Characterization and genome sequencing of two Propionibacterium acnes phages displaying pseudolysogeny

Lood, Rolf; Collin, Mattias

Published in:
BMC Genomics

DOI:
10.1186/1471-2164-12-198

2011

Link to publication

Citation for published version (APA):
Lood, R., & Collin, M. (2011). Characterization and genome sequencing of two Propionibacterium acnes phages displaying pseudolysogeny. BMC Genomics, 12. https://doi.org/10.1186/1471-2164-12-198

Total number of authors:
2

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Characterization and genome sequencing of two Propionibacterium acnes phages displaying pseudolysogeny

Rolf Lood* and Mattias Collin

Abstract

Background: Propionibacterium acnes is a Gram positive rod inhabiting the human skin that also infects orthopaedic implants and is associated with acne vulgaris. Previously, one lytic bacteriophage, PA6, from P. acnes has been sequenced and partially characterized. We recently isolated several inducible phages from P. acnes classified as Siphoviruses based on morphology and partial genome sequencing.

Results: In this study we sequenced the inducible P. acnes phages PAD20 and PASS50, isolated from deep infection and from skin, respectively. The genomes of PAD20 and PASS50 are 29,074 and 29,017 bp, respectively, compared with the 29,739 bp of PA6. The phage genomes have 87.3-88.7% nucleotide sequence identity. The genes are divided into clusters with different levels of similarity between the phages. PAD20 and PASS50 share four genes encoding identical amino acid sequences. Some deletions and insertions in the genomes have occurred, resulting in lack of genes, frame shifts, and possible regulatory differences. No obvious virulence factor gene candidates were found. The phages are inducible, but bacteria can be cured of phages by serial colony isolations and lose their phages during stationary phase, but are still sensitive to new phage infections. Construction of a phylogenetic tree based on more than 459 phage genomes, suggested that P. acnes phages represent a new lineage of Siphoviruses.

Conclusions: The investigated P. acnes Siphovirus genomes share a high degree of homology to other P. acnes phages sequenced, but not to genomes of other phages isolated from Propionibacteria. The phage genomes are not integrated in the bacterial genome, but instead, most likely have a pseudolysogenic life cycle.

Background

Propionibacterium acnes is a Gram positive rod normally found on human skin [1], but has also been isolated from infections of orthopaedic implants [2,3], and has been suggested to be involved in the development of acne vulgaris. Recently, the first genome of a P. acnes strain was sequenced [4] and several putative virulence factors were identified [5]. The sequenced P. acnes genome only contains one cryptic prophage, but it has earlier been shown that some P. acnes strains are carriers of inducible phages [6]. Other groups have also studied Propionibacterium phages both in order to establish a tool to differentiate between serotypes, but also to further characterize phages mainly infecting bacteria important for the dairy industry [7-16].

Recently, Farrar et al, sequenced the first genome of a P. acnes phage, PA6 (GenBank accession no. DQ431235) [17]. The genome consists of 29,739 bp dsDNA with 48 putative coding genes. The phage was determined to be a lytic phage since no integrase or repressor could be detected in the genome. Eighteen of the genes were assigned predicted functions based on their similarities to other known phage genes. Recently, we isolated and characterized inducible phages from P. acnes isolated from both deep infections and from skin of healthy individuals [6]. By comparing the genomes of these different phages, a better understanding of how the phage influences its host and an insight in the great genetic diversity seen among phages, as shown
by other comparative studies on phage genomes can be achieved [18,19]. Phages can enter two principally different life cycles, e. g. a lytic life cycle or a lysogenic cycle. In the lysogenic cycle most phages integrate their genome into the host genome. This is most often regulated by a phage-borne integrase, but there is also a recent report of a filamentous phage, CTXφ from *Vibrio cholerae*, that uses the host recombinases XerC and XerD to integrate into the host genome [20,21]. Furthermore, there are a few phages known to form stable linear or circular plasmids as a part of their integration mechanism [22-24], but most known temperate phages encode their own integrases. Other than those two well-defined phage life cycles, phages are also known to be able to interact with their host in other life cycles, termed as persistent infections, pseudolysogeny, carrier-state, and chronic infections [25,26]. Those phage life cycles, even though common, are less defined compared to the lytic and lysogenic life cycles. Common for all states, except the chronic infection, is that the bacteria carry the phage DNA, but the DNA is not integrated into the bacterial chromosome, but rather exists as a plasmid. The chronic infection is characterized by the release of phages from the bacteria without causing lysis [25]. In this study, we have sequenced the genome of one phage induced from a *P. acnes* strain isolated from deep infection (PAD20), and the genome of one phage induced from a *P. acnes* strain isolated from skin (PAS50), and we have compared the two genomes with the reported lytic phage PA6 [17]. The phages have earlier been shown to have dissimilarities in the genes encoding a putative major head protein and an amidase [6], but no other genes have been examined. This offers an opportunity to investigate if phages isolated from *P. acnes* contribute to the virulence seen in different strains of *P. acnes*, and to examine the genetic diversity seen among phages.

**Results and discussion**

**Genomic organization and homology** The genomic organization of the three examined phages are very similar, even though some genes differ (see Figure 1). Both phages have a 13 bp ssDNA extension at both ends of the genome, identical to that of phage PA6. These extensions enable the formation of potential concatamers and circularization of the phage genome. The genome of PAD20 is 29,074 bp and has a GC content of 54.1%, compared to the genome of PAS50 that is 29,017 bp and has a GC content of 53.9%. PAD20 has two larger deletions in the genome compared to PA6. The first is a 226 bp deletion resulting in the deletion of the main part of gp22 and thereby in a frame shift giving rise to a combination of the gene products Gp22 and Gp23 (Gp22.23, see additional file 1). The second larger deletion is a 378 bp deletion of the entire gp42 gene. PAS50 has two larger deletions and one insertion. The first deletion in PAS50 is a 207 bp deletion resulting in the deletion of the gp45 gene, and a second deletion of almost 500 bp in a non-coding region, corresponding to bp 27,932-28,419 in PA6. The insert in PAS50 is an 84 bp insertion in gp37 resulting in a 28 amino acid larger exonuclease.

