

The intron-containing genome of the lytic *Pseudomonas* phage LUZ24 resembles the temperate phage PaP3

Non Peer-reviewed author version

Ceyssens, PJ; Hertveldt, K; Ackermann, Hans-W.; NOBEN, Jean-Paul; Demeke, M; Volckaert, G & Lavigne, R (2008) The intron-containing genome of the lytic *Pseudomonas* phage LUZ24 resembles the temperate phage PaP3. In: *VIROLOGY*, 377(2). p. 233-238.

DOI: 10.1016/j.virol.2008.04.038

Handle: <http://hdl.handle.net/1942/8377>

RUNNING TITLE: Genome analysis of *Pseudomonas aeruginosa* phage LUZ24

RAPID COMMUNICATION

The intron-containing genome of the lytic *Pseudomonas* phage LUZ24 resembles the temperate phage PaP3.

Pieter-Jan Ceysens¹, Kirsten Hertveldt¹, Hans-W. Ackermann², Jean-Paul Noben³,
Mekonnen Demeke¹, Guido Volckaert¹ and Rob Lavigne^{1*}

¹ Division of Gene Technology, Katholieke Universiteit Leuven,

Kasteelpark Arenberg 21, Leuven, B-3001, Belgium

² Felix d'Herelle Reference center for Bacterial Viruses, Department of Medical Biology,
Faculty of Medicine, Laval University, Québec, QC G1K 7P4, Canada

³ Hasselt University, Biomedical Research Institute and transnational
University Limburg, School of Life Sciences, Diepenbeek, Belgium

*Corresponding author. Division of Gene Technology, Katholieke Universiteit Leuven,
Kasteelpark Arenberg 21, Leuven, B-3001, Belgium

Fax: +32 (0)16 32 19 65; E-mail: rob.lavigne@biw.kuleuven.be

ABSTRACT

The virulent *Pseudomonas aeruginosa* bacteriophage LUZ24 (45,625 bp) was isolated from hospital sewage. It belongs to the family of the *Podoviridae*, and carries a bidirectionally transcribed dsDNA genome delineated by two direct terminal repeats of 184 bp. *In vitro* transcriptional analysis identified seven σ^{70} promoters, revealing a bias towards stronger promoter strength in the late genomic region. Reverse transcription demonstrated *in vivo* splicing of a 668 bp Group I intron embedded inside the DNA polymerase gene. Using mass spectrometry, nine structural proteins were identified as part of the phage particle. The lytic characteristics of LUZ24 are evaluated against its genomic content, which displays an overall 71% sequence similarity to the temperate phage PaP3.

INTRODUCTION

In current databases, five types of *Pseudomonas*-infecting *Podoviridae* can be distinguished based on genome organization and behaviour during infection. While gh-1 is a *stricto sensu* T7-like lytic phage (Kovalyova and Kropinski, 2003), the ϕ KMV-like phages form a genus of widespread and highly virulent phages with a genome architecture which only roughly resembles T7 (Lavigne *et al.*, 2003). A third group includes uncharacterized temperate phages Pap2 ([NC 005884](#)) and 119X ([NC 007807](#)), which display 93% nucleotide sequence identity (Kwan *et al.*, 2006). As a fourth type, the temperate and pilus-specific phage F116 (65,195 bp) replicates extrachromosomally as a plasmid during its lysogenic cycle and is considered as a genomic orphan (Byrne and Kropinski, 2005). Finally, phage PaP3 ([NC 004466](#)) is also described as a temperate phage and unrelated to any known phage. The PaP3 genome has 5'-protruding cohesive ends, and is able to integrate into the $tRNA^{lys}$ gene of its host (Tan *et al.*, 2007). We present here a detailed analysis of the particle structure, infectivity and genomic content of *P. aeruginosa* phage LUZ24, which was isolated as a virulent phage, but proved to be closely related to PaP3.

RESULTS

Phage characteristics. Phage LUZ24 was isolated from hospital sewage using the clinical *P. aeruginosa* strain Li010 (Pirnay *et al.*, 2002) as enrichment host. Phage particle morphology clearly classifies LUZ24 within the *Podoviridae* family of short-tailed phages, consisting of an icosahedral head with a 63 nm diameter and a short and banded tail of 12 x 8 nm. Some decoration proteins can be distinguished, protruding from the phage capsid (Figure 1A). Standard stability assays show that LUZ24 retains its infectivity from pH 4 up to pH 11, but titer levels drop 80% after 9 months of storage at 4°C in standard phage buffer (pH 7.5). When LUZ24 is incubated with *P. aeruginosa* strain Li010, large (5 mm) and clear plaques arise, and spontaneous resistant *P. aeruginosa* colonies only appear after 48 hours of incubation. One-step growth experiments show an efficient adsorption (81% in 5 min) and a

short latent period (approx. 40 minutes), which leads to an average release of 64 phage particles for each infected cell (Supplementary figure S1).

