Acta virologica **66**: 110 – 126, 2022 doi: 10.4149/av_2022_201

# Selective human cells *in vitro* transduction with porcine endogenous retrovirus (PERV)

Sabina Gałka, Ewa Nowak, Ilona Bednarek

Department of Biotechnology and Genetic Engineering, Faculty of Pharmaceutical Sciences in Sosnowiec, Medical University of Silesia in Katowice, Poland, ul. Jedności 8, 41-200 Sosnowiec, Poland

*Received April 27, 2021; revised November 16, 2021; accepted December 12, 2021*

**Summary. –** Porcine xenograft transplantation raises concerns in humans about the risk of infection with porcine endogenous retroviruses (PERV) as they are an integral part of the pig genome and are therefore very difficult to exclude. In this study, for the first time, a relationship between the provirus genes sequences and released virions from pig cell line and the embedded sequence of this retrovirus in infected human cells was analyzed. PERV infection of human cells HEK-293 and HeLa and detection of PERV in pig PK-15 cells and supernatant were assessed by QPCR or RT-QPCR using primers specific for *envA*, *envB*, *gag*, *pol* genes and LTR region. Sequence analysis was performed at the DNA level and changes in the amino acid sequence were deduced *in silico.* Fifty nucleotide substitutions (45 in *pol*, 3 in *gag* and one each in *envA* and *envB*) were detected and most of these were heterozygous (42), which were present mainly in PK-15 cells. Our results show that sequence of the *pol* gene and the Pol protein is less conserved compared to the other *PERV* genes and PERV with some polymorphisms were not released from pig cells or/and do not infect human cells. PERV virions with a homozygous allele system were released from PK-15 cells, although their sequence replicated on the basis of the heterozygous PERV provirus sequence in PK-15. The newly discovered selective transduction of human cells with PERV will be helpful in studying the characteristics and genetic variability of the retrovirus genes to ensure safe xenotransplantation.

**Keywords:** PERV; porcine endogenous retroviruses; infection; genetic polymorphism; xenotransplantation

## Introduction

Xenotransplantation has become recognized as an alternative method of response to the increasing shortage of organs for allotransplantation. Due to physiological, anatomical, economical and ethical reasons, pigs constitute the most suitable organ and tissue donor species (Yoon *et al*., 2020). While the most common pathogens in pigs can be eliminated by breeding, treatment and containment, porcine endogenous retroviruses (PERV) cannot, as it could be integrated into the genome of all pig strains. PERV are very difficult or even impossible to eliminate due to a large number of integration sites in the animal's genome. In the pig genome, between 10 to 100 copies of the PERV provirus with a high degree of polymorphism can become embedded (Denner, 2016; Krüger *et al*., 2020a,b). However, recent studies have indicated that the pig genome contains fewer endogenous retroviruses compared to many other vertebrates, including humans and mice, and these sequences, for the most part, have been additionally characterized as damaged (Groenen *et al*., 2012). In pigs, PERV virions are actively produced in the kidney, skin, heart, liver, pancreas, spleen, thymus, lymph nodes, lungs, endothelial cells, and most importantly peripheral

E-mail: sgalka@sum.edu.pl; phone: +48-(32)-364 1257.

**Abbreviations:** env = envelope gene/protein; gag = group-specific antigen gene/protein, GAPDH = glyceraldehyde 3-phosphate dehydrogenase, HERV = human endogenous retroviruses; LTR = long terminal repeat; NTC = no template control; PBMC = peripheral blood mononuclear cells; PERV = porcine endogenous retroviruses; pol = polymerase gene/protein

blood mononuclear cells (PBMC). The number of released virions depends on the organ, individual, and even breed (Machnik *et al*., 2010; Mazurek *et al*., 2013, Semaan *et al.,* 2013; Sypniewski *et al*., 2005). Moreover, at least two subtypes of PERV, that is PERV-A and PERV-B, can infect human cells *in vitro*. Endogenous pig retroviruses in a xenograft can be transmitted to human cells, integrate into the genome, leading to changes causing various types of mutations, including point mutations, deletions, duplication, inversions and recombinations, especially when the cells are infected by two unrelated or related, exogenous or endogenous retroviruses (Denner, 2017; Kuddus *et al*., 2003). Such genomic rearrangements often result in inactive genomes of embedded retroviruses and changes in the regions of non-coding sequences. In the long terminal repeat (LTR), nucleotide sequences can dramatically affect retroviral replication and pathogenicity (Xie *et al*., 2019). The integration of a retrovirus can silence or activate nearby genes and trigger oncogenesis. In addition, it can also affect the activity of human endogenous retroviruses (HERV) or induce recombination in these viruses, leading to the formation of PERV/HERV viral hybrids (Denner, 2016; Łopata *et al*., 2018). In the genome of HEK-293 cells, 189 sites of PERV integration were observed. This integration was unevenly distributed throughout the whole set of chromosomes and most frequently occurred near transcription start sites and CpG islands, less often within genes. Additionally, it was shown that *PERV* integration was not completely accidental, since it most often took place within the human genome with the TG (int) GTACCAGC sequence, as was the case with the murine leukemia virus, suggesting that gammaretroviruses have a common integration pattern and that mechanisms for selecting target location within the retrovirus family may be similar (Mager and Stoye, 2014). Endogenous retroviruses are categorized as "not harmful" since most of them do not induce diseases in their natural hosts. PERV transmission from pig cells to human cells may result in the conversion of the endogenous nature of PERV in pigs into an exogenous state in humans. The loss of endogeneity and, at the same time, "not harmful" character may be associated with pathogenic potential occurring in closely linked, PERV-related exogenous gammaretroviruses, such as feline leukemia virus (FeLV), murine leukemia virus (MLV) and gibbon leukemia virus (GALV), which are capable of inducing tumors and immunodeficiency in the infected host (Mager and Stoye, 2014). Integration of the MLV vector near the promoter region contributed to the activation of the LMO2 proto-oncogene in patients and induced leukemia during gene therapy trials. The results showed that a similar risk of tumor induction may occur in the case of PERV integration that takes place during xenotransplantation (Moalic *et al*., 2006). The introduction of new antigens onto the cell surface constitutes another very important consequence of numerous viral infections. This is particularly important with regard to enveloped viruses, including retroviruses, which are released from the surface of the infected cell and induce the presentation of foreign intracellular antigens. Since these antigens are virus-specific and not host cell-like, they change the infected cell so that it becomes sensitive to cytotoxic T-lymphocyte attack, which is an important element of the immune response (Collier and Oxford, 2001).

PERV from pig PBMCs (coculture, inoculum) (Machnik *et al.,* 2004), different cell lines, miniature pig PERV-NIH (NIH, National Institutes of Health) (Ritzhaupt *et al*., 2002), Mv1Lu PERV, 293 (32) PERV (Specke *et al*., 2002), and most often the supernatants from PK-15 cells (porcine kidney epithelial cell line) (Kim *et al*., 2019; 2020, Scheef *et al*., 2002; Yu *et al*., 2005) were used to infect various human cell lines, including PBMCc, HEK-293 (Specke *et al*., 2002; Yu *et al*., 2005), HeLa (Ritzhaupt *et al*., 2002), C8166 C lymphocytes, THP-1 monocytes (human acute monocytic leukemia cell line), WIL2. NS.6TG (spleen lymphoblasts), HPAEC (human pulmonary artery endothelial cells), HAEC (human aortic endothelial cells), HCASMC (human coronary artery smooth muscle cells) (Specke *et al*., 2002). Infection, genome integration, replication and release of new PERV virions from human cells infected *in vitro* were confirmed (Machnik *et al*., 2004). In this study, for the first time with regard to the literature data, the relationship between the proviral gene sequences and virions released from the pig's PK-15 cell line, as well as the embedded sequence of this retrovirus in infected human cells HEK-293 and HeLa were analyzed. These findings will facilitate research on the characteristics and genetic variability of the retroviral genus (*envA, envB, gag* and *pol*) of the PERV provirus from PK-15 cells, in released virions and of the provirus in infected human cells: HEK-293 and HeLa. In order to perform safe xenotransplantation, it is necessary to select animals characterized by a low expression of PERV-A and PERV-B, and not harboring PERV-C (Semaan *et al*., 2013). Moreover, retroviral gene sequences should be taken into account, since retroviruses with polymorphisms were not released from pig cells or/and do not infect human cells. The obtained results may also help to improve the application of genetic engineering techniques based on the retroviral genomic sequence used to reduce PERV infection in xenograft recipients.

## Materials and Methods

*Infection-competent cell culture model*. Human embryonic kidney cell line HEK-293 (ATCC: CTL-1573) and HeLa cell line (ATCC: CCL-2) derived from human cervical cancer, were in-

