Transcriptome analysis of Pseudorabies virus reveals selective regulation of different kinetic classes of genes

Ph.D. thesis booklet

Nándor Póka

University of Szeged
Faculty of Medicine
Department of Medical Biology

Supervisor: Prof. Zsolt Boldogkői

Szeged, 2017

Intorduction

The Pseudorabies virus (or by its taxonomic name: Suid herpes virus I – SuHV-I) is a member of the Alphaherpesvirinae subfamily within the Herpesviridae family. It was described by a hungarian veterinary Aladar Ajueszky in 1902 (Ajueszky, 1902). The virus is a pathogen of swine and can cause serious economical losses at pig farms. It is also a close relative of human herpesviruses, it can naturally spread through synaptically connected neurons that makes it a useful live neural tracer and can be used to track neural activity (Boldogkoi *et al.*, 2009). The lifecycle of the virus is governed by an expression cascade in which genes are under tight temporal control. The genes are classified into 4 (previously 3) classes based on their peak expression after infection (so called kinetic classes): immediate early – IE, early – E, early-late – E/L and ate – L. The aim of our research was to get a better picture of the viral transcriptome and to better understand viral gene expression.

Aims of research

1. How does the deletion of the us1 gene affect the global gene expression pattern?
2. Acquiring a more detailed map of the PRV transcriptome using Next-Generation Sequencing techniques
3. What is the effect of deletion of the gE/gI heterodimer complex on global gene expression?

Materials and methods

We used the Kaplan strain of PRV. PK-15 porcine kidney epithelial cells were used as host cells for viral growth. Cells were grown on DMEM supplied with 5% FBS and 80ug/ml gentamicin. Incubation was carried out at 37 ^{o}C with 5% CO_2.

The mutant virus strains were produced with the common tools of molecular biology (digestion with restriction endonucleases, molecular cloning, cotransfection).

To assess the changes in gene expression we harvested cells that were infected with mutant viruses at the following time points: 0.5 (only at US1 experiment), 1, 2, 4, 6, 8, 12, 18 and 24 h after infection. We used 3 biological parallels in every

experiment.

Gene expression was measured with reverse transcription coupled real-time PCR. RNA was extracted using NucleoSpin RNA II Kit according to manufacturer's instructions. Reverse transcription was carried out using SuperScript III enyme, while for the real-time PCR Absolute QPCR SYBR Green Mix and Corbett Roto-Gene 6000 machine were used.

For our sequencing experiment we created two different cDNA libraries: one random hexamer primed total RNA library for 100 bp pair-end sequencing, and one with custom oligoT10(VN) anchored primers for polyA sequencing. Sequencing was carried out by an external partner on Illumina HiScanSQ platform. Data analysis was carried out with software commonly used in bioinformatics (FastQC, Tophat, HOMER, Bowtie 2, IGV v2.2), and the most current reference genomes. For the evaluation of gene expression data, we used the so called R-value (and its variants) that shows the ratio relative gene expression.

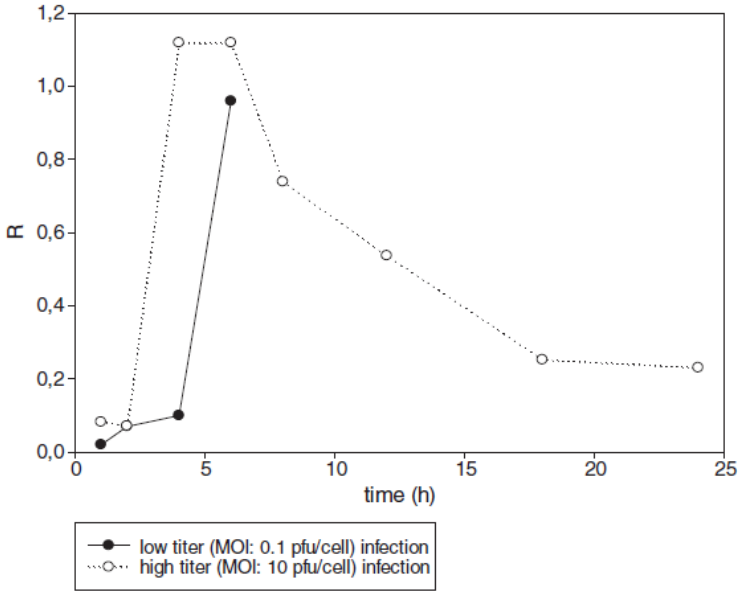

Figure 1. The expression kinetics of us1 at low (0,1 pfu/cell) and high (10 pfu/cell) titer infection in the wild type virus.. In low titer infections after 6-8 hours secondary infection occurs by the newly formed virus particles, that severely distorts gene expression data, thus for low titer infections data can be used only for the first 6 hours of infection. We have not seen similar titer dependent kinetics for other genes.