To further investigate the genomic organization of the different *P. acnes* phages, the genomes were screened for putative sigma70-dependent promoters and transcriptional terminators (see Figure 1 and Table 1). Several putative promoters were detected in all phages upstream of gp7, gp11 and gp46, as well as upstream of gp45 and gp30, potentially enabling the phage to transcribe different parts of the genome at different stages of infection. One of the promoters with the highest score was upstream of a hitherto unannotated ORF (3693-3301, gp50). However,
the predicted gene product does not show any similarity to known sequences or domains.

Except for the six identified promoters in each genome, one putative transcriptional terminator was also identified (see Figure 1). The putative terminators are placed immediately downstream of the amidase and the holin in the lysis gene cassette.

Even though most of the genes are very similar in the examined phages, a certain pattern of similarity between gene clusters can be observed. Packaging genes and putative head genes have a higher similarity between PA6 and PAS50 than between the other examined phages, a certain pattern of similarity between PAD20 and PAS50 than between the other examined phages, a certain pattern of similarity between PAD20 and PAS50 than between the other examined phages.

By searching for homologues corresponding to PAD20 and PASS0 proteins using BLASTp [27], most hits were in the Actinobacteria phylum, with Streptomyces, Mycobacterium and phages infecting Mycobacterium, Propionibacterium, Nocardia, and Bifidobacterium (see Table 2). This is not surprising since these species are closely related and thereby also might have similarities among their phages. Most hits were on bacterial proteins, most likely indicating that these bacterial strains have prophages that have not yet been examined. More interestingly, several phage proteins showed similarity to phages isolated from low G+C bacteria as Streptococcus and Lactococcus (see Table 2). These species are not very closely related to P. acnes, but the results might indicate that their phages have had several possibilities to exchange genes. Surprisingly, there were no hits on any other Propionibacterium phages or prophages. If this merely reflects the lack of sequenced phages from Propionibacterium genus, or indicate that phages isolated from P. acnes are distinct from those infecting other Propionibacteria is unknown. Furthermore, the putative cryptic prophage in

### Table 1 Putative regulatory elements

| Phage | Position | gp   | Sequence          | Score |
|-------|----------|------|-------------------|-------|
| Promoters |          |      |                   |       |
| PA6   | 5620-5655 | gp7  | GTGCAGGGGTTGTGTGATGGTATCATTTGAAG | 73/16 |
| PA6   | 7124-7159 | gp11 | TTAGATGTTGTGTTTGGTTGATATGTTT | 33/58 |
| PA6   | 28708-28742 | gp46 | TTAAAGCTTATAAGTTATATAATAATAAGTA | 68/31 |
| PA6   | 3740-3704 | gp50 | TAGTAAGGCTCCCTCTTGTGCTATAATCCGACCT | 80/5 |
| PA6   | 20693-20656 | gp30 | TTGACACAGCAGCTAGACAGTCGCA | 72/66 |
| PA6   | 28763-28724 | gp45 | TTATATACCTTTAAAGCCTTTATATAATAATAATTTAT | 61/39 |
| PAD20 | 5618-5653 | gp7  | GTGCAGGGGTTGTGTGATGGTATCATTTGAAG | 73/16 |
| PAD20 | 7104-7140 | gp11 | CTGAGTAAACGCTTTATTTTGTGTTTTATGATTTG | 55/3 |
| PAD20 | 28024-28063 | gp46 | TTAGATGTTGTGTTTGGTTGATATGTTT | 61/31 |
| PAD20 | 3738-3704 | gp50 | GTGAGTGGCCTCTCTTGTGCTATAATCCGACCT | 80/5 |
| PAD20 | 20473-20436 | gp30 | TTGACACAGCAGCTAGACAGTCGCA | 72/66 |
| PAD20 | 28084-28048 | gp45 | TTATAAGCTTTAATCTTTAAGGTTATATAT | 90/24 |
| PASS0 | 5615-5650 | gp7  | GTGCAGGGGTTGTGTGATGGTATCATTTGAAG | 73/16 |
| PASS0 | 7101-7137 | gp11 | CTGAGTAAACGCTTTATTTTGTGTTTTATGATTTG | 55/3 |
| PASS0 | 27980-28014 | gp46 | TTAGATGTTGTGTTTGGTTGATATGTTT | 68/31 |
| PASS0 | 3740-3704 | gp50 | TTATATAATCTTATAGGTATATAAAAT | 61/39 |
| Terminators |          |      |                   |       |
| PA6   | 16665-16713 | gp21 | GTGCCCCAGGGGGTTGCTGAGGTATGTGTGTTGCTGCTGGCCGGCTGGGAC | -39.9 |
| PAD20 | 16682-16730 | gp21 | GTGCCCCAGGGGGTTGCTGAGGTATGTGTGTTGCTGCTGGCCGGCTGGGAC | -39.9 |
| PASS0 | 16604-16652 | gp21 | GTGCCCCAGGGGGTTGCTGAGGTATGTGTGTTGCTGCTGGCCGGCTGGGAC | -39.9 |