**Table 1. Primers used for *PERV* and control pig and human genes for QPCR/RT-QPCR analysis**

| Genome | Gene/region | Primer[a] | Sequence (5'–3') | Product lenght (bp) | Location (bp) | Acc. No. |
|---|---|---|---|---|---|---|
| *PERV* | *gag* | *gag* F | TGATCTAGTG AGAGAGGCAGAG | 262 | 2230-2251 | AY099324 |
| | | *gag* R | CGCACACTGG TCCTTGTCG | | 2491-2473 | |
| | *envA* | *env* A-F | GAGATGGAAAGATTGGCAACAGCG | 364 | 6572-6595 | AY099323 |
| | | *env*A-R | AGTGATGTTAG GCTCAGTGGGGAC | | 6935-6912 | |
| | *envB* | *env* B-F | AATTCTCCTTTG TCAATTCCGGCCC | 270 | 6508-6532 | AY099324 |
| | | *env*B-R | CCAGTACTTTAT CGGGTCCCACTG | | 6777-6754 | |
| | *pol* | POL1F[b] | TACCCCTTGAGTAGAGAGGCTCGA | 600 | 3128-3151 | AY099324 |
| | | POL1R[b] | CCTGCAAATCTGGGCCTTCTTAGC | | 3727-3704 | |
| | | POL2F[b] | GAGCATTCAGTGCTGCTACAACCA | 588 | 2675-2698 | |
| | | POL2R[b] | TCATTGGTCCCAGGCTTCCTAACC | | 3261-3238 | |
| | | POL2RNA2F[b,c] | AGACTTACTGACCAAGATGGGAGCT | 407 | 2854-2878 | |
| | | POL2R[b] | TCATTGGTCCCAGGCTTCCTAACC | | 3261-3238 | |
| | | PRTF2 | GAGAGGTCAATAAACGGGTGCAGG | 379 | 3285-3308 | |
| | | PRTR2 | TGCCTTCGTGCCTTCCTAAGCAGTC | | 3664-3641 | |
| | | POL3F[b] | CGGGAGCCACCAAACAGGAC | 558 | 3624-3643 | |
| | | POL3R[b] | CTGCGATAGCCTTCAGACATACGG | | 4181-4159 | |
| | | POL4F[b] | GGAGACCTGTTGCCTACCTGTCA | 541 | 4104-4126 | |
| | | POL4R[b] | GCTTGCGTGAGGGCCATGA | | 4626-4644 | |
| | | POL5F[b] | CAGCCTGCCGGAAGGAACTT | 537 | 4588-4607 | |
| | | POL5R[b] | TGGGTGAGACGATGTATCTGTTGGA | | 5124-5100 | |
| | | POL6F[b] | GGGACCTGCTATACCTCATATGGGAAG | 561 | 5039-5065 | |
| | | POL6R[b] | CCTCTCTACCTGTCCTGAGCTTTGG | | 5599-5575 | |
| | | POL7F[b] | TGGGGATTGATTGGAAACTGCATTGTG | 661 | 5538-5564 | |
| | | POL7R[b] | GGGTTCGAGCTGTCTATAAGGCGTTTA | | 6199-6173 | |
| | | POL7F[b] | TGGGGATTGATTGGAAACTGCATTGTG | 323 | 5538-5564 | |
| | | POL7RNA2R[b,c] | CTTCCACGCTCGTTGCCTCAC | | 5860-5840 | |
| | LTR | PLTRF | ATTACAGACCCTGCTGGCTGCCAGT | 500 and/or 539 and/or every 539 | 6-30 LTR5', 8138-8163 LTR3' | AY099324 |
| | | PLTRR | TCACTCACAAGAAGCGGTCTTGAT | | 503-480 LTR5', 8677-8654 LTR3' | |
| Pig | mitochondrial *cytochrome B* | PMITF1 | CACCCGTTCATCATCATCGGCCA | 547 | 15196-15218 | NC_012095 |
| | | PMITR2 | GGTGGTGATATGCATGTTGACTG | | 15740-15717 | |
| | *GAPDH* | pGAPDH-F | TGTCGCCATCAATGACCCC | 116 (RNA) 195 (DNA) | 30-48 | exon 3 DQ648565.1 exon 4 |
| | | pGAPDH-R | TGACAAGCTTCCCATTCTC | | 224-206 | |
| Human | *β-actin* | AKT-F | TCACCCACACTGTGCCCATCTACGA | 295 | 560-584 | NM_001101 exon 4 |
| | | AKT-R | CAGCGGAACCGCTCATTGCCAATGG | | 830-854 | |

[a]F = forward primer, R = reverse primer, [b]*pol* gene primer walking, [c]primer specific for RNA from PK-15 supernatant.

fected by *PERV* spontaneously released from the porcine kidney cell line PK-15 (ATCC: CCL33). Infection assays were made according to Sypniewski *et al.* (2012). After the first passage, the DNA isolation from cells was performed.

*DNA and RNA isolation*. After 3 days post infection, DNA from the 1x10$^6$ HEK-293, HeLa and PK-15 cells was isolated using the lysis buffer with proteinase K and RNase A. The subsequent steps involved the method of salting out proteins

and DNA precipitation with isopropanol. RNA was isolated by phenol-chloroform method using TRI-Reagent according to the manufacturer's protocol, from 800 µl of supernatant used for infection, which was collected from PK-15 cells. To remove DNA from RNA extract, 1 µg of isolated RNA was hydrolyzed with 1 U of DNAse I.

*Detection of PERV genes.* PERV infection of human cell lines, HEK-293 and HeLa, and detection of PERV in PK-15 cells was evaluated by QPCR using primers specific for *PERV envA, envB, gag, pol* genes and LTR region (Table 1) (Sypniewski *et al*., 2012) and new additional designed primers used for overlapping primer-based PCR *pol* gene walking. Detection of viral RNA from supernatant of PK-15 cells was performed by RT-QPCR with the same primers. To avoid false-positive results induced by remaining porcine cellular DNA or RNA in the human's cells and in the supernatant from PK-15 cells, a control QPCR or RT-QPCR using the primers for pig cytochrome B mitochondrial DNA (mtDNA), and pig GAPDH (*pGAPDH,* to detect both DNA and RNA) was performed (Table 1). Amplification was carried out in 25 µl with SYBR Green QPCR or RT-QPCR Master mix (Stratagene, USA) using 0.2 µM primers, and 100 ng of DNA or RNA in standard conditions as stated by Sypniewski *et al*. (2012).

*Sequence analysis.* The obtained amplicons from *PERV* genes were separated on the 1% agarose gel and were isolated and purified by use of the Agarose Gel Extract mini kit (5 Prime/ Enzo Life Sciences, USA) according to the manufacturer's instructions. The sequences of the *PERV* DNA and RNA amplimers were obtained with commercial DNA sequencing services (using ABI Prism™ BigDye™ terminator cycle sequencing kit) at the Institute of Biochemistry and Biophysics of the Polish Academy of Sciences in Warsaw (Poland). Sequence analysis was performed using the Chromas 1.45 software. The sense and antisense sequences were compared among themselves and with the reference sequences AY099323 (for *envA*) and AY099324 (for *envB, gag, pol*) using the BLAST Clustal Omega software. Comparisons of the obtained gene sequences with all available sequences contained in the NCBI database were made using the MEGA BLAST software. At the protein level, changes in the amino acid sequence were deduced *in silico* and compared with the reference sequences AAM29191.1 (for *envA*), AAM29193.1 (for *envB*), and AAM29194.1 (for *gag, pol*).

## Results

*Real-time PCR detection of PERV genes*

The Real-Time™ PCR/RT-PCR method revealed differences in the presence of *envA, envB, gag, pol* and LTR coding sequences in the PK-15 cell supernatant and PK-15, HEK-293 and HeLa cells. According to previous data, PERV proviruses in PK-15 and HEK-293 cell lines, as well as virions released from PK-15 cells contain *envA, envB, gag,* and *pol* sequences. As expected, HeLa cells have become infected only by the PERV-A subtype, however, they did not contain an LTR region in their genome. Human kidney cells (HEK-293) were susceptible to infection with both PERV-A and PERV-B subtypes. Correct detection of human *β-actin* and the lack of amplification of pig *GAPDH* (neither DNA nor RNA), pig *mtDNA* and NTC reagent control in the PK-15 supernatant, HEK-293 and HeLa cells indicated no contamination of the supernatant and human cells with genetic material from pig cells (Fig. 1).

*Characterization of genomic variation*

According to reference sequences AY099323 (of the *envA* gene) and AY099324 (of the *envB, gag* and *pol* genes), 26 nucleotide substitutions (including most, that is 21, in *pol*, three in *gag*, one in *envA* and one in *envB*) were detected. Due to a large number of polymorphisms obtained in the amplimer of the *pol* gene, it was decided to analyze the complete sequence (3441 bp) of this gene in PK-15, HEK-293 cells and the supernatant from PK-15 cells. The primer walking strategy adopted in the sequence analysis of the *pol* gene revealed another 24 polymorphisms.

Out of all 50 substitutions, 18 were non-synonymous, 15 of which occurred in the *pol* gene, two in the *gag* gene,

**Table 2. Number of substitutions of the *PERV* genes in in the PK-15, HEK-293, HeLa cell lines and the supernatant of PK-15 cells**

| Cell/supernatant | *PERV* genes/polymorphism DANN (polymorphism protein) | | | |
|---|---|---|---|---|
| | ***envA*** | ***envB*** | ***gag*** | ***pol*** |
| PK-15 cell | 1HT (1)[ab] | 1HT[b] | 1HM, 1HT (1)[b] | 39HT (13), 5HM (1)[c] |
| PK-15 supernatant | - | - | 1HM[b] | 3HT, 8HM (3)[c] |
| HEK-293 cell | - | - | 1HM[b] | 3HT, 8HM (3)[c] |
| HeLa cell[b] | - | - | 3HM (2)[b] | 2HM[b] |

[a]The numbers in parentheses represent non-synonymous polymorphism – amino acid substitution refers to a [b]part or to the [c]complete sequence of a gene.