Results

- Comparison of infections with us1 mutant in low and high titer (0,1 and 10 pfu/cell respectively, pfu = plaque forming unit) revealed that onset and peaking of gene expression is delayed by 2 hours in low titer infections for the us1 gene. (Figure 1.)

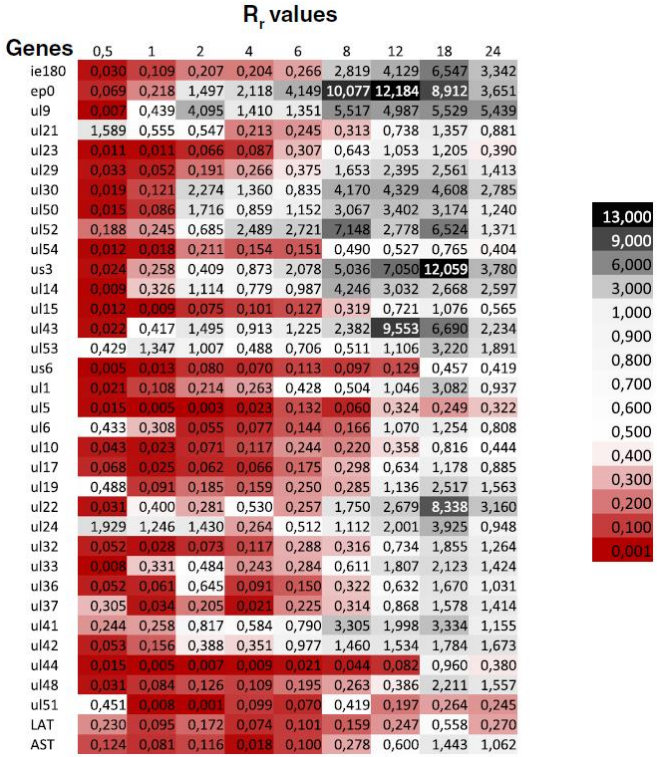

Figure 2. The heat map visualization of the ratio of gene expression (mutant/wild type) for the selected genes in the us1 experiment. Shades of red mean that at the given time point the expression of the gene in the mutant was lower than in the wild type, white means equal amounts, while shades of grey represent that the gene was overexpressed in the mutant compared to the wild type. It is clear that the genes of the E kinetic class are overexpressed compared to the wild type after 6 hours of infection.

- In the us1 mutant the average expression of genes belonging in the IE, E and E/L classes surpasses the same values of the wild type, in contrast to the genes of the L

kinetic class, whose average relative expression is below that of the wild type genes' expression, except for the 18 h sample. (Figure 2.).

- In the us7/us8 mutant the average expression of all kinetic classes are lower than that of the wild type in the early phase of the infection, however the inverse is true as the infection progresses. Ep0 produces the greatest overexpression (almost 24-fold) (Figure 3.).
- The deletion of us7/us8 gene pair affects the genes in the L kinetic class the most in the first hours of the infection, while later the genes of the IE and E classes are the most affected.
- In the us7/us8 mutant, the dynamics of the gene classes also differ from the wild type, except for the genes of the L kinetic class. (Figure 3.).
- In the us7/us8 mutant the expression of the ie180 gene strongly correlates with the expression of the other genesm similarly to what was seen previously. r values for Pearson's correlation: r=0,330 for wt, and r=0,893 for the mutant.
- We showed that all of the viral genes are transcribed. All intergenic regions were inactive, except for the ones between genes ul44-ul26 and the genes ul35-ul36 that showed leaky transcription. (Figure 4.).