Extensive host range screenings on environmental and clinical *P. aeruginosa* strains (Pirnay *et al.*, 2002) show that LUZ24 lyses 36 out of 123 (29%) strains, including multi-drug resistant *P. aeruginosa* strains Br642 and Br776. Surprisingly, only small and turbid plaques (1 mm) arise after infection of *P. aeruginosa* PAO1, although a slightly more efficient adsorption (91% in 5 min) and no significant difference in latent period or burst size could be noted (Supplementary figure S1). Despite the suspicion of lysogeny, all attempts to isolate stable lysogenic *P. aeruginosa* clones failed. DNA was extracted from phage-resistant *P. aeruginosa* clones isolated from turbid plaques and batch cultures. No phages could be induced, and no integrated LUZ24 DNA sequences could be demonstrated by PCR or restriction analysis.

Genome sequence and regulation. The LUZ24 particle encapsulates a linear dsDNA molecule consisting of 45,625 bp (52% GC). **The genome was sequenced with an average redundancy of 5.6, and every base position was covered by at least three independent sequence reactions.** Phage LUZ24 encodes 68 proteins, 47 of which are arranged in rightward orientation while 21 aim leftward (Table 1, Figure 2), and is delineated by two 184 bp direct terminal repeats (DTRs) (Supplementary Figure S2). As found in similarly organized phages, the opposing transcriptional gene clusters are separated by a strong bidirectional terminator (Table 2).

The genome of LUZ24 displays an overall nucleotide identity of 71% to phage PaP3. Eight small insertions/deletions are present and 88% of the encoded gene products are related between both phages (Table 1, Figure 2). However, the *cos* site identified in PaP3 is not conserved in LUZ24, **and no difference in restriction patterns could be observed if samples were heated to 80°C prior to separation on a 1% agarose gel (data not shown).** Moreover, only two tRNA genes (tRNA^{Asn} and tRNA^{Pro}) are present at the right end of the LUZ24 genome, compared to four in PaP3. The evolutionary advantage of these tRNA genes seems

rather small, since no clear codon preference in LUZ24 compared to *P. aeruginosa* could be detected.

To experimentally identify bacterial promoters in LUZ24, a threefold redundant fragment (200-300 bp) library was generated in the broad-spectrum promoter trap vector pTZ110 (Schweizer and Chuanchuen, 2001). Promoter activity was determined quantitatively in *P. aeruginosa* cells by measuring the β -galactosidase activity in a Miller assay (Miller, 1992). Seven β -galactosidase-specific promoter regions were identified, spread throughout the genome and displaying a clear *P. aeruginosa* consensus sequences (TTGACa-N₁₇-TATaaT, Table 2) (Dominguez-Cuevas and Sylvia, 2004). This indicates the dependency of LUZ24 on the transcriptional machinery of its host throughout the entire infection cycle. Promoters located in the right arm are clearly stronger, which might be explained by the presence of A-rich tracts upstream the -35 and -10 regions, a phenomenon also observed in phages T4 and T5 (Miller *et al.*, 2003).

Genes involved in host conversion and metabolism. Since LUZ24 (and PaP3) diverged strongly from other currently investigated phages, only a limited number of protein functions could be predicted by similarity searches. Especially proteins presumed to be involved early in the infection process can differ considerably, even between very closely related phages (Figure 2) (Ceyssens *et al.*, 2006). One exception is ORF2, which is conserved in PaP3, in the entire β -KMV genus and in phage PA11, suggesting a crucial role in infection and/or conversion of *P. aeruginosa*.

The middle genome region of LUZ24 encodes proteins involved in biosynthesis (gp18-22) followed by a phage T7-type genome replication module, with a primase-helicase (gp24) encoded directly upstream the DNA polymerase. Based on mutual sequence comparisons, we propose the annotation of a new ORF in PaP3 (ORF23.1, 21674-21465), which is perfectly well conserved in LUZ24 (ORF46). While the PaP3 DNA polymerase is composed of two separate polypeptides (Tan *et al.*, 2007). LUZ24 carries an endonuclease VII (ORF35) inserted within the 'core' polymerase, splitting the DNA polymerase gene in three separate

parts (ORFs 25, 34 and 36, Figure 3). However, closer inspection revealed the presence of a group I intron-like sequence, which may have inserted into the phage genome through the process of homing ([Sandegren and Sjöberg, 2007](#)). To confirm the presence of this predicted intron, total RNA was isolated from infected *P. aeruginosa* cells PCR performed on first-strand cDNA yielded smaller products compared to control reactions on LUZ24 DNA template (Figure 3). Subsequent cycle sequencing using ORF primer P_F and internal primers P_{f1}, P_{f2} and P_{ir} revealed the absence of the predicted 668 bp fragment (19145-19812) in the cDNA. The self-splicing of this intron completely restores the reading frame of the large DNA polymerase subunit.