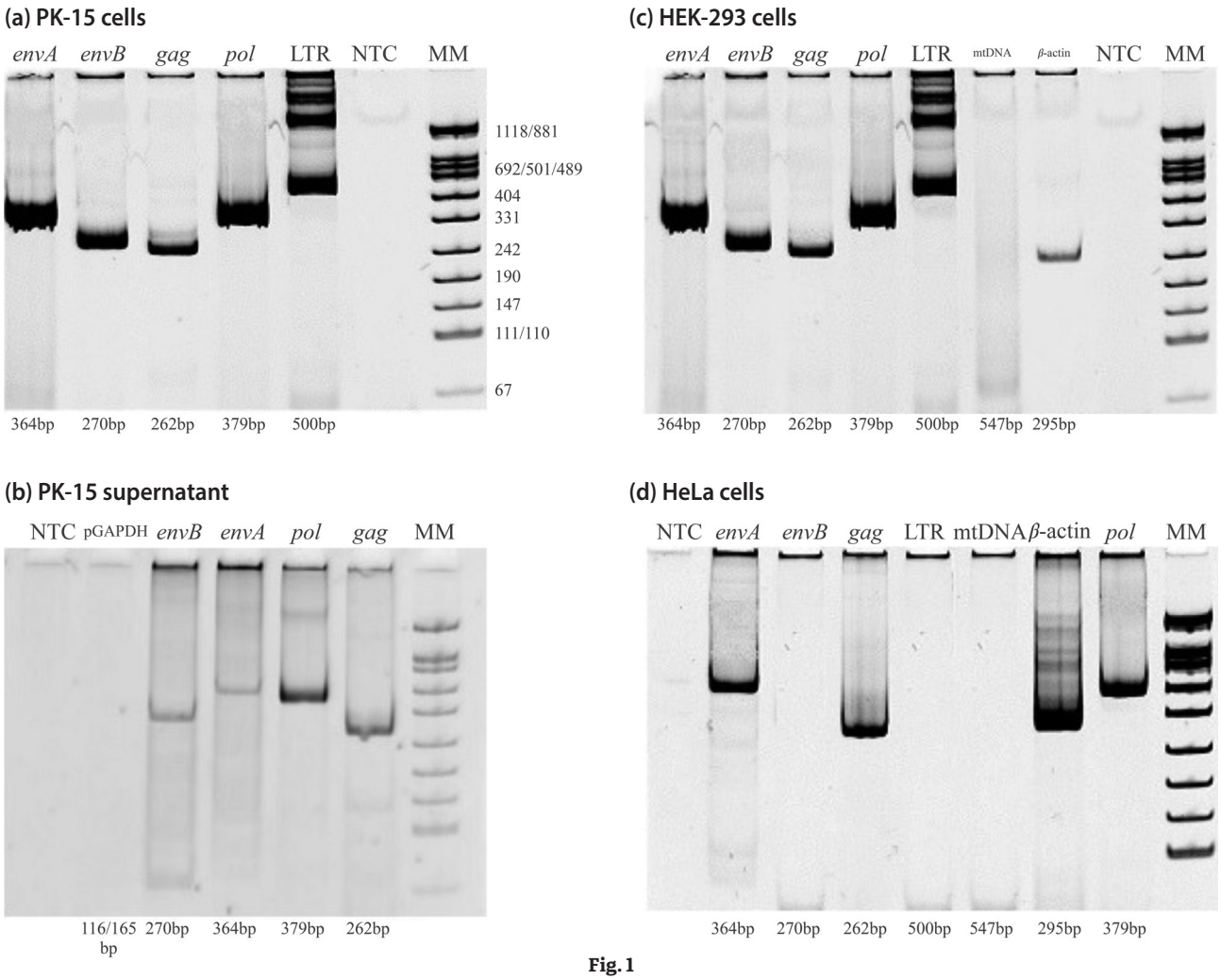


**Fig. 1**

**Polyacrylamide electrophoretic separation of PCR products of *PERV* genes**
Polyacrylamide electrophoretic separation of PCR products of *PERV* genes, LTR region, pig *pGAPDH,* pig *mtDNA*, human *β-actin* and NTC reagent control in PK-15 cells **(a)**, PK-15 cell supernatant **(b)**, HEK-293 **(c)** and HeLa cells **(d)**.

and one in the *envA* gene. Moreover, the majority of these polymorphisms were heterozygous (HT) (42 of 50), present primarily in PK-15 cells, and eight were homozygous (HM) (Table 2 and 3).

*PERV envA gene polymorphism*

One substitution, in the heterozygote form [6716 A>G]; [6716=], was identified in the *envA* gene, but it was only present in the PK-15 cell line. The allele [6716 A>G], in addition to the wild variant coding for isoleucine, could code for methionine at position 224 of the amino acid p. [Ile224Met]. This heterozygote was not detected in the *envA* gene of the virions from the supernatant taken from the aforementioned PK-15 cells and the provirus in the infected HEK-293 and HeLa cells, which revealed the presence of a wild-type homozygote [6716=] encoding isoleucine p. [Ile224=] (Fig. 2, Table 3). The observed polymorphic variant was located in the region of structural proteins (surface SU and transmembrane TM glycoproteins) corresponding to the location of the TLV coat region. Based on the comparison of obtained DNA sequence fragments of the *envA* gene with all 3,307 sequences of *PERV* nucleotide available in the NCBI/BLAST database (14. December 2020), their similarity to 65 sequences was demonstrated. The presence of adenine was detected in 14 cases (in infected HEK-293 and HeLa cells, and in PERV virions released from PK-15 cells) while guanine was found in 51 other cases. The heterozygote [6716 A>G]; [6716=] was identified only in the PK-15 cells.

**Table 3. Substitutions in the PERV genes in *PERV* provirus from pig PK-15 cells; PERV-virions released from pig PK-15 cells and in *PERV* provirus from infected human cells HEK-293 and HeLa**

| Analyzed gene/ GenBank reference gene's Acc. No. | Allel nucleotide/ amino acid sequence nt[a] / aa[b] | | PK-15 cell - changes within the provirus | PK-15 supernatant - changes within the released virions | HEK-293 cell - changes within the provirus | HeLa cell - changes within the provirus |
|---|---|---|---|---|---|---|
| *envA* AY099323 (DNA) [nt] AAM29191.1 (Protein) [aa] | 6716[a] | A | 6716A>G; 6716= | | 6716= | |
| | 224[b] | Ile | p.Ile224Met;Ile224= | | p.Ile224= | |
| *envB* AY099324 (DNA) [nt] AAM29193.1 (Protein) [aa] | 6740 | A | 6740A>G; 6740= | 6740= | | |
| | 232 | | | p.Thr232= | | |
| *gag* AY099324 (DNA) [nt] AAM29194.1 (Protein) [aa] | 2331 | G | | 2331= | | 2331G>A |
| | 441 | | | p.Arg441= | | p.Arg441His |
| | 2398 | A | | 2398A>G | | |
| | 463 | | | p.Arg463= | | |
| | 2400 | A | 2400A>G; 2400= | 2400= | | 2400A>G |
| | 464 | Glu | p.Glu464Gly; Glu464= | p.Glu464= | | p.Glu464Gly |
| *pol* AY099324 (DNA) [nt] AAM29194.1 (Protein) [aa] | 2916 | T | 2916T>C; 2916= | 2916= | | |
| | 636 | Val | p.Val636Ala; Val636= | p.Val636= | | |
| | 3091 | A | 3091A>G; 3091= | 3091A= | | |
| | 694 | Glu | p.Glu694= | 694 | | |
| | 3243 | G | | 3243G>A;3243= | | |
| | 745 | Arg | p.Arg745Lys; Arg745= | 745 | | |
| | 3299 | A | | 3299A>C | | |
| | 764 | | | p.Arg764= | | |
| | 3310 | C | | 3310C>T; 3310= | | |
| | 767 | | | p.Asp767= | | |
| | 3322 | G | | 3322G>A; 3322= | | |
| | 771 | | | p.Thr771= | | |
| | 3347 | A | 3347A>T; 3347= | | 3347= | |
| | 780 | Ser | p.Ser780Cys; Ser780= | | p.Ser780= | |
| | 3349 | C | 3349C>T; 3349= | | 3349= | |
| | 780 | Ser | p.Ser780Cys; Ser780= | | p.Ser780= | |
| | 3352 | C | 3352C>T; 3352= | | 3352= | |
| | 781 | | | p.Ala781= | | |
| | 3358 | G | 3358G>A; 3358= | | 3358= | |
| | 783 | | | p.Pro783= | | |
| | 3361 | T | 3361T>C; 3361= | | 3361= | |
| | 784 | | | p.Pro784= | | |
| | 3362 | G | 3362G>C; 3362= | | 3362= | |
| | 785 | Glu | p.Glu785Gln;Glu785= | | p.Glu785= | |
| | 3369 | A | 3369A>G; 3369= | | 3369= | |
| | 787 | Asn | p.Asn787Ser;Asn787 | | p.Asn787 | |
| | 3376 | C | 3376C>T; 3376= | | 3376= | |
| | 789 | | | p.Tyr789= | | |
| | 3394 | A | 3394A>G; 3394= | | 3394= | |
| | 795 | | | p.Lys795= | | |
| | 3416 | T | 3416T>C; 3415= | | 3415= | |

**Table 3. Continued 1**

| Analyzed gene/ GenBank reference gene's Acc. No. | Allel nucleotide/ amino acid sequence nt[a] / aa[b] | | PK-15 cell - changes within the provirus | PK-15 supernatant - changes within the released virions | HEK-293 cell - changes within the provirus | HeLa cell - changes within the provirus |
|---|---|---|---|---|---|---|
| | 803 | | | p.Leu803= | | |
| | 3421 | C | 3421C>T; 3421= | | 3421= | |
| | 804 | | | p.His804= | | |
| | 3467 | A | 3467A>G; 3467= | | 3467= | |
| | 820 | Thr | p.Thr820Ala;Thr820= | | p.Thr820= | |
| | 3490 | C | 3490C>T; 3490= | | 3490= | |
| | 827 | | | p.Thr827= | | |
| | 3496 | C | 3496C>T; 3496= | | 3496= | |
| | 829 | | | p.Thr829= | | |
| | 3517 | G | 3517G>A; 3517= | | 3517= | |
| | 836 | | | p.Lys836= | | |
| | 3553 | G | 3553G>A; 3553= | | 3553= | |
| | 848 | | | p.Arg848= | | |
| | 3583 | T | 3583T>C; 3583= | | 3583= | |
| | 858 | | | p.Pro858= | | |
| | 3655 | T | | 3655T>C | | |
| | 882 | | | p.Gly882= | | |
| | 3703 | T | 3703T>C; 3703= | | 3703= | |
| | 898 | | | p.Ser898= | | |
| | 4021 | C | 4021C>T, 4021= | | 4021= | |
| | 1004 | | | p.Asp1004= | | |
| | 4024 | A | 4024A>G; 4024= | | 4024= | |
| | 1005 | | | p.Val1005= | | |
| | 4082 | A | 4082A>G; 4082= | | 4082= | |
| | 1025 | Thr | p.Thr1025Ala; Thr= | | p.Thr= | |
| | 4090 | C | 4090C>T; 4090= | | 4090= | |
| | 1027 | Thr | p.Thr1025Ala; Thr= | | p.Thr= | |
| | 4348 | T | 4348T>G; 4348= | | 4348= | |
| | 1113 | | | p.Thr1113= | | |
| | 4735 | A | 4735A>G; 4735= | | 4735= | |
| | 1242 | | | p.Lys1242= | | |
| | 4869 | T | 4869T>C; 4869= | | 4869= | |
| | 1287 | Leu | p.Leu1287Pro; Leu= | | p.Leu= | |
| | 4876 | T | 4876T>C; 4876= | | 4876= | |
| | 1289 | | | p.Ser1289= | | |
| | 4923 | C | 4923C>G; 4923= | | 4923= | |
| | 1305 | Ala | p.Ala1305Gly; Ala1305= | | p. Ala1305= | |
| | 4942 | A | 4942A>G; 4942= | | 4942= | |
| | 1311 | Ile | p.Ile1311Met; Ile1311= | | p.Ile1311= | |
| | 4951 | G | | 4951G>A; 4951= | | |
| | 1314 | | | p.Thr1314= | | |
| | 4970 | A | 4970A>G; 4970= | | 4970A>G | |
| | 1321 | Arg | p. Arg1321Gly; Arg1321= | | p. Arg1321Gly | |
| | 5119 | C | | 5119C>A | | |

**Table 3. Continued 2**

| Analyzed gene/ GenBank reference gene's Acc. No. | Allel nucleotide/ amino acid sequence nt[a] / aa[b] | | PK-15 cell - changes within the provirus | PK-15 supernatant - changes within the released virions | HEK-293 cell - changes within the provirus | HeLa cell - changes within the provirus |
|---|---|---|---|---|---|---|
| | 1372 | His | | p.His1372Asn | | |
| | 5431 | A | | 5431A>G; 5431= | | |
| | 1474 | | | p.Lys1474= | | |
| | 5438 | C | 5438C>T; 5438= | | 5438= | |
| | 1477 | | | p.Leu1477= | | |
| | 5503 | T | 5503T>C;5503= | | 5503T>C | |
| | 1498 | | | p.Phe1498= | | |
| | 5455 | G | | 5455G>A | | |
| | 1482 | | | p.Pro1482= | | |
| | 5479 | C | | 5479C>A | | |
| | 1490 | | | p.Ile1490= | | |
| | 5665 | A | 5665= | | 5665A>G | |
| | 1552 | | p.Ile1552= | | p.Ile1552Met | |
| | 5695 | T | | 5695T>G; 5695= | | |
| | 1562 | | | p.Val1562= | | |

[a]nt - nucleotide, [b]aa - amino acid, p - protein, > - change or = - no change in sequence relation to the GeneBank reference sequences.