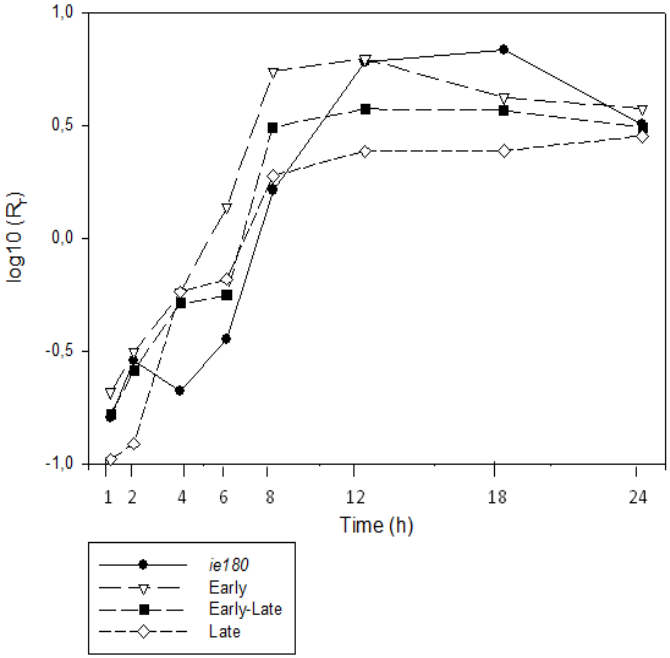

Figure 3. The $\log_{10}$ of the R_r values of the gene classes of the us7/us8 mutant virus $R_r = R_{mutant} / R_{wt}$. An R_r value greater than 1 indicates overexpression while if it's less than one it means underexpression compared to the wild type.

- We discovered a new non-coding RNA, termed CTO.
- We were able to show the presence of the predicted ORF1.2 sequence, but the exact transcription start could not be identified.

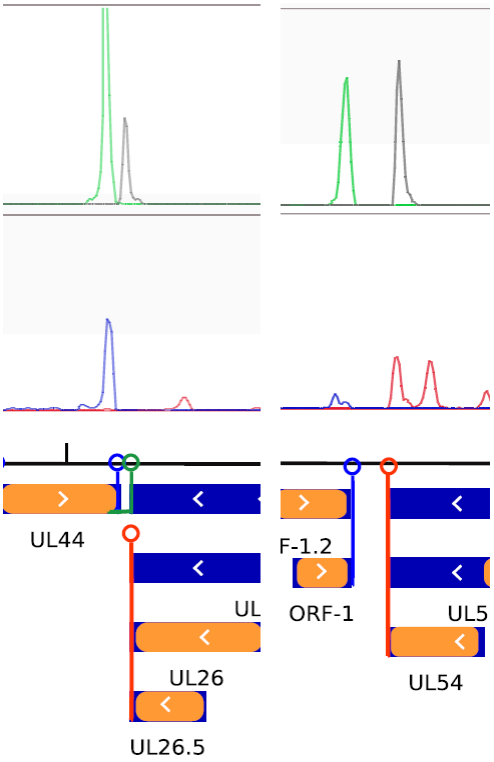

Figure 4. An example of leaky transcription. Legend: orange: coding regions, blue: transcripts, blue circle: polyA seq. on the + strand, red circle: polyA seq. on the – strand, green circle: alternative polyA seq. on the a + strand. Middle and top boxes, coverage/bp, top: PA-sequencing, green: + strand's coverage, black: − strand's coverage, middle: total RNA sequencing, blue: + strand's coverage, red: − strand's coverage.

Conclusions

The titer dependent expression of the us1 gene was not published previously. In closely related viruses it was shown that multiple RNA forms can be transcribed from the us1 gene (Holden *et al.*, 1992), however sequencing did not show new RNA species from this region, thus the reason of this phenomenon is still unclear. Our results suggest that the ICP22 protein that's encoded by us1 excites gene expression in the early phase of the infection. The excitatory effect can be carried either by directly influencing viral promoters or blocking host genes that would digest viral RNAs. In HSV-1 it was shown that ICP22 can interact with cyclin-dependent kinases and through this can block expression of host genes. Both mechanisms would require the protein to be readily present at early times of infection and this is in agreement with fact that the protein can be found in the tegument of the virus (Kramer *et al.*, 2011), thus it enters infected cell with the nucleo-capsid of the virus.