Despite the lack of nucleotide similarity to other phage introns, the embedded endonuclease is related to the intron-associated endonuclease of *Synechococcus* phage S-PM2 (Mann *et al.*, 2005) and resembles (among others) free-standing endonucleases of T7-like *Yersinia* phages Berlin and phiYeO3-12. A comparable 787 bp insertion containing a similar endonuclease is also present in the DNA polymerase gene of T7-like phage Pf-WMP3 (**EF537008**), strongly suggesting the presence of a self-splicing intron in the genome of this cyanophage.

The presence of this intron within the DNA polymerase gene of LUZ24 confirms a common theme seen in many phage genomes, i.e., that phage introns target highly conserved regions within functionally important genes. It is speculated that these introns might confer a selective advantage to the phage by offering the possibility of regulating the expression of the intron-containing genes by the regulation of splicing (Mann *et al.*, 2005). By providing the first experimental evidence of a self-splicing intron in the genome of a *Pseudomonas* phage, we support the current hypothesis that there seems no barrier for intron maintenance within the phages of the Proteobacteria (Bonocora and Shub, 2004).

Particle formation and host lysis. The genes transcribed from the right arm are involved in virion assembly and host lysis, and follow a conserved gene order with the large subunit of the terminase (gp66) encoded downstream the portal (gp65), scaffolding (gp63) and the capsid protein (gp62). Following this conserved genomic organization, upstream ORFs

probably encode the internal core, tail tube and tail fibers, but only a tail collar domain (gp57) could be predicted through similarity searches. To obtain a more comprehensive view of the structural region, individual proteins within the mature phage particle were identified with ESI-MS/MS. Mass spectrometric analysis of denatured phage particles, with and without prior fractionation on a 1D-SDS PAGE gel (Figure 1B), identified nine structural gene products with a total protein coverage ranging from 13 to 88% (Table 1). This analysis delineates the structural region from gp49 to gp65, and confirms gp62 as the most abundant protein in the particle. Surprisingly, gp57 has a tail collar domain and displays similarity to tail fiber proteins of many *Yersinia* phages, but is not identified in this analysis. This is most likely due to the underrepresentation of arginine and lysine residues in its primary protein sequence, which precludes suitable trypsinization for peptide identification. All other non-identified predicted proteins in this region are rather small (<12 kDa) and may have fallen below the threshold for detection.

The encoded terminases of LUZ24 (gp66) and PaP3 (gp3) are strongly related to the well-characterized terminase of phage P22 (Blastp e-value of 4E-69), which catalyses the sequential headful packaging of P22 DNA leading to both terminal redundancy and circular permutation (Wu *et al.*, 2002).

The lysozyme of LUZ24 (gp67) is encoded between both subunits of the terminase, and carries the conserved catalytic residues E32 and D41. It lacks a cell wall binding module, but is probably able to reach the peptidoglycan layer with the help of gp31. This short protein has three transmembrane domains and carries a positively charged intracellular C-terminal end (DRVFHKKA), typical for class I holins (Wang *et al.*, 2000).

DISCUSSION

It seems unlikely that the newly isolated phage LUZ24 has a temperate nature, despite being very homologous to PaP3. LUZ24 shares its bidirectional genome organization with a vast number of virulent phages like cyanophages P60, Pf-WMP3 and Pf-WMP4, roseophage SIO1 and vibriophage VpV262 (Liu *et al.*, 2007). Corresponding to the latter phages, it carries

DTRs at the genomic ends and lacks any gene or regulatory systems which suggest integration or lysogenic behaviour in its host bacterium. The *P. aeruginosa* tRNA^{lys} gene is proposed as PaP3 integration site (Tan *et al.*, 2007), but in contrast to other tRNA-integrating phages like mycobacteriophage Ms6 (Freitas-Vieira *et al.*, 1998) and *Streptococcus* phage T12 (McShan and Ferretti, 1997), no site-specific recombinase is encoded upstream or downstream the PaP3 *attP* site. Although filamentous ssDNA phages like ctxI† exist which recruit host-encoded recombinases, underlying mechanisms differ profoundly and integration-associated proteins like RstB are encoded in the ctxI† genome ([Davis and Waldor, 2000](#)).

Furthermore, neither immunity nor reactivation of the integrated PaP3 DNA was demonstrated by Tan *et al.* (2007). These arguments call into question the temperate nature of PaP3, although very closely related temperate and lytic phages have been reported before, i.e. phages L5 and D29 (Ford *et al.*, 1998).