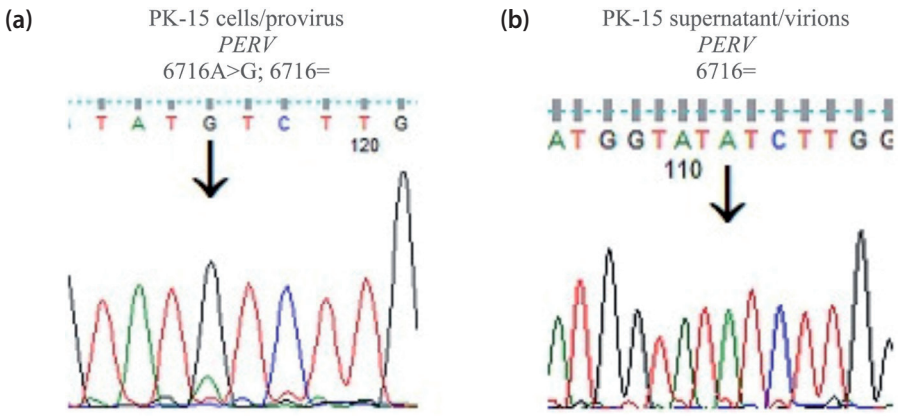


**Fig. 2**

***EnvA* gene sequencing**

*EnvA* gene sequencing chromatogram representing (a) heterozygote [6716 A>G]; [6716=] in the provirus PERV in PK-15 cells and (b) homozygote [6716=] in virions *PERV* released from PK-15 cells.

*PERV envB gene polymorphism*

In the *envB* gene sequence of the PERV provirus, a heterozygote [6740 A>G]; [6740=] was identified only in the PK-15 cell line, which did not change the 232$^{nd}$ amino acid coding as in the wild variant threonine [Thr232=] (Fig.3, Table 3). The obtained *envB* gene sequence was similar to 42 sequences (out of 3,307 *PERV* nucleotide sequences from the NCBI/BLAST database from 14. December 2020), 40 of which showed the presence of adenine (similarly to infected HEK-293 and HeLa cells, and PERV virions from the above-mentioned PK-15 cells), one revealed cytosine and one guanine.

*PERV gag gene polymorphism*

In the sequences of amplified *gag* gene fragments, three substitutions were identified at nucleotide positions

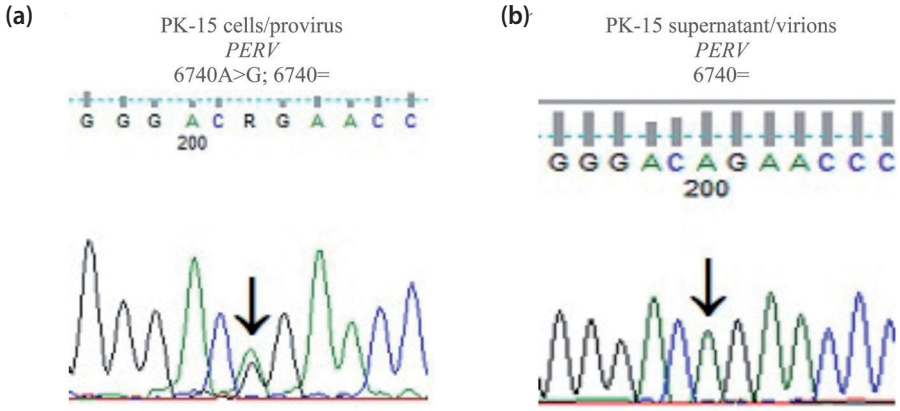


**Fig. 3**

***EnvB* gene sequencing**

*EnvB* gene sequencing chromatogram representing (a) heterozygote [6740 A>G]; [6740=] in the provirus PERV in PK-15 cells and (b) homozygote [6740=] in virions *PERV* released from PK-15 cells.

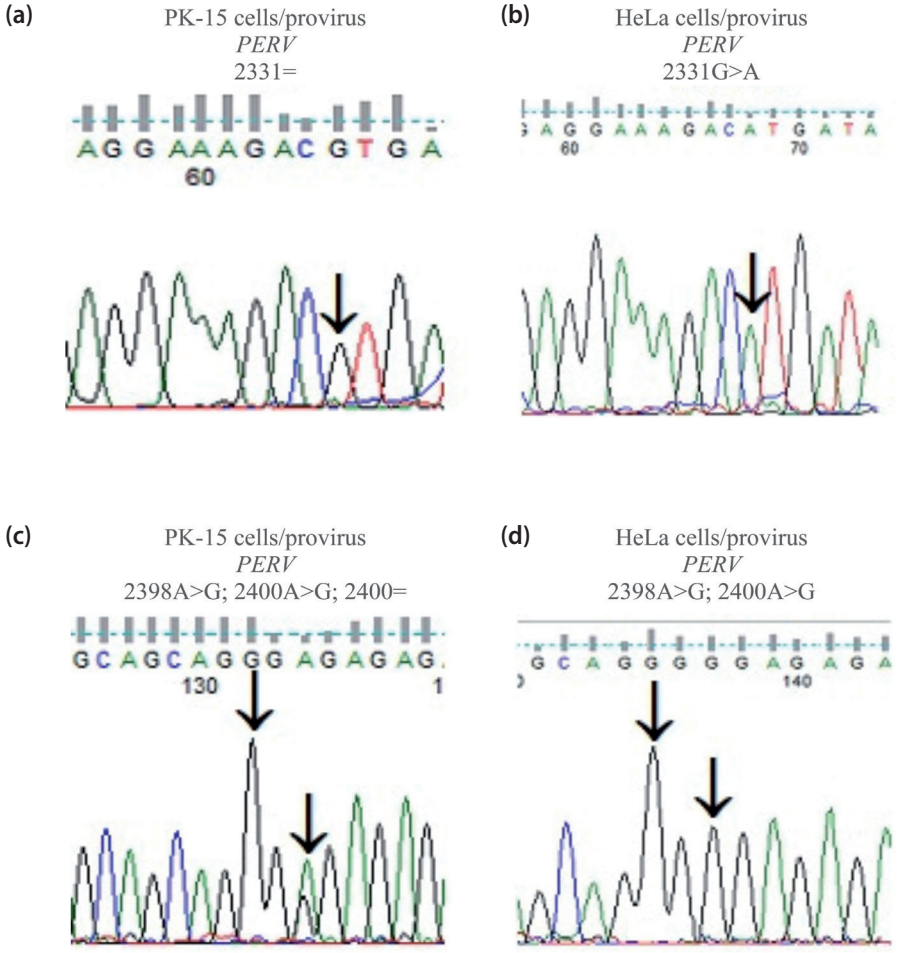


**Fig. 4**

***Gag* gene sequencing**

*Gag* gene sequencing chromatogram showing: (b) new polymorphic [2331G>A] and (d) homozygotes: [2398 A>G], [2400 A>G] in the provirus PERV in HeLa cells in comparison with (a) homozygotes: 2331=, (c) 2398 A>G and (c) heterozygote [2400 A>G]; [2400=] in the provirus *PERV* in PK-15 cells.