The gp number represents the ORF directly after the regulatory element. Bold letters indicate the putative -10 and -35 site of the promoters, and the palindromes in the terminators. The score is given as -10/-35 for the promoters based on BPROM as described in Methods, and as minimal free energy (MFE) in kcal/mol for the terminators.
| Gene  | Location (PAD20/PASS50) | Function                      | Hits                                              | E-value PAD20 | E-value PASS50 |
|-------|-------------------------|-------------------------------|--------------------------------------------------|---------------|---------------|
| gp1   | 2-373, 2-373            | Small terminase               | Mycobacterium phage Brujita gp1                  | 8e-06         | 3e-05         |
|       |                         |                               | Streptomyces phage VWB gp32                      | 3e-05         | 5e-05         |
| gp2   | 378-1889, 378-1889      | Large terminase               | Bifidobacterium hypothetical protein BFDEN_01707 | 9e-79         | 2e-78         |
|       |                         |                               | Streptococcus phage terminase                    | 2e-37         | 2e-37         |
|       |                         |                               | Lactococcus phage r1t putative terminase         | 5e-37         | 4e-36         |
| gp3   | 1886-3211, 1886-3211    | Portal protein                | Corynebacterium hypothetical protein cur_1413    | 2e-70         | 6e-69         |
|       |                         |                               | Mycobacterium phage Lockley gp12                 | 3e-36         | 5e-33         |
|       |                         |                               | Lactococcus phage phiLC3 putative portal protein | 3e-29         | 2e-28         |
|       |                         |                               | Streptococcus phage SM1 gp34                     | 7e-27         | 3e-26         |
| gp4   | 3218-3973, 3218-3973    |                               | Mycobacterium phage TM4 gp6                      | 2e-06         | 3e-05         |
|       |                         |                               | Rhabdovirus nucleocapsid protein                 | 7.3e-02       | 2e-01         |
| gp5   | 4084-4638, 4081-4635    | Scaffold protein              | Strongylucrotus hypothetical protein             | 5.9e-02       | 1.1e-01       |
|       |                         |                               | Mycobacterium phage Tweety gp5                   | 9.5e-01       | 9e-03         |
|       |                         |                               | Streptococcus phage SM1 gp39                     | 1.1e-00       | 1.1e-00       |
| gp6   | 4645-5592, 4642-5589    | Major head protein            | Mycobacterium phage Che9d gp7                    | 8e-49         | 7e-50         |
|       |                         |                               | Streptococcus phage SM1 gp40                     | 1e-30         | 2e-30         |
|       |                         |                               | Lactococcus phage phiLC3 MHP                      | 2e-30         | 1e-29         |
| gp7   | 5636-6097, 5633-6094    |                               | Mycobacterium phage Halo gp9                     | 5e-26         | 4e-26         |
| gp8   | 6099-6446, 6096-6443    |                               |                                                |               |               |
| gp9   | 6453-6743, 6450-6740    |                               |                                                |               |               |
|       |                         | Streptococcus conserved      |                                                | 8.3e-01       | 2.5e-00       |
| gp10  | 6740-7111, 6737-7108    | Amidase domain                |                                                | 4.7e-01       | 6.1e-01       |
|       |                         |                               |                                                |               |               |
| gp11  | 7164-7793, 7161-7799    | Major tail protein            |                                                | 4e-21         | 6e-21         |
|       |                         |                               |                                                |               |               |
|       |                         | Streptococcus major tail     |                                                | 2e-10         | 3e-09         |
|       |                         | protein                      |                                                | 5e-04         | 1e-04         |
|       |                         |                                |                                                | 2e-03         | 2e-03         |
|       |                         | Mycobacterium phage          |                                                | 1.2e-02       | 7e-03         |
| gp12  | 7821-8114, 7827-8123    |                               |                                                | 3.7e-01       | -             |
|       |                         | Ralstonia DNA topoisomerase  |                                                |               |               |
|       |                         | IV subunit A                 |                                                |               |               |
| gp13  | 8213-8500, 8222-8509    | Putative transmembrane protein |                                                | 1.8e-01       | -             |
| gp14  | 8508-11273, 8517-11282  | Tape measure protein         |                                                | 5e-51         | 3e-54         |
|       |                         |                                |                                                |               |               |
|       |                         |                               |                                                | 7e-32         | 1e-40         |
|       |                         |                                |                                                | 2e-28         | 2e-28         |
|       |                         | Tape measure protein         |                                                |               |               |
| gp15  | 11289-12230, 11298-12239| Minor tail protein            |                                                | 9e-06         | 2e-08         |
| Protein ID | E-value (Phage Lxx24170) | E-value (Phage Lxx24170) |
|------------|--------------------------|--------------------------|
| gp16       | 1.00e-03                 | 0.00e-00                 |
| gp17       | 3.90e-02                 | 1.00e-03                 |
| gp19       | 2.00e-03                 | 1.00e-03                 |
| gp20       | 4.90e-03                 | 2.30e-02                 |
| gp21       | 1.90e-00                 | 2.20e-00                 |
| gp24       | 9.00e-03                 | 1.90e-02                 |
| gp25       | 7.10e-00                 | 0.00e-01                 |
| gp26       | 6.70e-00                 | 3.30e-00                 |
| gp27       | 1.00e-27                 | 6.60e-27                 |
| gp28       | 1.00e-23                 | 3.60e-22                 |
| gp30       | 3.90e-00                 | 2.10e-01                 |
| gp31       | 3.10e-31                 | 4.00e-27                 |
| gp32       | 2.00e-05                 | 1.00e-04                 |
| gp33       | 1.10e-02                 | 3.00e-02                 |
| gp34       | 1.60e-01                 | 1.60e-01                 |
| gp35       | 1.70e-01                 | 4.10e-01                 |

**Table 2 Phage protein similarities (Continued)**

Leifsonia phage-related protein Lxx24170

Bacillus phage minor structural protein

Streptomyces hypothetical protein SAV_S506

Salinospora hypothetical protein Strop_0531

Geobacillus phage-related protein GK0542

H-type lectin domain

Streptococcus collagen-like surface protein

Propionibacterium putative amidase

Mycobacterium phage Chah gp50

Mycobacterium phage Tweety gp32

Lactococcus phage rht holin

Sigma 70, region 4

HTH transcription regulator

Prochlorococcus phage P-SSM4 gp7

Prochlorococcus hypothetical protein A9001_14091

Mycobacterium phage TM4 gp36

Nocardioida hypothetical protein nfa39310

Hyaluronidase domain

Aspergillus hypothetical protein An14g04790

Salinospora hypothetical protein Strop_0519

Mycobacterium phage Solon gp33

Topoisomerase primase

DNA primase

DNA primase

DNA primase

DNA primase

NERD (nuclease related domain)

Llaj restriction endonuclease

Holliday junction resolvase
P. acnes strain KPA171202 [5] does not show any similarity to PAD20, PAS50 or PA6. However, most recently, Machuca et al [28] isolated a phage infecting Fusobacterium nucleatum. This phage was partially sequenced (390 bp), and showed 90% sequence identity to parts of gp3 (portal protein) and gp4 in PA6. However, the whole phage genome of the Fusobacterium phage needs to be sequenced to conclude how related this phage is to Propionibacterium phages.