Another intriguing finding is that the end structures of virion chromosomes of both PaP3 (5'-protuding cohesive ends) and LUZ24 (184 bp DTRs, Supplementary Figure 2) differ from those observed in P22, despite the fact that their terminases are strongly related. The same observation was made in the recently sequenced *E. coli* phage phiEco32, which carries 193-bp direct repeats and also encodes a P22-like terminase (Savalia *et al.*, 2008). At first glance, this seems to contradict current theory that the structure of virion DNA ends can be accurately predicted based on amino acid similarity of the predicted terminase (Casjens *et al.*, 2005). However, it is common among all the relevant terminases that the large terminase subunit does not determine the cut sites through sequence specificity. All four nicks in LUZ24 and both nicks in PaP3 are at different sequences, suggesting that some other component of the packaging apparatus leads the terminase to its cut sites. This might explain how the large terminase subunits can be 97% identical between LUZ24 and PaP3 and yet participate in a different end cutting strategy. This hypothesis implies however that another (yet unspecified) function that must differ between LUZ24 and PaP3.

Despite the differences stated above, LUZ24 (isolated in Belgium) and PaP3 (isolated in China) remain closely related phages and represent a new genus within the *Podoviridae*

family. The world-wide spread of closely related phages has been reported repeatedly, and maybe best illustrated by the occurrence of nearly identical structural gene sequences in environments as far-reaching as the Southern Ocean, the Gulf of Mexico and a melt-water pond on an Arctic ice shelf (Short and Suttle, 2005). LUZ24 and PaP3 lack sequence similarity to other phages, but the closely related gene map and a strikingly conserved order of structural genes suggest (very ancient) divergence from phages like P60, Pf-WMP4, SIO1 and VpV262. Despite the enormous global phage diversity which arose through the course of evolution, the number of truly different phages infecting a given host appears to be finite. Our efforts to build a comprehensive view of the global genomic diversity among *Pseudomonas* phages will continue.

MATERIALS AND METHODS

Phage manipulations. Phage LUZ24 was isolated by spotting of a cleared and filtered (0.45 μm) hospital sewage sample (UZ Leuven, Belgium) on lawns of *P. aeruginosa* strain Li010. The phage was amplified on solid plate cultures, concentrated by PEG8000 precipitation and purified by two successive CsCl gradient centrifugations. **Phage particles were stained with 2% Uranyl acetate (pH 4.5)**, followed by electron microscope imaging performed according to Ackermann (2008). Analysis of structural proteins, host range, pH stability, adsorption efficiency, latent period and burst size were performed as described elsewhere (Lavigne *et al.*, 2006; Carlson, 2005; Ceysens *et al.*, 2008). LUZ24 was sequenced by a combination of shotgun sequencing (pJET1 vector, Fermentas, Germany) and primer walking, followed by *in silico* characterization as described by Ceysens *et al.*, 2006. Phage particles were destabilized by four successive rounds of freeze-thawing and sonication, heated for 10 min at 95 $^{\circ}\text{C}$ and reduced in the presence of 10 mM DDT for 1 h at 56 $^{\circ}\text{C}$. Disulfide bonds were blocked by alkylation with 10 mM Iodoacetamide, followed by trypsinisation of the entire mixture and identification of structural phage particle proteins as described by Lavigne *et al.* (2006).

In vivo splicing assay. *P. aeruginosa* Li010 was grown to an OD_{600} of 0.4 and infected with LUZ24 at a multiplicity of 10 per cell. Ten minutes after infection, 4 ml cells were harvested

and quickly frozen to -80°C . Total RNA was isolated using an RNAeasy mini-prep kit (QIAGEN), removing contaminating DNA with RNase-free DNase. Two microgram of total RNA was incubated with 20 pmol of the sequence-specific primer P_R (5'-GCCAAAGGACAGTTAACCCCGAGGAGTTC-3'), followed by first-strand cDNA synthesis using the RevertAidTM H Minus kit (Fermentas). Resulting cDNA was purified, ten times diluted and used for PCR amplification and subsequent cycle sequencing with primers P_R, P_F (5'-ATCATCAGCCGCCAGAGTAAACGTGGAG-3') and internal primers P_{f1} (5'-GGGTAAGTGGAAGACTGTAGGAAAG-3'), P_{f2} (5'-GGGCTATGGCT-GATGAACGAG-3') and P_{ir} (5'-CAATGACTTGTGGATCATCACGTCGCC-3').

The genome of bacteriophage LUZ24 is deposited at Genbank under accession number [NC 010325](#).