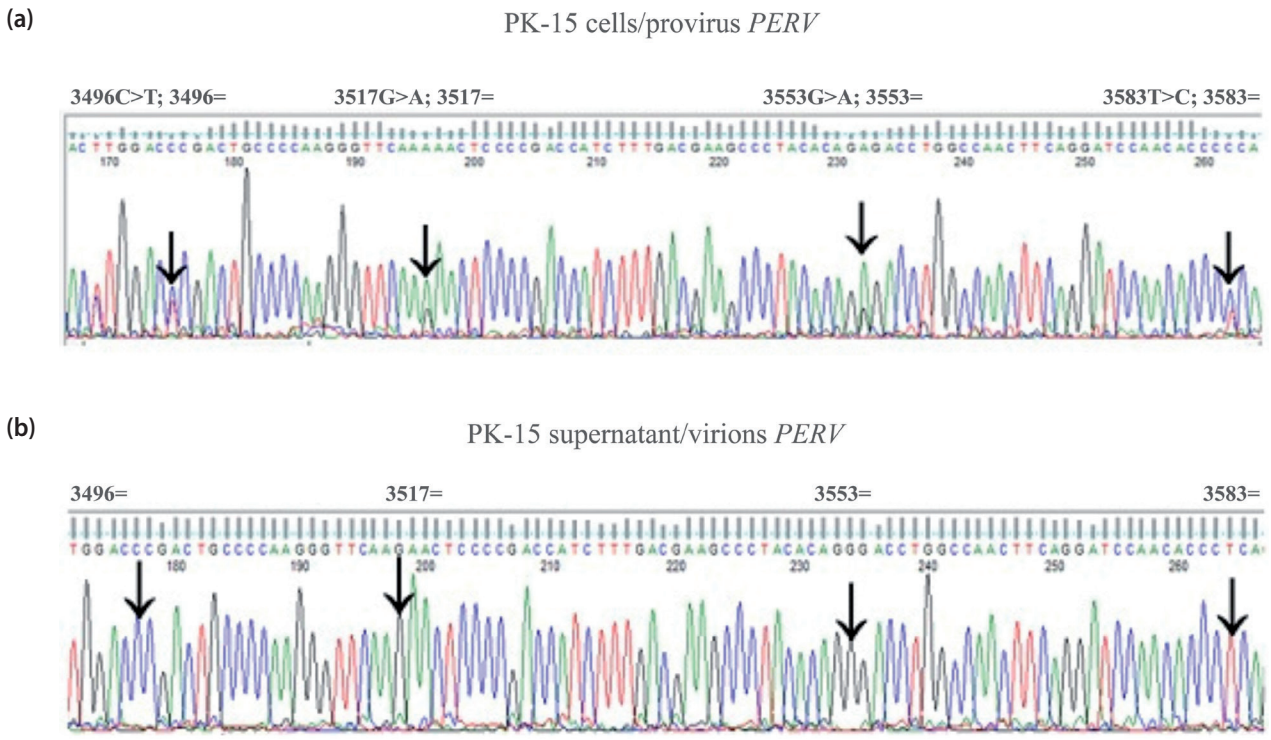


**Fig. 5**

***Pol* gene sequencing**

*Pol* gene sequencing chromatogram showing: (a) heterozygotes: [3496C>T]; [3496=], [3517G>A]; [3517=], [3553G>A]; [3553=], [3583 T>C]; [3583=] in the provirus PERV in PK-15 cells and (b) polymorphic variants: [3496=], [3517=], [3553=], [3583=] in virions PERV released from PK-15 cells.

2331, 2398 and 2400 compared to the reference sequence AY099324 (Fig. 4, Table. 3). The first [2331G>A] was present only in HeLa cells and could cause the coding of histidine at position 441 instead of arginine p. [Arg441His]. Another substitution in the *gag* gene sequence [2398 A>G] was identified in all tested samples however, it was a synonymous change p. [Arg463=] (Fig. 4, Table 3). A third polymorphic site in the *gag* gene sequence was identified as heterozygous [2400 A>G]; [2400=], and occurred only in PK-15 cells. The G allele was found in the HeLa cells. In the sequence of virions from the PK-15 cells and the provirus from HEK-293 cells, only the A allele was present. The transition [2400 A>G] caused the coding of glycine instead of glutamic acid p. [Glu464Gly) (Fig. 4, Table 3).

The above-mentioned polymorphic variants (*gag* gene) were located in the region of structural proteins (matrix protein MA, capsid protein CP and nucleocapsid protein NP) corresponding to the location of the region between Gag_p30 and ZnF_C2Hc.

In almost all cases, that is in the tested pig PK-15 cells, infected human HEK-293 cells, PERV virions in the supernatant of PK-15 cells and all 53 sequences (out of 3,307 *PERV* nucleotide sequences from the NCBI/BLAST database from 14. December 2020), the presence of guanine allele was found at the 2331$^{st}$ nucleotide. The only exception was the sequence of the *gag* gene in infected HeLa cells under study, characterized by a non-synonymous substitution [2331G>A]. At another polymorphic site, i.e. at the 2,398$^{th}$ nucleotide of the *gag* gene, guanine was found in pig PK-15 cells, infected human cells of HEK-293 and HeLa cell lines, PERV virions in the supernatant from PK-15 cells and 46 (out of 53) database sequences. Adenine was present in the remaining seven cases. In the next polymorphic site, which is the 2,400$^{th}$ nucleotide position of the *gag* gene, adenine was revealed in virions from PK-15 cells and infected HEK-293 cells, similarly to 51 (out of 53) sequences from the NCBI database. Guanine was present in infected HeLa cells at the same nucleotide position, as well as in two (out of 53) sequences from the NCBI database. The heterozygote [2400 A>G]; [2400=] was found only in the tested PK-15 cells.

*PERV pol gene polymorphism*

Forty-five nucleotide substitutions were identified by comparing the obtained complete sequence of the *PERV*

*pol* gene with reference sequence AY099324. All of these polymorphic sites (except for one, [5665 A>G], found only in the PK-15 supernatant and the HEK-293 cells) were present in the proviral sequence in tested PK-15 cells, and only two substitutions, [3299 A>C] and [3655 T>C], also occurred in the *pol* gene sequence in PERV virions from PK-15 cells, as well as proviruses in HEK-293 and HeLa cells. Nine remaining polymorphisms, i.e. [4951G>A; 4951=], [4970 A>G], [5119C>A], [5431 A>G; 5431=], [5503 T>C], [5455G>A], [5479C>A], [5665 A>G], [5695 T>G; 5695=], were also found in the PK-15 supernatant and HEK-293 cells. The substitutions identified in the proviral *pol* gene sequence in PK-15 cells were mainly (39 out of 45) heterozygotes. Five of them were homozygotes, three were transversions ([3299 A>C], [5119C>A], [5479C>A]) and two were transitions [3655 T>C], [5455G>A]) (Table 2 and 3, Fig. 5 and 6).

Among 39 heterozygous genotypes (except for the wild type allele), the second allele was most often identified as a transition: ten variants as C>T (3310, 3349, 3352, 3376, 3421, 3490, 3496, 4021, 4090, 5438$^{th}$ nt), in six cases as G>A (3243, 3322, 3358, 3517, 3553, 4951$^{th}$ nt), eight in the form of T> C (2916, 3361, 3416, 3583, 3703, 4869, 4876, 5503$^{th}$ nt), ten as A>G (3091, 3369, 3394, 3467, 4024, 4082, 4735, 4942, 4970, 5431$^{th}$ nt) and four types of transversions: one A>T (3347$^{th}$ nt), one G>C (3362$^{nd}$ nt), two T>G (4348, 5695$^{th}$ nt) and one C>G (4923$^{rd}$ nt) (Fig. 5 and 6, Table 3). By analyzing all 45 identified nucleotide polymorphic sites in the *pol* gene, it was found that most of them (30 out of 45) were synonymous changes while in 15 (out of 45) cases, observed primarily among heterozygotes in PK-15 cells (14 out of 15), were non-synonymous (Table 3). The above-mentioned changes in the amino acid sequence took place within the enzymatic proteins (corresponding to the location of integrase IN, protease PRO, and reverse transcriptase RT) encoded by the *pol* gene.

45 polymorphic sites observed in the tested samples were also found in sequences of the *pol* gene contained in the database, but in various numbers. For example, in all samples under study, cytosine was always detected at the 3,299$^{th}$ nucleotide, but in 56 sequences (out of 3,307 *PERV* nucleotide sequences from the NCBI/BLAST database from 4. December 2020) cytosine appeared in 30, adenine in 24, while guanine and thymine only in one case (Table 4).

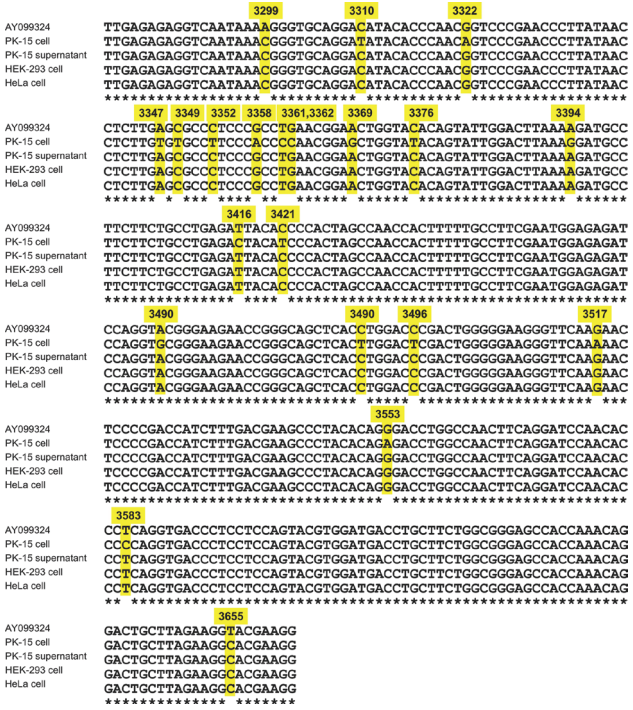


**Fig. 6**

**Multiple alignment of the part of *pol* gene sequence**

Multiple alignment of the part of *pol* gene sequence of the PERV provirus in porcine PK-15 cells, in infected human cells HEK-293 and HeLa, and PERV virions in the supernatant of the PK-15 cells with the reference sequence using the Clustal Omega software. The polymorphic sites along with their nucleotide locations are highlighted.

**Table 4. Comparison of selected *pol* alleles with sequences contained in the NCBI database (analysis based on MEGA BLAST software)**

| Cell/supernatant | | Nucleotide position | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 3299 | 3310 | 3322 | 3347 | 3349 | 3352 | 3358 | 3361 | 3362 | 3369 | 3376 | 3394 | 3416 | 3421 | 3467 | 3490 | 3496 | 3517 | 3553 | 3583 | 3655 |
| PK-15 cell | | A>C | C>T C | G>A G | A>T A | C>T C | C>T C | G>A G | T>C T | G>C G | A>G A | C>T C | A>G A | T>C T | C>T C | A>G A | C>T C | C>T C | G>A G | G>A G | T>C T | T>C |
| PK-15 supernatant | | A>C | C | G | A | C | C | G | T | G | A | C | A | T | C | A | C | C | G | G | T | C |
| HEK-293 cell | | A>C | C | G | A | C | C | G | T | G | A | C | A | T | C | A | C | C | G | G | T | C |
| HeLa cell | | A>C | C | G | A | C | C | G | T | G | A | C | A | T | C | A | C | C | G | G | T | C |
| NCBI database-number of cases | C | 30 | 26 | - | - | 27 | 25 | - | 30 | 29 | - | 27 | - | 14 | 42 | - | 42 | 45 | - | - | 17 | 15 |
| | A | 24 | - | 30 | 27 | - | - | 29 | - | 1 | 27 | - | 27 | - | - | 43 | - | - | 15 | 35 | - | - |
| | G | 1 | - | 26 | - | - | - | 27 | - | 26 | 29 | - | 29 | - | - | 13 | - | - | 41 | 21 | - | - |
| | T | 1 | 30 | - | 29 | 29 | 31 | - | 26 | - | - | 29 | - | 42 | 14 | - | 14 | 11 | - | - | 39 | 41 |

## Discussion

For the first time in terms of literature data, relationships between *PERV* proviral gene sequences (*envA*, *envB*, *gag*, *pol*) and released virions in pigs and the embedded sequence of this retrovirus in infected human cells were analyzed. Sequence analysis of *PERV* gene amplicons (*envA*– 364 bp, *envB*– 270 bp, *gag*– 262 bp and the complete *pol* gene sequence – 3441 bp long) was carried out on the level of proviral DNA obtained from pig PK-15 cells and from infected human HEK-293 and HeLa cells, as well as based on RNA sequences of virions found in the supernatant of *in vitro*-cultured PK-15 cells. The reference sequences, that is AY099323 (for the *envA* gene) and AY099324 (for the *envB, gag* and *pol* genes), were used for comparison, which resulted in the detection of experimental changes in the nucleotide at the DNA level and *in silico* changes at the protein level.