The role of the gE/gI heterodimer in gene expression regulation was not well known previously. The considerable overexpression of the ep0 (one of the virus' transcription factor that may also play role in synchronization of the gene expression(Tombácz, Tóth and Boldogkői, 2012)) in the mutant

virus, and the fact that expression of the gene ie180 (the primer trans-activator of the PRV) highly correlates with the expression of other genes, may explain, that even though the level of expression of the gene classes are different, their dynamics during the infection is still similar. We hypothesize that the gE/gI heterodimer has a yet unknown influence on the transcription of ep0, which in turn affects the activity of ie180 and/or other genes. Since it is known that the DNA replication of us7/us8 is intact, and our results show no difference in this matter, thus we assume that changes in DNA replication is not the reason behind our findings in the us7/us8 mutant.

The function of the novel non-coding RNA is not yet known, but its abundance is our samples is remarkable. Further experiments are required to assess the function and exact kinetics. Finding leaky expression between two pairs of genes can also be an important evidence for supporting the Transcriptional Interference Network hypothesis. The sequencing also showed that our strain is free of laboratory induced mutations.

Summary

Results of our research can be summarized as follows:

- The expression of the us1 gene depends on the titer of the infection
- Deletion of the us1 gene has different effect on the genes belonging to different kinetic classes
- In the late phase of infection every kinetic class had greater average expression in the us7/us8 mutant than in the wt
- The deletion of the us7/us8 gene pair affects kinetic classes differently
- The expression of the ie180 gene correlates with the expression of other genes in the us7/us8 mutants
- Brief summary of the high-resolution profile of the PRV genome (detailed description of these results is part of a another Ph.D thesis)

References

Ajueszky, A. (1902) "A contagious disease, not readily distinguishable from rabies, with unknown origin," *Veterinarius*, 12, pp. 387–396.

Boldogkoi, Z. *et al.* (2009) "Genetically timed, activity-sensor and rainbow transsynaptic viral tools.," *Nature methods*,

6(2), pp. 127–30.

Holden, V. R. *et al.* (1992) "ICP22 homolog of equine herpesvirus 1: expression from early and late promoters.," *Journal of virology*. American Society for Microbiology (ASM), 66(2), pp. 664–73.

Kramer, T. *et al.* (2011) "Proteomic Characterization of Pseudorabies Virus Extracellular Virions," *Journal of Virology*, 85(13), pp. 6427–6441.

Tombácz, D., Tóth, J. S. and Boldogkői, Z. (2012) "Effects of deletion of the early protein 0 gene of pseudorabies virus on the overall viral gene expression," *Gene*, 493(2), pp. 235–242.

Publications used for this thesis

I. Takács, I.F., Tombácz, D., Berta, B., Prazsák, I., **Póka, N.**, and Boldogkői, Zs., 2013. The ICP22 protein selectively modifies the transcription of different kinetic classes of pseudorabies virus genes. BMC molecular biology, 14, p.2. IF: 2,057

II. Oláh, P., Tombácz, D., **Póka, N.**, Csabai, Zs., Prazsák, I., and Boldogkői, Zs., 2015. Characterization of pseudorabies

virus transcriptome by Illumina sequencing. BMC Microbiology, 15(1), p.130. IF: 2,96

III. **Póka, N.**, Csabai, Zs., Pásti, E., Tombácz, D., and Boldogkői, Zs., 2017. Deletion of the us7 and us8 genes of pseudorabies virus exerts a differential effect on the expression of early and late viral genes. Virus Genes, pp.1–10. IF: 1,431

Acknowledgements

I'd like to show my gratitude to Zsolt Boldogkői for his supervision and his help in me getting this far. Results shown here are all products of teamwork for which I am grateful to, Dóra Tombácz, Zsolt Csabai, Péter Oláh, and Irma Takács. Special thanks to Ernő Duda, János Szabad, Kata Filkor and Csilla Magyarné Papdi for all the talks and their help as a friend. I would also like to thank the members of the Boldogkői group: István Belecz, István Prazsák, Mariann Ábrahám, György Seprényi, Gabriella Teleki and Antal Mai the good times we spent together.

This research were funded by the following grants:

TÁMOP-4.2.1.B- 09/1//KONV, TÁMOP-4.2.2/B-10/1-2010-0012, Swiss-Hungarian Cooperation Programme grant SH/7/2/8 and the Bolyai János Scholarship of the Hungarian Academy of Sciences: 2015-18.