ACKNOWLEDGEMENTS

PC holds a predoctoral fellowship of the 'Instituut voor de aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen' (I.W.T., Belgium). KH and RL are postdoctoral fellows of the 'Fonds voor Wetenschappelijk Onderzoek-Vlaanderen' (FWO-Vlaanderen, Belgium). This work was financially supported by the research council of the KULeuven grant OT/05/47 and the flemish FWO research grant G.0307.05.

REFERENCES

Ackermann, H.-W., 2008. Basic phage electron microscopy. In: Kropinski, A.M., Clokie, M. (Eds.), *Bacteriophages: Methods and Protocols*. Humana Press, USA, in press.

Bonocora, R.P., Shub, D.A., 2004. A Self-Splicing Group I Intron in DNA Polymerase Genes of T7-Like Bacteriophages. *Journal of Bacteriology* 163, 8153-8155.

[Byrne, M.](#), [Kropinski, A.M.](#), 2005. The genome of the *Pseudomonas aeruginosa* generalized transducing bacteriophage F116. [Gene](#) 346, 187-194.

Carlson, K., 2005. Working with Bacteriophages: Common techniques and methodological approaches. In: Kutter, E., Sulakvelidze, A. (Eds.), *Bacteriophages, Biology and Applications*. CRC Press, Florida, pp. 437-494.

[Casjens, S.R.](#), [Gilcrease, E.B.](#), [Winn-Stapley, D.A.](#), [Schicklmaier, P.](#), [Schmieger, H.](#), [Pedulla, M.L.](#), [Ford, M.E.](#), [Houtz, J.M.](#), [Hatfull, G.F.](#), [Hendrix, R.W.](#), 2005. The generalized transducing Salmonella bacteriophage ES18: complete genome sequence and DNA packaging strategy. *Journal of Bacteriology* 187, 1091-1104.

[Ceysens, P.J.](#), [Lavigne, R.](#), [Mattheus, W.](#), [Chibeu, A.](#), [Hertveldt, K.](#), [Mast, J.](#), [Robben, J.](#), [Volckaert, G.](#), 2006. Genomic analysis of *Pseudomonas aeruginosa* phages LKD16 and LKA1: establishment of the phiKMV subgroup within the T7 supergroup. *Journal of Bacteriology* 188, 6924-6931.

[Ceysens, P.J.](#), [Mesyanzhinov, V.](#), [Sykilinda, N.](#), [Briers, Y.](#), [Roucourt, B.](#), [Lavigne, R.](#), [Robben, J.](#), [Domashin, A.](#), [Miroshnikov, K.](#), [Volckaert, G.](#), [Hertveldt, K.](#), 2008. The genome and structural proteome of YuA, a new *Pseudomonas aeruginosa* phage resembling M6. *Journal of Bacteriology* 190, 1429-1435.

[Davis, B.M.](#), [Waldor, M.K.](#), 2000. CTXphi contains a hybrid genome derived from tandemly integrated elements. *Proceedings of the National Academy of Sciences U S A* 97, 8572-8577.

[Dominguez-Cuevas P.](#), [Sylvia, S.](#), 2004. Compiling sigma-70 dependent promoters. In: [Ramos, J.L.](#) (ed.), *Pseudomonas: virulence and gene regulation, vol. II*. Kluwer Academic/Plenum, Boston, MA, pp 319-343.

Ford, M.E., Sarkis, G.J., Belanger, A.E., Hendrix, R.W., Hatfull, G.F., 1998. Genome structure of mycobacteriophage D29: implications for phage evolution. *Journal of Molecular Biology* **279**, 143-164.

[Freitas-Vieira, A.](#), [Anes, E.](#), [Moniz-Pereira, J.](#), 1998. The site-specific recombination locus of mycobacteriophage Ms6 determines DNA integration at the tRNA(Ala) gene of *Mycobacterium* spp. [Microbiology](#) 144, 3397-3406.

Kovalyova, I.V., Kropinski, A.M., 2003. The complete genomic sequence of lytic bacteriophage gh-1 infecting *Pseudomonas putida*-evidence for close relationship to the T7 group. [Virology](#) 311, 305-315.

[Kwan, T.](#), [Liu, J.](#), [Dubow, M.](#), [Gros, P.](#), [Pelletier, J.](#), 2006. Comparative genomic analysis of 18 *Pseudomonas aeruginosa* bacteriophages. [Journal of Bacteriology](#) 188, 1184-1187.

[Lavigne, R.](#), [Burkal'tseva, M.V.](#), [Robben, J.](#), [Sykilinda, N.N.](#), [Kurochkina, L.P.](#), [Grymonprez, B.](#), [Jonckx, B.](#), [Krylov, V.N.](#), [Mesyanzhinov, V.V.](#), [Volckaert, G.](#), 2003. The genome of bacteriophage phiKMV, a T7-like virus infecting *Pseudomonas aeruginosa*. [Virology](#) 312, 49-59.