Compared to the reference sequences, most of the identified substitutions (45 cases, i.e. 90%) concerned the 3,441-nucleotide sequence of the complete *pol* gene. Three (out of 50) changes were found in the 262-nucleotide sequence of the *gag* gene fragment, one in the 364-nucleotide sequence of the *envA* gene fragment and one in the 270-nucleotide sequence of the *envB* gene fragment. On this basis, it can be concluded that most polymorphic changes per one nucleotide occurred in the case of the *pol* gene ($1.45x10^{-2}$), less in *gag* ($1.1x10^{-2}$) and the least in *envB* ($0.4x10^{-2}$) and *envA* ($0.3x10^{-2}$). Therefore, the highest coefficient of variation determined at the nucleotide sequence level was recorded for the *pol* gene (1.45%), lower for *gag* (1.1%) and the lowest for *envB* (0.4%) and *envA* (0.3%). Among these 50 variants, most of them (64%) were synonymous – silent changes that did not lead to any modifications in the amino acid sequence of the protein through a codon change. Among 18 remaining substitutions affecting the amino acid sequence, most (15) were identified in the *pol* gene, two in the *gag* gene and one in the *envA* gene. One polymorphic site located in the *envB* gene was synonymous. Thus, the coefficient of variation determined at the amino acid sequence level (including only non-synonymous polymorphisms) was about 1.3% for the complete sequence of the *pol* gene (1147 aa), 2.3% for the *gag* gene (fragment of 87 aa) and 0.8% for the *envA* gene (fragment of 121 aa). The above results were consistent with the literature data, in particular with analysis of the entire genome of PERV-A-BM isolated from the Chinese breed of pigs, Guangxi Bam (Tang *et al*., 2016). Within the 8,774-nucleotide-long PERV-A-BM, the authors revealed three substitutions in *gag* and *env* genes (each causing a change in the amino acid sequence). 91.3% (63 out of 69) of all detected substitutions were polymorphisms located in the *pol* gene (3,438 bp), which caused 24 changes in the amino acid sequence of the Pol (1145 aa) protein. Based on this finding, it was determined that the *pol* gene was also characterized by the highest coefficient of variation, both at the level of nucleotide sequence (of 1.83%) and the non-synonymous amino acid substitution (of 2.10%).

In all analyzed gene amplicons, i.e. *envA, envB, gag* and *pol*, a total of 50 nucleotide polymorphic changes were identified, most of which (44) occurred in PK-15 cells. It is worth noting that these cells were treated using the same conditions, i.e. isolation, amplification and sequencing, as all other cells. Moreover, nearly all of them (39 out of 44) were heterozygotes, which were observed mainly in PK-15 cells. That is why it could be concluded that PK-15 cells contain the most polymorphic *PERV* sequences compared to others, i.e. the *PERV* sequence of virions released into the supernatant from PK-15 and the *PERV* sequence of the provirus in HEK-293 and HeLa cells. Furthermore, the presence of a high total percentage of homozygous sequences was observed in the supernatant of these PK-15 cells, as well as infected HEK-293 and HeLa cells. Therefore, it is possible to assume that *PERV* virions with a homozygous allele system were released from PK-15 cells, although their sequences replicated on the basis of the heterozygous *PERV* proviral sequence in PK-15 cells. Perhaps in PK-15 cells, the translation and PERV transcription of one allele was inhibited, and there was a preferential formation and release of virions with the sequence of the second allele from a given locus. Additional studies analyzing sequences of PERV mRNA transcriptions in PK-15 cells could facilitate the verification of this issue in the future.

Six remaining substitutions (out of 44) were homozygous and identified in each tested sample. PERV virions infecting HEK-293 and HeLa cells resulted in homozygous sequences containing three nucleotide polymorphic changes: *gag:* [2398 A>G], *pol*: [3299 A>C], [3655 T>C], and tested only in HEK-293 *pol*: [5119C>A], [5455G>A], [5479C>A]. Apart from the three aforementioned cases, two other substitutions occurred in the sequence of the *gag* gene: [2331G>A] and [2400 A>G]. The former was present only in HeLa cells, and the latter took place also as one of the variants in the heterozygote in PK-15 cells but was not identified in the PERV virion sequence. On this basis, it can be assumed that the genome of HeLa cells infected with the same supernatant as HEK-293 cells was affected by two polymorphic changes in the region of the *gag* gene that arose *de novo* after infection, before or during PERV's integration into the HeLa genome. Both of these changes resulted in the encoding of other amino acids in the Gag protein p. [Arg441His] and p. [Glu464Gly]. It is possible to assume that the determining polymorphisms located in the *gag* gene may have altered the conformation and function of PERV Gag structural proteins. The

structural proteins consist of the capsid, nucleocapsid and matrix proteins. The matrix protein (MA, p15) is the outermost protein in the viral particle that binds to the viral envelope, i.e. the host cell lipid membrane, during the viral release. The capsid protein (CA, p27) is the major structural protein, while the nucleocapsid protein (NC, p10) is responsible for the efficient packing of RNA into the virion. In gammaretroviruses, the p12 protein is also distinguished. It is located in the Gag polyprotein, between MA and CA proteins, and participates in the integration and the process of releasing new virions (Rein, 2011; Lee *et al*., 2011). In the later phase, Gag polyproteins are bound to viral genomic RNA in the cytoplasm. The binding occurs through the NC portion of the Gag polyprotein and partially through the interaction with the MA. The resulting Gag multimers bind to the cell membrane, and the process of budding new virions begins. The p12 protein plays an important role in viral collection and release, namely its late (L) domain. Inactivation of this domain results in the incomplete maturation of released viruses, reduction of their shedding and accumulation of the Gag polyprotein under the cell membrane (Martin-Serrano and Neil, 2011). It is possible that based on the *PERV* proviral sequence with an altered *gag* gene sequence, altered Gag proteins that did not function properly were produced in HeLa cells, preventing the formation and release of new virions.

Despite the fact that the literature on the subject states that sequence multimerization in the LTR and recombination in the *env* gene are associated with the transferability of PERV to human cells and their potential pathogenicity (Denner, 2021), the authors did not identify the LTR region in HeLa cells. The study results are consistent with the information provided by Zwolinska (2006) and Kim *et al*. (2009) on the presence of retroviruses with fragment deletions or the absence of the entire LTR sequence. Granowitz and Goff (1994) investigated the effect of a series of substitution mutations on the MA protein of another retrovirus – Moloney murine leukemia in NIH/3T3 mouse fibroblast cells line and COS cells derived from monkey kidney fibroblasts. Many of the mutated viruses replicated similarly to wild-type ones, but the remaining seven, which were characterized by changes in the small region of the *gag* gene (amino-terminal residues 7–14), were replication-defective. These viruses were blocked from releasing particles from NIH/3T3 cells. It was also observed that the function of both the *gag* and *gag-pol* genes was defective. In contrast, in COS cells, three of the defective mutant DNAs were capable of forming virion particles, but only with an increased copy number of this gene in COS cells (Granowitz and Goff, 1994). The results of Granowitz and Goff's research suggest that mutations in this region inhibit the targeting of virions to the plasma membrane.

By analyzing all 45 identified polymorphic sites within the complete 3,441-bp *pol* gene sequence (compared to reference sequence AY099324, with a *gag-pol* polyprotein length of 5,162 bp), it was found that in 30 cases, these changes were synonymous while 13 (out of 45), observed only among heterozygotes in PK-15 cells, included modifications of nucleotides that potentially took place due to the codon change, which in turn led to a modification in the amino acid sequence within the enzyme proteins (respectively: integrase IN, protease PRO, reverse transcriptase RT) encoded by the *pol* gene. Perhaps these changes in the amino acid sequence influenced the functioning of the protease catalyzing the proteolysis of the Gag-Pol polyprotein into individual proteins. In addition, the observed polymorphisms could alter the action of retroviral reverse transcriptase, which is responsible for transcribing viral RNA into double-stranded DNA in the host cell's cytoplasm. This process takes place in the so-called reverse transcription complex (RTC), which includes genomic RNA, capsid, p12 protein, nucleocapsid, reverse transcriptase integrase and cellular proteins. Reverse transcription includes the process of flattening, and then, after synthesizing double-stranded DNA, the RTC complex transforms into PIC (the preintegration complex), which travels to the cell nucleus. In the nucleus, the viral DNA is connected with the host genome (creating a provirus) by the enzyme integrase, also encoded by the *pol* gene (Fassati, 2012).