**Table 2 Phage protein similarities (Continued)**

| gp | Mycobacterium phage Cjw1 gp87 | Phage PAD20 and PAS50 protein sequences were screened for similar proteins using BLAST. All proteins showed the highest similarity to P. acnes phage PA6. Those hits were not further analyzed since they only indicate that the phages are most related to each other. | 1.6e-00 | 4.3e-01 |
| gp34 | 22529-21666 | DNA helicase | 2e-47 | 3e-45 |
| gp35 | 23836-22572 | 23201-22737 | 3.3e-01 | 1.1e-00 |
| gp36 | 23488-23078 | 23652-23242 | Prochlorococcus possible ABC transporter related | 1.8e-00 | 1.8e-00 |
| gp37 | 24432-23503 | 24680-23649 | Exonuclease | Mycobacterium phage Bethlehem gp69 | 2e-16 | 6e-16 |
| gp38 | 24785-24432 | 25033-24680 | Bartonella phage-related protein Btr_0991 | 6e-11 | 3e-13 |
| gp39 | 25094-24891 | 25342-25139 | Mycobacterium phage TM4 gp52 | 8e-10 | 4e-11 |
| gp40 | 25318-25091 | 25566-25339 | | | |
| gp41 | 25876-25343 | 26124-25591 | Alkalilimnicola outer membrane lipoprotein LolB | 2.6e-02 | - |
| gp42 | - | 26492-26208 | Rhodobacterales methylamine utilization protein MauG | - | 2.3e-01 |
| gp43 | 26311-26000 | 26932-26621 | | | |
| gp44 | 26653-26387 | 27274-27008 | | | |
| gp46 | 28207-28362 | 28150-28305 | Borderella probable transcriptional regulator | 4.8e-00 | 4.8e-00 |
| gp47 | 28492-28626 | 28435-28569 | Endonuclease | HNH endonuclease | 5e-07 | 1e-06 |
| gp48 | 28723-29025 | 28663-28068 | Mycobacterium phage TM4 gp92 | 1e-03 | 3e-03 |
| gp49 | 28626-75 | | Topoisomerase II-associated protein PAT1 | 3.9e-02 | 5e-03 |
| gp50 | 3693-3301 | 3693-3301 | Zaire Ebola Virus VP30 | - | 1.7e-01 |
| | 3693-3301 | 3693-3301 | Zaire Ebola Virus minor nucleoprotein | - | 1.9e-01 |

**Phages encoding packaging proteins and structural proteins**

The packaging and structural proteins (Gp1-Gp19) are generally very well conserved, with high sequence
similarities between all three phages, ranging from 89.7% to 99.3% amino acid identity (see Figure 1 and additional file 2) with one exception - gp19. Gp19 has 67.1% to 77.4% amino acid identity and is predicted to function as a tail protein, with high similarity to collagen. This protein is likely to be involved in host specificity, so a more diverged gene would imply that the phages might have differences in host specificity. The difference in host specificity between PAD20 and PAS50 is however not very pronounced, since both PAD20 and PAS50 are equally potent to initiate lysis in most examined *P. acnes* isolates, even though some isolates can be used to differentiate the phages [6]. This does not rule out that a larger screen would detect more differences in host specificity and more clearly point out the genetic varieties that determine the host specificity.

Another gene of particular interest is gp17 encoding a protein that has not yet been assigned a putative function. Gp17 has a H-type lectin domain with E-values ranging from 0.001-0.04 in the phages examined. Even though the E-values are rather low, they might still give an indication of the putative function. H-type lectins are known to interact with carbohydrates, and the genomic position of gp17, close to genes encoding tail proteins, makes it possible that this protein is also a tail protein interacting with carbohydrates in order to bind to certain structures and potentially mediate a first binding site for the phage.

**The lysis mechanism**

All three phages encode an amidase, Gp20, with very high similarity to a *P. acnes* autolysin. They also encode a putative holin (Gp21) with more pronounced sequence divergence. Phage PA6 and PAS50 have 94% amino acid identity in the holin protein, while PA6 and PAD20 only have 84.3% amino acid identity. Interestingly, phage PAD20 has a deletion in gp22 resulting in a mainly deleted gp22 and a new open reading frame encoding a protein with similarities to both Gp22 and Gp23 in PA6 (see additional file 1). Gp22 will not be fully translated in PAS50 due to two internal stop codons.

**Identical proteins and putative virulence factors**

Several genes (gp39, gp40, gp44 and gp47) encode identical amino acid sequences in PAD20 and PAS50, but none of these proteins have predicted functions, nor any significant BLASTp hits, so these genes seem to be specific for *P. acnes* phages, and might, due to their sequence similarity be a good tool to use in phylogenetic analyses of phages isolated from *P. acnes*. Gp29, a gene in PA6, is absent in both PAD20 and PAS50, but another ORF in the same region (gp29.1, 19,490-20,056 in PA6) exists in all three phages and might thereby be more likely to be a true coding gene. However, none of the putative gp29.1 genes have similarities to other genes, neither has gp29 in PA6.

Some proteins did get BLAST hits indicating that the proteins could be putative virulence factors. Among those are Gp34 that is annotated as a DNA helicase due to a DnaB helicase domain (E-value 2e-04 for PAD20). More interesting, this protein also has a Multidrug Resistance Protein-like transporter domain (E-value 7e-06 for PAD20). This domain is less pronounced in PAS50 (E-value 0.15) and PA6 (E-value 0.001). PAS50 also has a weak hit (E-value 0.21) in Gp30 on a hyaluronidase. However, due to the low similarities to other proteins, and thereby the low E-values, we can not assign a function to the phage proteins based on the similarities to other domains, only determine which domains they are most similar to.

The phages might also be able to circularize and form concatamers due to the ssDNA extensions on each side of the genome. When circularized, a new putative gene (gp49) can be transcribed. This gene encodes a protein with some similarity to different nucleoproteins from Ebola virus, and has a topoisomerase II-associated protein PAT-1 domain (E-value 0.005 for PAS50). Thus, this gene can only be transcribed when the phage has circularized and potentially integrated, or formed concatamers.