[Lavigne, R.](#), [Noben, J.P.](#), [Hertveldt, K.](#), [Ceysens, P.J.](#), [Briers, B.](#), [Dumont, D.](#), [Roucourt, B.](#), [Krylov, V.N.](#), [Mesyanzhinov, V.V.](#), [Robben, J.](#), [Volckaert, G.](#), 2006. The structural proteome of *Pseudomonas aeruginosa* bacteriophage phiKMV. [Microbiology](#) 152, 529-534.

[Liu, X.](#), [Shi, M.](#), [Kong, S.](#), [Gao, Y.](#), [An, C.](#), 2007. Cyanophage Pf-WMP4, a T7-like phage infecting the freshwater cyanobacterium *Phormidium foveolarum*: complete genome sequence and DNA translocation. [Virology](#) 366, 28-39.

[Mann, N.H.](#), [Clokie, M.R.](#), [Millard, A.](#), [Cook, A.](#), [Wilson, W.H.](#), [Wheatley, P.J.](#), [Letarov, A.](#), [Krisch, H.M.](#), 2005. The genome of S-PM2, a "photosynthetic" T4-type bacteriophage that infects marine *Synechococcus* strains. [Journal of Bacteriology](#) 187, 3188-3200.

McShan WM, Ferretti JJ., 1997. Genetic diversity in temperate bacteriophages of *Streptococcus pyogenes*: identification of a second attachment site for phages carrying the erythrogenic toxin A gene. [Journal of Bacteriology](#) 179, 6509-6511.

Miller, J.H., 1992. A Short Course in Bacterial Genetics. CSH Laboratory Press, Cold Spring Harbor, NY.,

[Miller, E.S.](#), [Kutter, E.](#), [Mosig, G.](#), [Arisaka, F.](#), [Kunisawa, T.](#), [RÄ¼ger, W.](#), 2003. Bacteriophage T4 genome. [Microbiology and Molecular Biology Reviews](#) 67, 86-156.

[Pirnay, J.P.](#), [De Vos, D.](#), [Cochez, C.](#), [Bilocq, F.](#), [Vanderkelen, A.](#), [Zizi, M.](#), [Ghysels, B.](#), [Cornelis, P.](#), 2002. *Pseudomonas aeruginosa* displays an epidemic population structure. [Environmental Microbiology](#) 4, 898-911.

[Sandegren, L.](#), [SjÄ¶berg, B.M.](#), 2007. Self-splicing of the bacteriophage T4 group I introns requires efficient translation of the pre-mRNA in vivo and correlates with the growth state of the infected bacterium. [Journal of Bacteriology](#) 189, 980-990.

[Savalia, D.](#), [Westblade, L.F.](#), [Goel, M.](#), [Florens, L.](#), [Kemp, P.](#), [Akulenko, N.](#), [Pavlova, O.](#), [Padovan, J.C.](#), [Chait, B.T.](#), [Washburn, M.P.](#), [Ackermann, H.W.](#), [Mushegian, A.](#), [Gabisonia, T.](#), [Molineux, I.](#), [Severinov, K.](#) 2008. Genomic and proteomic analysis of phiEco32, a novel *Escherichia coli* bacteriophage. [Journal of Molecular Biology](#) 28, 774-789.

[Schweizer, H.P.](#), [Chuanchuen, R.](#), 2001. Small broad-host-range lacZ operon fusion vector with low background activity. [Biotechniques](#) 31, 1258-1260.

Short, C.M., Suttle, C.A., 2005. Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. [Applied and Environmental Microbiology](#) 71, 480-486.

[Tan, Y.](#), [Zhang, K.](#), [Rao, X.](#), [Jin, X.](#), [Huang, J.](#), [Zhu, J.](#), [Chen, Z.](#), [Hu, X.](#), [Shen, X.](#), [Wang, L.](#), [Hu, F.](#), 2007. Whole genome sequencing of a novel temperate bacteriophage of *P. aeruginosa*: evidence of tRNA gene mediating integration of the phage genome into the host bacterial chromosome. [Cellular Microbiology](#) 9, 479-491.

[Wang, I.N.](#), [Smith, D.L.](#), [Young, R.](#), 2000. Holins: the protein clocks of bacteriophage infections. [Annual Reviews Microbiology](#) 54, 799-825.

Wu, H., Sampson, L., Parr, R., Casjens, S., 2002. The DNA site utilized by bacteriophage P22 for initiation of DNA packaging. *Molecular Microbiology* 45, 1631-1646.