This study presents a small number of polymorphic variants – only one, in *PERV envA* and *envB* gene sequences. In both cases, heterozygotes were observed in the *PERV* sequence, but found only in PK-15 cells. Moreover, the change in [6716 A>G] resulted in a coding change with regard to amino acid p. [Ile224Met] only in the case of the *envA* gene. The *env* gene encodes one *PERV* envelope protein, that is protease, into the transmembrane protein TM p15E and surface protein SU gp70. The SU protein allows the virus to contact the appropriate receptor on the surface of the host cell, while the transmembrane MT protein is responsible for the fusion of the viral envelope with the cellular membrane of the target cell. The Env protein is a glycoprotein with several glycosylation sites that may affect the process of binding to a host receptor. Watanabe *et al.* (2005) indicated that the entire SU region, not only the N terminal portion with the VRA and VRB regions, is necessary for the PERV-A binding to the host cell receptor to take place. It has also been shown that in terms of receptor binding, the PRR region is also necessary for PERV-A and PERV-C (Gemeniano *et al*., 2006). Additionally, in the case of PERV-C, 100 final amino acids in the SU protein were found to inhibit its capacity to bind to human cells. By using the SU protein without its C-terminal part (containing only 360 amino acids from

the N-terminus), PERV-C was modified, which allowed it to bind to human cells (HEK-293) through a receptor different than *PERV-A*. Therefore, it has been shown that this terminal portion of 100 amino acids, which differs only in terms of nine amino acids between PERV-C and PERV-A, plays a key role in modulating PERV-C binding capacity and infectivity (Gemeniano et al., 2006). Other researchers have proven that the replacement of one of the two amino acids with appropriate ones found in PERV-A, e.g. Q374R (glutamine→arginine) or I412V (isoleucine→valine), restores the infectivity of PERV-A/C in relation to human cells to a similar level as it was in the case of PERV-A (Argaw and Wilson, 2012). However, this modification did not cause PERV-C itself to become capable of infecting human cells. On the other hand, a change of only four amino acids of NHRQ436YNRP, together with K441R (asparagine↔tyrosine, histidine↔asparagine, glutamine↔proline and lysine↔arginine), resulted in PERV-C's affinity to human cells, similarly to that of PERV-A (Argaw and Wilson, 2012). In another study, Ito *et al.*, (2015) indicated that when a reverse mutation had been induced in the envelope (*env*) genes, ERV in cats did not generate infectious viral particles. ERV-DC7 fl and ERV-DC16 fl were highly expressed in cells but were not cleaved into surface (SU) and transmembrane subunits, nor incorporated into virions. The substitutions G407R / N427I-A429T and Y431D within the SU C-terminal domain of ERV-DC7 fl and ERV-DC16 fl, respectively resulted in these dysfunctions. Glycine 407 and tyrosine 431 residues are relatively conserved among infectious gammaretroviruses, and their substitution results in the same dysfunction as in the tested retroviruses. It reveals that specific mutations within the C-terminal domain of SU suppress Env cleavage and virion incorporation (Denner, 2016).

Retroviruses are susceptible to intra-genomic rearrangements (point mutations, deletions, or duplications) and recombinations, particularly when the cell is infected with two exogenous or endogenous retroviruses. Undamaged PERV is very active in terms of transcription and replication (Kuddus *et al.*, 2003). New changes generated in the viral genome constitute one method of adapting to the replication in a new environment or a different host. Knowledge of the frequency and speed of the formation of viral mutations is important to understand processes such as the evasion of the host's immune responses, drug resistance, and the ability of viruses to induce new diseases (Sanjuán and Domingo-Calap, 2016). In HIV-1 and other retroviruses, the mutation is primarily caused by the error-prone transcription, but there are two other possible sources of mismatching: host RNA polymerase II-mediated transcription and G to A transcription induced by restriction factor ABOBEC3G (Coffin and Swanstrom, 2013). Coffin *et al.* reported that this selection resulted from one variant being more efficient in terms of replication than the other, which in the case of viruses resulted in variations in the abundance of daughter particles in each replication cycle. These disproportions could arise from discrepancies occurring at any stage of the virus' life cycle, ranging from different virus adsorption capacity to the cell surface, disproportions concerning the entry of the virus into the cell, production rate of progeny virions, and the productive viability and stability of the infectious virus (Coffin and Swanstrom, 2013). Retroviruses mutate at an average rate of about $3x10^{-5}$ substitutions per nucleotide per cell infection (s / n / c), which is approximately 10 to 1,000 times more frequent than in the case of DNA viruses ($10^{-6}$–$10^{-8}$ s / n / c) and at least 10,000 times more common than in the case of eukaryotic DNA (≤$10^{-9}$ s / n / c). The presence of reverse transcriptase in retroviruses, or the lack of 3'–5' exonuclease activity, reduces the accuracy of DNA synthesis and causes more frequent errors ($10^{-4}$–$10^{-5}$ mutations/base pair). In addition, viruses are characterized by a negative correlation between genome size and mutation frequency (Rawson and Mansky, 2014; Sanjuán and Domingo-Calap, 2016; Schlub *et al.*, 2014). The polymorphisms observed in *PERV* gene sequences may cooperate with cellular restriction factors, e.g. APOBEC3, which catalysis the deamination of cytosine to uracil in both DNA and RNA that in turn causes instability of DNA strands, which, as a result, are unable to encode viral proteins and thus progeny viruses (Dörrschuck *et al.*, 2008). In this study, it was observed that polymorphic changes in the form of A>G transitions in the following gene sequences, *envA, envB, gag* and *pol* as well as T>C in *pol* occurred mostly in the *PERV* provirus in PK-15 cells, and less frequently in the supernatant of these cells. It is, therefore, possible that APOBEC3 proteins may have deaminated the cytosine to uracil in these genes, resulting in the fact that no progeny virions with this sequence or virions of incomplete RNA length were produced (Dörrschuck *et al.*, 2008). Similarly, Nowarski *et al.* (2014) proved that cytosine deamination in HIV-1 reduced the production of full-length HIV-1 transcriptions. This mechanism was based on the inactivation of the TAR sequence (trans-activation response element) by deamination of cytosine to uracil in the newly formed minus strand DNA. This change prompted the G to A mutation in the TAR RNA sequence, which in turn inhibited the function of the transcriptional activator protein Tat (transcriptional activator), blocking HIV-1 RNA extension of proviral transcripts, accumulating incomplete (short) viral transcripts and consequently limiting viral replication of HIV-1 (Nowarski *et al.*, 2014). Dörrschuck *et al.* (2008) revealed that PERVs were susceptible to the action of human, porcine and mouse APOBEC3 proteins. After introducing the molecular clone PERV-B, as

well as human A3F and A3G into 293 T cells, Lee *et al.* observed a reduction in *PERV* replication by 69–90%, which confirmed the inhibitory effect of APOBEC3 on PERV infectivity in human cells (Lee *et al.*, 2011). Porcine A3 proteins inhibited PERV through a mechanism dependent on cytosine deamination (Dörrschuck *et al.*, 2008), however, Jonsson *et al.* demonstrated that human A3G can affect *PERV* in a manner independent of deamination (Jónsson *et al.*, 2007). APOBEC3 can also reduce the infectivity of retroviruses by inhibiting primer hybridization, strand transfer during the transcription, accumulation of viral transcriptions, and proviral integration (Bélanger *et al.*, 2013; Nowarski *et al.*, 2014).

To reduce the risk of *PERV* infection in humans during xenotransplantation, it is necessary to test xenograft donor candidates for retrovirus levels, so that the organs can only be harvested from animals with low levels, while carriers of the PERV-C subtype should be completely eliminated (Fishman, 2019). Another suggested solution refers to using small interfering RNA (siRNA) and other genome editing techniques (ZFN, TALEN and CRISPR/Cas) to remove *PERV* coding sequences from the animal genome (Carithers, 2020; Hryhorowicz *et al.*, 2017, Lu *et al.*, 2020). In order for this strategy to be successful, it is essential to know the course of the infection itself and develop highly specific retrovirus detection techniques.

In conclusion, this study revealed that endogenous porcine PERV retroviruses are able to infect human cells *in vitro*, which can also occur in the case of xenotransplantation, therefore inducing negative health effects in the xenograft recipient. Observation of such cases is especially important due to the increased interest in xenogeneic and genetically modified stem cell-based therapy – the potential of porcine cells in the field of stem cell-based therapy and regenerative medicine is subject to intensive research (Galow *et al.*, 2020). In porcine cells, PERV was characterized by high sequence variability, mainly of the *pol* gene, which may result in less accurate verification of its presence in the organs of donors and less effective therapy using modern *PERV* gene interference techniques (CRISPR, RNAi, siRNA shRNA) based on the nucleotide sequence. Therefore, with regard to xenotransplantation, it is recommended to select donors with incomplete proviral and PERV virion sequences and specific polymorphic variants in the genome, possibly inhibiting the release of virions with such sequences. The results obtained in this publication may be useful for designing new method of PERV treatment.

**Acknowledgments.** We would like to thank Dr Daria Matczyńska and Dr Daniel Sypniewski for their help with the issues of *in vitro* infection and cell culture. The research was financed by the Medical University of Silesia in Katowice, Poland, under research grants: KNW-1-095/N/9/B, PCN-1-174/N/0/1 and PCN-2-002/N/1/I.

## References

Argaw T, Wilson CA (2012): Detailed mapping of determinants within the porcine endogenous retrovirus envelope surface unit identifies critical residues for human cell infection within the proline-rich region. J. Virol. 86(17), 9096–90104. https://doi.org/10.1128/JVI.00738-12

Bélanger K, Savoie M, Rosales Gerpe MC, Couture J-F, Langlois M-A (2013): Binding of RNA by APOBEC3G controls deamination-independent restriction of retroviruses. Nucleic Acids Res. 41, 7438–7452. https://doi.org/10.1093/nar/gkt527

Carithers RL Jr (2020): Liver transplantation: will xenotransplantation be the answer to the donor organ shortage?. Trans. Am. Clin. Climatol. Assoc. 131, 270–285.

Coffin J, Swanstrom R (2013): HIV Pathogenesis: Dynamics and Genetics of Viral Populations and Infected Cells. Cold Spring Harb. Perspect. Med. 3, a012526. https://doi.org/10.1101/cshperspect.a012526

Collier L, Oxford J (2001): Wirusologia. Wydawnictwo Lekarskie PZWL, Warszawa, pp. 63–64

Denner J (2016): How Active Are Porcine Endogenous Retroviruses (PERVs)? Viruses 8(8), 215. https://doi.org/10.3390/v8080215

Denner J (2017): The porcine virome and xenotransplantation. J. Virol. 14(1), 171. https://doi.org/10.1186/s12985-017-0836-z

Denner J (2021): The origin of porcine endogenous retroviruses (PERVs). Arch. Virol. 166, 1007–1013. https://doi.org/10.1007/s00705-020-04925-8

Dörrschuck E, Münk C, Tönjes RR (2008): APOBEC3 proteins and porcine endogenous retroviruses. Transplant. Proc. 40, 959–961. https://doi.org/10.1016/j.transproceed.2008.03.032

Fassati A (2012): Multiple roles of the capsid protein in the early steps of HIV-1 infection. Virus Res. 170, 15–24. https://doi.org/10.1016/j.virusres.2012.09.012

Fishman JA (2019): Infection in xenotransplantation: opportunities and challenges. Curr. Opin. Organ. Transplant. 24(5), 527–534. https://doi.org/10.1097/MOT.0000000000000682

Galow A-M, Goldammer T, Hoeflich A (2020): Xenogeneic and Stem Cell-Based Therapy for Cardiovascular Diseases: Genetic Engineering of Porcine Cells and Their Applications in Heart Regeneration. Int. J. Mol. Sci. 21, 9686. https://doi.org/10.3390/ijms21249686

Gemeniano M, Mpanju O, Salomon DR, Eiden MV, Wilson CA (2006): The infectivity and host range of the ecotropic porcine endogenous retrovirus, PERV-C, is modulated by residues in the C terminal region of its surface envelope protein. Virology 346, 108–17. https://doi.org/10.1016/j.virol.2005.10.021

Granowitz C, Goff SP (1994): Substitution mutations affecting a small region of the Moloney murine leukemia virus MA gag protein block assembly and release of virion particles. Virology 205(1), 336–344. https://doi.org/10.1006/viro.1994.1650

Groenen MAM, Archibald AL, Uenishi H et al. (2012): Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491(7424), 393–398.

