detection of translation-terminating mutations. *Hum. Mol. Genet.* **2**:1719–1721.
 37. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 38. Schumacher, J. M., S. A. Ellias, E. P. Palmer, H. S. Kott, J. Dinsmore, P. K. Dempsey, A. J. Fischman, C. Thomas, R. G. Feldman, S. Kassissieh, R. Raineri, C. Manhart, D. Penney, J. S. Fink, and O. Isacson. 2000. Transplantation of embryonic porcine mesencephalic tissue in patients with PD. *Neurology* **54**:1042–1050.
 39. Shih, C. C., J. P. Stoye, and J. M. Coffin. 1988. Highly preferred targets for retrovirus integration. *Cell* **53**:531–537.
 40. Siebert, P. D., A. Chenchik, D. E. Kellogg, K. A. Lukyanov, and S. A. Lukyanov. 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* **23**:1087–1088.
 41. Stewart, M. A., M. Warnock, A. Wheeler, N. Wilkie, J. I. Mullins, D. E. Onions, and J. C. Neil. 1986. Nucleotide sequences of a feline leukemia virus subgroup A envelope gene and long terminal repeat and evidence for the recombinational origin of subgroup B viruses. *J. Virol.* **58**:825–834.
 42. Stoye, J. P., and J. M. Coffin. 1987. The four classes of endogenous murine leukemia virus: structural relationships and potential for recombination. *J. Virol.* **61**:2659–2669.
 43. Stoye, J. P., and J. M. Coffin. 1995. The dangers of xenotransplantation. *Nat. Med.* **1**:1100.
 44. Takeuchi, Y., S. H. Liang, P. D. Bieniasz, U. Jager, C. D. Porter, T. Friedman, M. O. McClure, and R. A. Weiss. 1997. Sensitization of rhabdo-, lenti-, and spumaviruses to human serum by galactosyl(α 1-3)galactosylation. *J. Virol.* **71**:6174–6178.
 45. Takeuchi, Y., C. Patience, S. Magre, R. A. Weiss, P. T. Banerjee, P. Le Tissier, and J. P. Stoye. 1998. Host range and interference studies of three classes of pig endogenous retrovirus. *J. Virol.* **72**:9986–9991.
 46. Tonjes, R. R., F. Czauderna, and R. Kurth. 1999. Genome-wide screening, cloning, chromosomal assignment, and expression of full-length human endogenous retrovirus type K. *J. Virol.* **73**:9187–9195.
 47. Towers, G., M. Bock, S. Martin, Y. Takeuchi, J. P. Stoye, and O. Danos. 2000. A conserved mechanism of retrovirus restriction in mammals. *Proc. Natl. Acad. Sci. USA* **97**:12295–12299.
 48. van der Laan, L. J., C. Locky, B. C. Griffith, F. S. Frasier, C. A. Wilson, D. E. Onions, B. J. Hering, Z. Long, E. Otto, B. E. Torbett, and D. R. Salomon. 2000. Infection by porcine endogenous retrovirus after islet xenotransplantation in SCID mice. *Nature* **407**:90–94.
 49. Weiss, R. A., S. Magre, and Y. Takeuchi. 2000. Infection hazards of xenotransplantation. *J. Infect.* **40**:21–25.
 50. Wilson, C. A., S. Wong, J. Muller, C. E. Davidson, T. M. Rose, and P. Burd. 1998. Type C retrovirus released from porcine primary peripheral blood mononuclear cells infects human cells. *J. Virol.* **72**:3082–3087.
 51. Wilson, C. A., S. Wong, M. Van Brocklin, and M. J. Federspiel. 2000. Extended analysis of the in vitro tropism of porcine endogenous retrovirus. *J. Virol.* **74**:49–56.