FIGURE LEGENDS



TABLES



<i>Location</i>	<i>Promoter sequence</i>	<i>Activity^a (Miller Units)</i>
511-556	CGGGGGTTGACAAGGGATCGGTGAAGCGGTATAGTTCGCCCCGGTCAAG	357
615-660	CACCGCTTGACACTGTAAGGCTCAGTCGGTATGATGGGCACCACTCCCT	237
4828-4873	TAGGCCTTGACATTCTGCTGGAAGTGTGGTATAATAACCTTAAGGTGTC	120
20312-20357	TGGGGGTTGACTTTCAGCCCCCTCTGTGGTATAATACCTTCTTCCCTAC	289
45121-45076	TCAAAAAGTGCTTGACAAGGTATCGAAAATGTAGTATAATAGACCTATA	684
44923-44878	AGAAAAAAGACTTGACAAAATAGAAAAAGTGTGATATAATAGTATTATA	845
30353-30308	ATAAAAGGTACTTGACAAAGTATCCAAAATGTAGTATAATAGACCTATA	1084
Terminator Sequences ($\hat{\alpha}^{\dagger}G$, kcal/mol)		
25725-25768	AAAAAAAAGGCCCAAGGGTATCATCCAAGGGGCCTTATCTTT (-21.5) ACGGGGCCCTTCGGGGCCCTTGTTT (-17.8)	

Table 2. Regulatory sequences of LUZ24.

^a Measurements were performed in threefold, activity units were calculated as described by Miller (1992).

11

Figure 1. A. Electron microscopic image of a negatively stained LUZ24 particle. Scale bar represents 50 nm, capsid decorations are marked by arrows. **B.** Phage particle proteins separated on a 12% SDS-PAGE gel, in parallel to a LMW-size ladder (kDa).

Figure 2. Side-by-side comparison of the LUZ24 and PaP3 genomes, marking the amino acid identities of the corresponding ORFs in different shades of blue. Insertions and deletions are indicated, as are the experimentally identified $\check{I}f^{70}$ promoters (arrows with length according to the promoter strength) and rho-independent terminators (open circles). One copy of the LUZ24 DTR (block arrow) is also present in the genome of PaP3. The identified structural proteins are marked with asterisks, and the newly annotated PaP3 gene is hatched.

Figure 3. Schematic representation of the genomic region encoding the three DNA polymerase polypeptides, with demonstration of *in vivo* splicing of the intron DNA interrupting ORF34. Lanes 1, 2 and 3 of the 1% agarose gel contain PCR product obtained on LUZ24 DNA with primers P_F-P_R, P_F-P_{ir} and P_{f2}-P_R, respectively. Lanes 4, 5 and 6 contain PCR product obtained from a cDNA template, using the same primer combinations. Lane 7, Smartladder (Eurogentec).

Table 1. Bioinformatic analysis of the LUZ24 genome.

<i>ORF</i>	<i>Location</i>	<i>Mw (kDa)</i>	<i>Strand</i>	<i>Best phage homolog</i>	<i>E-value</i>	<i>% AA coverage^a</i>	<i>Comments</i>
1	1020-1190	6.3	+	No similarity			
2	1208-1762	21.0	+	LKD16 gp3	3e-5		Also present in \check{I}^{\dagger} KMV, LKA1 and PaP3
3	1974-	8.2	+	PaP3 p69	2e-6		

	2192						
4	2189-2329	5.7	+	PaP3 p68	8e-14		
5	2326-2409	3.1	+	PaP3 p67	4e-5		
6	2509-2691	6.9	+	PaP3 p66	4e-6		
7	2823-3056	8.5	+	PaP3 p65	1e-10		
8	3058-3333	10.5	+	PaP3 p61	6e-30		
9	3487-3633	5.6	+	PaP3 p60	7e-21		
10	3642-3893	9.4	+	PaP3 p59	3e-25		
11	3883-4164	10.6	+	PaP3 p58	3e-47		
12	4164-4823	24.4	+	PaP3 p55	2e-27		C-terminal ATPase domain [Pfam domain PF02449, E= 0.75]
13	4979-5131	8.3	+	No similarity			
14	5443-6405	3.6	+	PaP3 p50	6e-39		
15	6424-7380	35.8	+	PaP3 p49	3e-50		
16	7377-8177	29.6	+	PaP3 p48	1e-94		Conserved in PA11 [E= 4e-30]
17	8170-8739	21.7	+	PaP3 p47	8e-47		
18	8715-9890	44.4	+	PaP3 p46	4e-115		Conserved in PA11 [E= 2e-22]
19	9902-11434	55.9	+	PaP3 p45	0.00		L-glutamine-D-fructose-6-phosphate amidotransferase GATase_2 domain [Pfam domain IPR000583, E= 2e-6]
20	11444-11665	8.6	+	PaP3 p44	5e-18		Conserved in PA11 [E= 1e-21]
21	11793-12677	32.2	+	PaP3 p43	9e-97		Glutathion synthase [Pfam domain RimK, E= 2e-3]
22	12677-13075	14.8	+	PaP3 p42	9e-43		Biosynthesis related [Pfam domain UPF0131, E= 2e-4]