Hryhorowicz M, Zeyland J, Słomski R, Lipiński D (2017): Genetically Modified Pigs as Organ Donors for Xenotransplantation. Mol. Biotechnol. 59, 435–444. https://doi.org/10.1007/s12033-017-0024-9

Ito J, Baba T, Kawasaki J, Nishigaki K (2015): Ancestral mutations acquired in Refrex-1, a restriction factor against feline retroviruses, during its cooption and domestication. J. Virol. 90(3), 1470–1485. https://doi.org/10.1128/JVI.01904-15

Jónsson SR, LaRue RS, Stenglein MD, Fahrenkrug SC, Andrésdóttir V, Harris RS (2007): The restriction of zoonotic PERV transmission by human APOBEC3G. PLoS One 2(9), e893. https://doi.org/10.1371/journal.pone.0000893

Kim JH, Choi E, Jung E-S, Kwon Y, Lee DS, Hwang DY, Hwang ES (2009): Characterization of clones of human cell line infected with porcine endogenous retrovirus (PERV) from porcine cell line, PK-15. Infect. Chemother. 41(1), 1–8. https://doi.org/10.3947/ic.2009.41.1.1

Krüger L, Kristiansen Y, Reuber E, Möller L, Laue M, Reimer Ch, Denner JA (2020): Comprehensive strategy for screening for xenotransplantation-relevant viruses in a second isolated population of Göttingen minipigs. Viruses 12(1), 38. https://doi.org/10.3390/v12010038

Krüger L, Stillfried M, Prinz C, Schröder V, Neubert LK, Denner J (2020): Copy number and prevalence of porcine endogenous retroviruses (PERVs) in German wild boars. Viruses 12(4), 419. https://doi.org/10.3390/v12040419

Kuddus RH, Ghandi CR, Rehman KK, Guo F, Watkins SC, Valdivia LA, Fung JJ (2003): Some morphological, growth, and genomic properties of human cells chronically infected with porcine endogenous retrovirus (PERV). Genome 46, 858–869. https://doi.org/10.1139/g03-064

Lee J, Cho YD, Heo YK, Kwon Y, Kim DG, Choi B-S, Kim SS, Kim TB (2011): Reduction of N-tropic mutant porcine endogenous retrovirus infectivity by human tripartite motif-containing 5-isoform alpha. Transplant. Proc. 43, 2774–2778. https://doi.org/10.1016/j.transproceed.2011.04.014

Lu T, Yang B, Wang R, Qin C (2020): Xenotransplantation: Current Status in Preclinical Research. Front. Immunol. 10, 3060. https://doi.org/10.3389/fimmu.2019.03060

Łopata K, Wojdas E, Nowak R, Łopata P, Mazurek U (2018): Porcine endogenous retrovirus (PERV) - molecular structure and replication strategy in the context of retroviral infection risk of human cells. Front. Microbiol. 9, 730. https://doi.org/10.3389/fmicb.2018.00730

Machnik G, Sypniewski D, Mazurek U, Rożek-Kostórkiewicz J, Wilczok T (2004): Endogenne wirusy świni (PERV) jako zagrożenie w ksenotransplantacjach. Med. Weter. 60(4), 345–348.

Machnik G, Sypniewski D, Gałka S, Loch T. Sołtysik D, Błaszczyk D, Bednarek I (2010): Zmiana poziomu ekspresji syncytyny I w komórkach linii HEK293 po infekcji endogennymi retrowirusami świni (PERV). Farm. Przegl. Nauk. 12, 14–20.

Mager DL, Stoye JP (2014): Mammalian endogenous retroviruses. Microbiol. Spectrum 3(1), 1–20. https://doi.org/10.1128/microbiolspec.MDNA3-0009-2014

Martin-Serrano J, Neil SJ (2011): Host factors involved in retroviral budding and release. Nat. Rev. Microbiol. 9(7), 519–531. https://doi.org/10.1038/nrmicro2596

Mazurek U, Kimsa MC, Strzalka-Mrozik B, Kimsa MW, Adamska J, Lipinski D, Zeyland J, Szalata M, Slomski R, Jura J, Smorag Z, Nowak R, Gola J (2013): Quantitative analysis of porcine endogenous retroviruses in different organs of transgenic pigs generated for xenotransplantation. Curr. Microbiol. 67(4), 505–514. https://doi.org/10.1007/s00284-013-0397-3

Moalic Y, Blanchard Y, Felix H, Jestin A (2006): Porcine endogenous retrovirus integration sites in the human genome: features in common with those of murine leukemia virus. J. Virol. 80(22), 10980–10988. https://doi.org/10.1128/JVI.00904-06

Nowarski R, Prabhu P, Kenig E, Smith Y, Britan-Rosich E, Kotler M (2014): APOBEC3G inhibits HIV-1 RNA elongation by inactivating the viral trans-activation response element. J. Mol. Biol. 426, 2840–2853. https://doi.org/10.1016/j.jmb.2014.05.012

Rawson JMO, Mansky LM (2014): Retroviral vectors for analysis of viral mutagenesis and recombination. Viruses 6(9), 3612–3642. https://doi.org/10.3390/v6093612

Rein A (2011): Murine leukemia viruses: objects and organisms. Adv. Virol. 403419. https://doi.org/10.1155/2011/403419

Ritzhaupt A, Van Der Laan LJ, Salomon DR, Wilson CA (2002): Porcine endogenous retrovirus infects but does not replicate in nonhuman primate primary cells and cell lines. J. Virol. 76(22), 11312–11320. https://doi.org/10.1128/JVI.76.22.11312-11320.2002

Sanjuán R, Domingo-Calap P (2016): Mechanisms of viral mutation. Cell. Mol. Life Sci. 73(23), 4433–4448. https://doi.org/10.1007/s00018-016-2299-6

Scheef G, Fischer N, Flory E, Schmitt I, Tönjes RR (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(24), 12553–12563. https://doi.org/10.1128/JVI.76.24.12553-12563.2002

Schlub TE, Grimm AJ, Smyth RP, Cromer D, Chopra A, Mallal S,Venturi V, Waugh C, Mak J, Davenport MP (2014): Fifteen to twenty percent of HIV substitution mutations are associated with recombination. J. Virol. 88(7), 3837–3849. https://doi.org/10.1128/JVI.03136-13

Semaan M, Rotem A, Barkai U, Bornstein S, Denner J (2013): Screening pigs for xenotransplantation: prevalence

and expression of porcine endogenous retroviruses in Göttingen minipigs. Xenotransplantation 20(3), 148–156. https://doi.org/10.1111/xen.12032

Specke V, Plesker R, Coulibaly C, Boller K, Denner J (2002): Productive infection of a mink cell line with porcine endogenous retroviruses (PERVs) but lack of transmission to minks in vivo. Arch. Virol. 147(2), 305–319. https://doi.org/10.1007/s705-002-8321-z

Sypniewski D, Machnik G, Mazurek U, Wilczok T, Smorąg Z, Jura J, Gajda B (2005): Distribution of porcine endogenous retroviruses (PERVs) DNA in organs of a domestic pig. Ann. Transplant. 10(2), 46–51

Sypniewski D, Bednarek I, Matczyńska D, Gałka S, Loch T, Sołtysik D, Machnik G, Nowak E (2012): Modulation of porcine endogenous retroviruses (PERVs) expression in vitro by shRNA in the presence of cyclosporine A and dexamethasone. Ann. Transplant. 17(4), 92–107. https://doi.org/10.12659/AOT.883699

Tang HB, Ouyang K, Rao GB, Ma L, Zhong H, Bai A, Qin S, Chen F, Lin J, Cao Y, Liao YJ, Zhang J, Wu J (2016): Characterization of complete genome sequences of a porcine endogenous retrovirus isolated from China Bama minipig reveals an evolutionary time earlier than that of isolates from European minipigs. Transplant. Proc. 48(1), 222–228. https://doi.org/10.1016/j.transproceed.2015.12.005

Watanabe R, Miyazawa T, Matsuura Y (2005): Cell-binding properties of the envelope proteins of porcine endogenous retroviruses. Microbes Infect. 7(4), 658–665. https://doi.org/10.1016/j.micinf.2005.01.008

Xie J, Ge W, Li N, Liu Q, Chen F, Yang X, Huang X, Ouyang Z, Zhang Q, Zhao Y, Liu Z, Gou S, Wu H, Lai Ch, Fan N, Jin Q, Shi H, Liang Y, Lan T, Quan L, Li X, Wang K, Lai L (2019): Efficient base editing for multiple genes and loci in pigs using base editors. Nat. Commun. 10(1), 2852. https://doi.org/10.1038/s41467-019-10421-8

Yoon CH, Choi HJ, Kim MK (2020): Corneal xenotransplantation: Where are we standing? Prog. Retin. Eye Res. 100876. https://doi.org/10.1016/j.preteyeres.2020.100876

Yu P, Zhang L, Li SF, Li YP, Cheng JQ, Lu YR, Bu H (2005): Long-term effects on HEK-293 cell line after co-culture with porcine endogenous retrovirus. Transplant. Proc. 37, 496–499. https://doi.org/10.1016/j.transproceed.2004.12.296

Zwolińska K (2006): Retroviruses-derived sequences in the human genome. Human endogenous retroviruses (HERVs). Postepy Hig. Med. Dosw. 60, 637–652.