**Identification of phage proteins**

To verify that some of the genes encoded proteins, PAD20 was precipitated with PEG 6000, the proteins were separated on a polyacrylamide gel, and then analyzed with MS/MS. All visible bands but band F were identified (see Figure 2 and Table 3). One of the bands, band E, was identified as a *P. acnes* succinate dehydrogenase flavoprotein subunit, thus indicating that this protein exists as a large complex, thereby being eluted in the same fractions as the whole phage particles during the FPLC gel filtration. Other than the *P. acnes* flavoprotein, all proteins were identified as phage proteins. Bands A-D were all identified as the major head protein of PAD20 (Gp6), which has a calculated molecular weight of approximately 33 kDa. However, the bands A-D have a much higher molecular weight, starting from approximately 160 kDa for band D, reaching over 500 kDa for band A. Interestingly, we did not get any other hits on the major head protein with lower molecular mass proteins, indicating that these proteins tend to form complexes of 6-8 monomers, based on the apparent molecular weight, that are not disrupted even when boiling the samples with SDS and reducing agents. A similar phenomenon has been described before with the *Enterobacteria* phage HK97, which has complexes of covalent linked major head proteins, and not a single major head protein exist as a monomer [29]. Both the...
N-and C-terminal part of the major head protein were covered, showing that this protein is not processed after its translation, as some major head proteins require in order to establish their final conformation.

Other proteins that were identified were Gp14 (tape measure protein), Gp3 (portal protein), Gp16, Gp15 (minor tail protein) and Gp11 (major tail protein). Furthermore, band K did not result in a significant hit. However, the peptide with the highest score was similar to Gp17 from P. acnes phage PA6. This is further strengthen by that the apparent molecular weight of the protein band (29 kDa) is close to the theoretical molecular weight (28,835 Da). Thus, all phage proteins identified seem to be structural proteins, dominated by Gp6 (major head protein) and Gp11 (major tail protein).

Phage life cycle

The phages used in this study were originally isolated by mitomycin C treatment of P. acnes isolates. However, those phages were able to infect their host strains, and thus did not confer superinfection immunity [6]. This data made us conclude that the phages were lysogenic. A similar phenomenon has been reported for Leuconostoc oenos phages, where more than 65% of the bacterial isolates could be induced with mitomycin C and where the phages were able to infect their host strains [30].

To further investigate how stable the interaction between the P. acnes phages and the bacterium was, we performed ten serial single-colony isolations and then screened for phages using a PCR assay. All examined colonies had been cured from phages (data not shown), indicating that the phages are not true lysogens. This ability to cure the bacterial isolates from phages suggests that the interaction between the phage and bacterium rather might be pseudolysogenic [31-33]. The inability to cure bacteria from phages was one of the criteria used to determine that phages from the P. acnes related bacterium Nocardia erythropolis were lysogenic, due to that the spontaneous curing of phages was less than 0.4% after 5 serial single-colony isolations [34]. These phages were also induced with mitomycin C, but they all showed superinfection immunity. It has also been shown in several experiments that the environment may influence the virus-host interaction. This is evident in the case of phages infecting the archael species Halobacterium, where the NaCl concentration determines if the phages enter a lytic cycle, or cause a persisting infection. At low NaCl concentration, the phages effectively kill the host, while at higher concentrations of NaCl the lysis is delayed for several bacterial generations [35-37], leading to a persistent infection.

To further investigate if P. acnes might harbour pseudolysogenic phages, single colonies from the P. acnes
isolates AD20 and AS50 were inoculated into BHI broth and different parameters such as OD600 and pfu were measured two times a day (see Figure 3 and Table 4). No free infectious phages were detected at any time point, indicating that if any phages were spontaneously induced, their number have not exceeded the detection limit of 100 pfu/ml. However, at time points corresponding to mid-log phase and late-log phase, both AD20 and AS50 had 10-60% colonies harbouring phage DNA, as judged by PCR amplification of the major head gene (see Table 4). However, this carriage of phages is lost during the stationary growth phase, thus indicating that the phage DNA is not being replicated at this time point. Furthermore, the cells at the various time points were sensitive to PAD20 and PAS50 infection, indicating that the loss of phages were not due to the development of resistance (data not shown). This implies that the phages are not strictly lytic, neither strictly lysogenic.

Much of the data indicates that the P. acnes phages might have a pseudolysogenic cycle, which is also supported by the genetic information. There are no genes on the phage genomes with similarities to known integrases or repressors. It is, however, known that certain filamentous phages can integrate their genomes using host recombinases [20,21]. However, when extracting P. acnes chromosomal DNA, phage DNA is lost, as judged by the absence of a positive PCR reaction for the phage major head gene (data not shown), indicating that the phage indeed does not integrate. Thus, it seems like P. acnes phages can persist as extra-chromosomal plasmids, however when in this state, they are not replicating, since they easily are cured after 10 serial single-colony isolations.

### Phage phylogenetic analysis

Bacteriophages can be divided into different groups based on different criteria as host range, morphology