23	13075-13452	14.2	+	PaP3 p41	2e-26		
24	13453-15162	64.4	+	PaP3 p40	0.00		Primase/helicase Conserved in PA11 [E= 8e-69]
25	15146-15655	19.7	+	PaP3 p39	4e-85		DNA polymerase (~T7 DNA pol. res. 2-186)
26	15668-15937	9.7	+	PaP3 p38	5e-20		Contains one transmembrane domain
27	15969-16202	8.8	+	No similarity			
28	16183-16374	7.2	+	PaP3 p37	2e-16		
29	16364-16525	5.8	+	PaP3 p36	3e-08		
30	16516-16935	15.8	+	Mx9 p02	5e-13		
31	16932-17216	10.7	+	PaP3 p35	8e-34		Putative holin; 3 transmembrane domains
32	17341-17796	17.9	+	PaP3 p34	1e-19		
33	17801-17998	7.5	+	PaP3 p33	2e-20		
34	17999-19180	45	+	PaP3 p32	0.00		DNA polymerase (~T7 DNA pol. res. 207-525)
35	19250-19666	16.1	+	phage Berlin	2e-21		Endonuclease_7 [Pfam family PF02945, E= 5e-15]
36	20024-20305	10.7	+	PaP3 p32	4e-11		DNA polymerase (~T7 DNA pol. res. 621-706)
37	20375-20929	20.7	+	PaP3 p31	8e-88		
38	20907-21410	18.3	+	PaP3 p30	9e-81		Also present in $\dot{\text{I}}\ddagger\text{KMV}$ and LKD16
39	21436-21666	8.8	+	PaP3 p29	7e-20		
40	21666-22115	17.1	+	ES18 gp47	8e-20		Related to both gp2.8 and gp7.7 of T7
41	22200-23030	31.8	+	PaP3p28	4e-121		5'-3' Exonuclease [Pfam family PF01367, E= 4e-9] Conserved in PA11 [E= 5e-29]
42	23008-23952	35.5	+	PaP3 p27	1e-59		

43	24002-24184	7.2	+	PaP3 p26	4e-20		
44	24177-24938	29.3	+	PaP3 p25	1e-127		
45	24935-25153	8.1	+	PaP3 p24	2e-28		
46	25157-25366	7.8	+	PaP3 p23.1	6e-45		Newly defined in the PaP3 genome
47	25353-25559	7.7	+	PaP3 p23	2e-09		
48	26123-25767	12.6	-	PaP3 p21	2e-40		
49	27025-26138	31.6	-	PaP3 p20	1e-146	61.1	Structural protein
50	30204-27037	112.0	-	PaP3 p19	0.00	30.5	Structural protein
51	31872-30355	53.5	-	PaP3 p18	2e-169	16.2	Structural protein
52	32257-31877	12.6	-	PaP3 p17	1e-25	60.6	Structural protein
53	33198-32257	32.0	-	PaP3 p16	2e-113	15.3	Structural protein
54	33613-33179	16.6	-	PaP3 p15	2e-76		
55	34299-33610	25.2	-	PaP3 p14	3e-107		
56	35837-34296	57.5	-	PaP3 p13	0.00	14.2	Structural protein
57	36493-35846	22.0	-	PaP3 p12	1e-114		Phage Tail Collar Domain [Pfam family PF07484, E= 5.6e-19]
58	36731-36483	8.8	-	PaP3 p11	2e-17		
59	36927-36715	7.7	-	PaP3 p10	4e-17		
60	37543-36917	24.0	-	PaP3 p09	3e-111	12.9	Structural protein
61	37867-37547	12.0	-	PaP3 p08	3e-49		
62	38868-37915	35.0	-	PaP3 p07	2e-168	88.4	Major head protein
63	39879-38887	37.2	-	PaP3 p06	2e-108		Scaffolding protein
64	40094-	8.2	-	PaP3 p05	4e-21		

	39879						
65	42244-40124	81.1	-	PaP3 p04	0.00	35.9	Portal protein
66	43692-42244	54.6	-	PaP3 p03	0.00		Terminase, large subunit [Pfam family PF03237, E= 1.2e-111]
67	44090-43692	14.8	-	PaP3 p02	1e-55		T4-like Lysozyme
68	44580-44122	16.7	-	PaP3 p01	1e-69		Terminase, small subunit (PsiBlast)
	44680-44607		-				tRNA-Pro [tRNAscan-SE]
	44852-44780		-				tRNA-Asn [tRNAscan-SE]

^aPart of the predicted structural protein which was identified as part of the phage particle by mass spectrometry.