| Table 4 Carriage of phages during growth of P. acnes |
| Time (h) | AD20 | AS50 |
|---|---|---|
| 12 | 0/10 | 0/10 |
| 24 | 1/10 | 0/10 |
| 36 | 1/10 | 6/10 |
| 48 | 4/10 | 0/10 |
| 60 | 0/10 | 0/10 |
| 72 | 0/10 | 0/10 |
| 84 | 0/10 | 0/10 |
| 96 | 0/10 | 0/10 |
| 108 | 0/10 | 0/10 |
and genetic material. Earlier attempts to classify phages according to their proteome have been conducted by Rohwer et al using 105 phage genomes [38]. We aligned 453 bacteriophage genomes, including PAD20 and PAS50, using the progressive Mauve algorithm [39] in order to investigate how the P. acnes phages were related to other phages. Mauve recognizes conserved nucleotide segments in the genomes, but has the advantage that it can rearrange the order of these segments, thereby also identifying genome rearrangements [40]. However, Mauve has several limitations when it comes to comparing several diverged phage genomes, due to that Mauve works best with related organisms, and can not identify conserved regions shared only by a fraction of the genomes. Thus, aligning the genomes using Mauve will not generate a reliable phylogenetic tree. Instead, we used the progressive Mauve algorithm, which is developed to be able to align more diverged genomes with less than 50% nucleotide identity. Our results are in good agreement with the results obtained by Rohwer et al, but some of the phages are placed differentially. The Sk1-like Siphoviruses with Lactococcus phage skl and phage c2 form two closely related lineages, and also the lambda-like Siphoviruses align in a similar way as Rohwer et al. However, Lactococcus phage r1t is more closely related to Lactococcus phage Tuc2009 and TPP901-1 in our alignment than was showed previously. Furthermore, Acinetobacter phage AP205 is not as diverged from the other Leviphages in our model as Rohwer shows. Other than that, most phages cluster in the same groups as was shown by Rohwer et al, even though we have several more phages and subgroups of phages, resulting in not as distinct groupings as in Rohwer’s model. We found that the P. acnes phage genomes sequenced so far form a new phylogenetic lineage, well separated from other phages, with no close relatives (see Figure 4 and additional files 3 and 4). Furthermore, the different families of phages are widely distributed over all classes of bacteria, with Siphoviruses being the most abundant family reported in the database. Remarkably, even though PAD20 and PAS50 have a morphology that classifies them as Siphoviruses, some of their closest related phages are Propionibacterium freudenreichii phage B5 (inovirus), Acinetobacter phage AP205 (Levivirus) and Pseudomonas phage phi6 segment L (Cystovirus). Furthermore, a group of podoviruses including Yersinia phage Berlin seem to be closely related to a group of myoviruses with Vibrio, Aeromonas and Enterobacteria phages. Also of interest is that most phages infecting Mycobacterium and Streptomyces were rather closely related to phages from Propionibacterium. This is not surprising, since the bacterial hosts also are related. Of more interest is that Proteobacteria such as Pseudomonas, Burkholderia and Ralstonia phages also were related to phages infecting the Actinobacteria phylum, while other Proteobacteria as Vibrio and Salmonella were not. Other than that, bacterial families tend to have closely related phages, even though the classification of the phage families can differ in a bacterial species, as the very closely related Mycobacterium phage Rizal and Rosebush, that are classified as a Myovirus and a Siphovirus, respectively. Thus, this clearly shows that classification of phages solely based on morphology is not sufficient to determine to which family a phage should be assigned, but that genomic data is also necessary.

Conclusions

We have found that phages isolated from P. acnes are highly similar and show high nucleotide sequence identity. Several of the phage structural proteins were identified using MALDI-TOF MS/MS. The sequenced phages do not seem to encode any putative virulence factors. Finally we could show that the phages seem to be able to cause pseudolysogeny in their hosts.

Methods

Bacteriophages and bacterial isolates

Bacteriophages PAD20 and PAS50 were isolated as previously described [6]. Stock solutions were stored at 4°C in SM-buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 60 mM MgSO4 (7H2O), 0.01% gelatin). Bacterial isolates KPA171202, AD12, AD20, AD21, AD33 and AS50 were plated from frozen stocks on Tryptic Soy Broth supplemented with 1.5% agar (Bacto, Mt Pritchard, NSW, Australia) and incubated at 37°C under anaerobic conditions until colonies were clearly visible.

DNA sequencing and analysis

Since we had problems getting high concentrations of pure phage DNA, we adapted a PCR based approach. Phage lysate (1010 PFU/ml) was boiled for 5 min and 1 μl added as a DNA template to Fermentas High Fidelity PCR Enzyme Mix (Vilnius, Lithuania). Primers were designed to generate 4-6 kbp long overlapping products based on the already sequenced P. acnes phage PA6 genome, covering the whole genome (see Table 5). The overlapping regions were in most cases more than 200 bp, and confirmatory PCRs were run over the whole genome to exclude that we by this approach missed any inserts. The PCR was run under the following conditions: 94°C 2 min, 10 cycles of 94°C 30 sec, annealing temperature 30 sec, 72°C 6 min. This was followed by 35 repeats of 94°C 30 sec, annealing temperature 30 sec and 72°C 6 min (+20 sec/repeat). PCR products were cleaned up with SpinPrep PCR Clean-up Kit (Novagen, Madison, WI, USA). Products were sequenced using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied
Biosystems, Foster City, CA, USA) and analyzed using an ABI 3100 Genetic Analyzer (Applied Biosystems). For each sequencing, a new PCR generated DNA template was used, minimizing the impact of the PCR. When assembled, a >3-fold coverage with identical nucleotides was achieved. Sequences were assembled and aligned using the Clustal W algorithm [41]. Analyses were performed using the MacVector v9.5.2 software package (Cary, NC, USA). The sequence of the ssDNA ends was determined by similarity to phage PA6 after a PCR over the ends of the phage genome.

Identification of promoter regions and terminator structures
The genome of PAD20, PAS50 and PA6 were screened for putative sigma70-promoters using SAK and BPROM

Table 5 PCR oligonucleotides used

| Name | Sequence (5'-3') |
|------|-----------------|
| G1 fwd | AGTGAATACCTCCCTTTTGTTTTT |
| G1 rev | TTCTCCCTCACACAGCAACCA |
| G2 fwd | ATTTGTAAGGCGACGATTTTCTT |
| G2 rev | AATAAGCATACCAATAACAGGCACCA |
| G3 fwd | ATATTTGGATTGCCATGGCTGGTTTGG |
| G3 rev | ATCATGGCGTTACACCCGTC |
| G4 fwd | TTTGGGAAGGTGCTCTACGCT |
| G4 rev | CTACTACAAAAACCATCAAAACAAAGCCAC |
| G5 fwd | TTCTGTGCTGTTGGTGGATGTTTGG |
| G5 rev | GACAAACTCTCATTCTCCAAACATT |
| G6 fwd | AATGTAGCTTTCTCTGCTGCTGCCC |
| G6 rev | ATGTGTGTTTGGTTTTGGAAT |

Figure 4 Phages isolated from *P. acnes* represent a distinct phylogenetic lineage of Siphoviruses. More than 450 phage genomes were aligned using a progressive Mauve alignment. For visualization, 28 of the phages representing different clusters of most closely related phages to *P. acnes* phages were realigned using the same settings. The different colours represent different families of phages, as indicated by the key.
(Softberry, Inc.). Using SAK [42], a value higher than 1 was regarded as positive. Terminator structures were identified using FindTerm (Softberry, Inc.) and EMBOSS Explorer [43]. Palindromes followed by a high content of U and with a stable secondary structure (less than -20.0 in minimum free energy) were considered to be transcriptional terminators.

**Phage protein identification**

The phages were purified according to Boulanger [44], with some modifications. Briefly, bacterial lysates were treated with 0.5 M NaCl for 1 h at 4°C to separate phage particles from bacteria. The solution was centrifuged (10 min, 8,000 g, 4°C) and the supernatant moved to a new tube. PEG 6000 was added to a final concentration of 10%, and incubated for 1 h at 4°C. The phages were pelleted by centrifugation at 10,000 g for 15 min at 4°C. The supernatant was discarded and the pellet suspended in PBS. Whole phage particles were then separated from bacterial proteins using gelfiltration (Superose 6 10/300 GL, GE Healthcare, Sweden, Uppsala) operated on a ÄKTAprime plus (GE Healthcare) screening for the well-conserved major head gene, as described elsewhere [6].

**Serial single-colony isolation**

AD20 and AS50 were cultured on Brain Heart Infusion Agar plates (BHIA-plates) for three to four days, until visible single colonies were seen. Ten single colonies from each bacterium were selected, and ten serial streak isolations were made. The presence of phages in the final isolated colonies was detected by a PCR assay screening for the well-conserved major head gene, as described elsewhere [6].

**Growth curve of P. acnes isolates**

*P. acnes* isolates AD20 and AS50 were plated on BHIA-plates for three to four days. Single colonies were inoculated into pre-reduced BHI broth, and samples were analysed every 12th hour for OD<sub>600</sub>, cfu, infectious phages (pfu<sub>i</sub>), and carriage of phages (pfu<sub>c</sub>). Colony forming units were measured by dilution of the bacterial culture and plating on BHIA-plates. Pfui was measured by sterile filtration of the undiluted growth medium, using overlay plates as described elsewhere [6]. The detection limit using this method was >100 pfu/ml. To investigate the percentage of bacteria carrying phage DNA (pfu<sub>c</sub>), single colonies at each time point were subjected to PCR for the identification of the major head gene, using the recA gene as a positive control, as described elsewhere [6]. All experiments were performed in triplicates.

**Phage phylogenetic analysis**

The software Geneious (Geneious Pro 4.6.6, Biomatters Ltd) was used to download 451 different phage genomes from GenBank, searching for genomes with 'phage' in the title. The genomes for *Propionibacterium* phage PAD20 and PAS50 were added, and aligned using the software Mauve [39] with progressive Mauve, using the default settings. The generated guide-tree was then visualized with Geneious. For better visualization, 28 phage genomes were selected based on their localisation and their relationship with *P. acnes* phages on the complete phylogenetic tree, and a new
alignment with the same settings was run with only these phages.

**Nucleotide sequence accession number**
The GenBank accession numbers for the bacteriophages PAD20 and PAS50 genomes are FJ706171 (protein accession number ACX30796-ACX30840) and FJ706172 (protein accession number ACX30841-ACX30886), respectively.

**Additional material**

Additional file 1: Nucleotide sequence alignment for gp22 and gp23 in *P. acnes* phage PA6, PAD20 and PAS50. Underlined sequences represent different genes, as indicated by the colour of the line. Phage PAS50 has two internal stop codons (red) in the gene encoding gp22 and will not generate a full-length protein. Phage PAD20 has a deletion in a major part of the gene encoding gp22 resulting in a frame shift and a combined gene gp22.23.

Additional file 2: Gene comparison between phage PA6, PAD20 and PAS50. An ‘x’ indicates that the gene is absent, while a minus (−) indicates that no comparison of the genes were possible.

Additional file 3: Bacteriophage genomes used in phylogenetic alignment. Bacteriophages used in the phylogenetic alignment are ordered together with their accession number for reference.

Additional file 4: Phylogenetic tree of more than 450 bacteriophage genomes. More than 450 bacteriophage genomes were aligned using progressive Mauve. The generated tree-file was then visualized using the software Geneious. The different colours represents different families of bacteriophages: brown (Microvirus), pale green (Corticovirus), orange (Plasmavirus), light blue (Levivirus), gray (Cystovirus), purple (Tectivirus), and pink (Inovirus).

**References**

1. McGinley KJ, Webster GF, Leyden JJ. Regional variations of cutaneous propionibacteria. *Appl Environ Microbiol* 1978, 35(1):62-66.

2. Zetter V, Ghorbani A, Strady C, Leonard P, Mousouny P, Desplaces N. Propionibacterium acnes: an agent of prosthetic joint infection and colonization. *J Infect* 2007, 55(2):119-124.

3. Tunney MM, Patrick S, Gorman SP, Nixon JR, Anderson N, Davis RI, Hanna D, Ramage G. Improved detection of infection in hip replacements: a currently underestimated problem. *J Bone Joint Surg Br* 1998, 80(4):568-572.

4. Brüggemann H, Henne A, Hoster F, Liesegang H, Groschinsky S, Söderberg N. Bacteriophages infecting dairy propionibacteria. *Lact* 1995, 75:427-434.

5. Lood R, Mörgelin M, Holmberg A, Rasmussen M, Collin M. Inducible bacteriophage typing to distinguish *Propionibacterium acnes* types I and II. *J Clin Microbiol* 1998, 37(1):84-90.

6. Pulverer G, Sorgo W, Ko HL. Bakteriophagen von *Propionibacterium acnes*. *Zentralbl Bakteriol Parasitenk Infektionskr Hyg* 1: Abt Orig A 1973, 225:353-363.

7. Pulverer G, Ko HL. Fermentative and serological studies on *Propionibacterium acnes*. *Appl Microbiol* 1993, 25(2):222-229.

8. Gauthier M, Rouault A, Sommer P, Brandret R, Cassin D. Bacteriophages infecting dairy propionibacteria. *Lact* 1995, 75:427-434.

9. Pulverer G, Ko HL. Fermentative and serological studies on *Propionibacterium acnes*. *Appl Microbiol* 1993, 25(2):222-229.

10. Gautier M, Rouault A, Sommer P, Brandret R, Cassin D. Bacteriophages infecting dairy propionibacteria. *Lact* 1995, 75:427-434.

11. Hervé C, Coste A, Rouault A, Fraslin JM, Gautier M. Studies on bacteriophages of *Propionibacterium acnes* species in the sheep rumen. *Appl J Agric Res* 2000, 51:119-123.

12. Cheong JPE, Brooker JD. Isolation of a virulent bacteriophage from a *Propionibacterium acnes* species in the sheep rumen. *Appl J Agric Res* 2000, 51:119-123.

13. Borisova TG, Baranova NA, Vorobyova LI, Trusova ZM, Istomina IB, Gavrina TP, Tereblovka LS. A phage of propionic bacteria (*Russian*). *Prikl Biokhim Mikrobiol* 1973, 9:246-249.

14. Zerdin CH. Properties of Corynebacterium *acnes* bacteriophages and description of an interference phenomenon. *J Virol* 1974, 14(5):1268-1267.

15. Jong EC, Ko HL, Pulverer G. Studies on bacteriophages of *Propionibacterium acnes*. *Med Microbiol Immunol* 1975, 161(4):263-271.

16. Lee CN, Hu RH, Chou TY, Lin JW, Chen HY, Tseng YH, Weng SF. Comparison of Genomes of Three *Xanthomonas oryzae* Bacteriophages. *BMC Genomics* 2007, 8(1):442.

17. Tokitoh H, Yamasaki M, Sato J, Satoh N, Iwamoto Y, Itoh S, Kamo A, Hidaka H, Tomita M, Matsuoka S. Comparison of bacteriophages of the gram-positive bacterium *Propionibacterium acnes*. *Appl Environ Microbiol* 2000, 67(1):231-238.

18. Kwan T, Liu J, DuBow M, Gros P, Pelletier J. The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. *Proc Natl Acad Sci USA* 2005, 102(14):5174-5179.

19. Huber KE, Waldor MK. Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* 2002, 417(6889):656-659.

20. McLeod SM, Waldor MK. Characterization of XerC-and XerD-dependent CTX phage integration in *Vibrio cholerae*. *Mol Microbiol* 2004, 49(4):393-404.

21. Inal JM, Karunakaran KV. A temperate bacteriophage isolated from *Bacillus anthracis* exists as a plasmidial prophage. *Curr Microbiol* 1996, 32(4):171-175.

22. Gronis IS, Bourhy P, Ottone C, Picardeau M, Yelton D, Hendrix RW, Glaser P, Charon N. The LE1 bacteriophage replicates as a plasmid within *Leptospira biflexa*: construction of an *L. biflexa*-Escherichia coli shuttle vector. *J Bacteriol* 2000, 182(20):5700-5705.

23. Eggers CH, Cains J, Hayes SF, Garon CF, Dammann CJ, Oliver DB, Samuel DS. Bacteriophages of spirochetes. *J Mol Microbiol Biotechnol* 2000, 2(4):365-373.

24. Weinbauer MG. Ecology of prokaryotic viruses. *EMS Microbiol Rev* 2004, 28(2):127-181.

25. Porter K, Russ BE, Dyall-Smith ML. Virus-host interactions in salt lakes. *Curr Opin Microbiol* 2007, 10(4):418-424.

26. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990, 215(3):403-410.
30. Tenreiro R, Santos R, Brito L, Paveia H, Vieira G, Santos MA: Bacteriophages induced by mitomycin C treatment of Leuconostoc oenos strains from Portuguese wines. Letters in Applied Microbiology 1993, 16:207-209.

31. Welch M: Lysogenicity in streptomycetes. Ann N Y Acad Sci 1959, 81:974-993.

32. Bastias R, Higuera G, Sierra M, Espeso RT: A new group of cosmopolitan bacteriophages induce a carrier state in the pandemic strain of Vibrio parahaemolyticus. Environ Microbiol 2010, 12(4):990-1000.

33. Jarling M, Bartkowiak K, Robenek H, Pape H, Weinhardt F: Isolation of phages infecting Actinoplanes SN223 and characterization of two of these viruses. Appl Microbial Biotechnol 2004, 64(2):250-254.

34. Crockett JK, Brownell GH: Isolation and characterization of a lysogenic strain of Nocardia erythropolis. J Virol 1972, 10(4):737-745.

35. Torsvik T, Dundas ID: Persisting phage infection in Halobacterium salinarium str. 1. Journal of general virology 1980, 47:29-36.

36. Daniels LL, Wais AC: Virulence in phage populations infecting Halobacterium cutirubrum. FEMS Microbiology Ecology 1996, 25:129-134.

37. Daniels LL, Wais AC: Ecophysiology of Bacteriophage SS100 Infecting Halobacterium cutirubrum. Appl Environ Microbiol 1990, 56(11):3605-3608.

38. Rohwer F, Edwards R: The Phage Proteomic Tree: a genome-based taxonomy for phage. J Bacteriol 2002, 184(16):4529-4535.

39. Darling AC, Mau B, Blattner FR, Perna NT: Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 2004, 14(7):1394-1403.

40. Introduction to Mauve. [http://asap.ahabs.wisc.edu/mauve-aligner/mauve-user-guide/introduction-to-mauve.html]

41. Thompson JD, Higgins DG, Gibson TJ: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994, 22(22):4673-4680.

42. Gordon L, Chervonenkis AY, Gammerman AJ, Shahmuradov IA, Solovyev VV: Sequence alignment kernel for recognition of promoter regions. Bioinformatics 2003, 19(15):1964-1971.

43. Rice P, Longden I, Bleasby A: EMBOS: the European Molecular Biology Open Software Suite. Trends Genet 2000, 16(8):276-277.

44. Boulanger P: Purification of bacteriophages and SDS-PAGE analysis of phage structural proteins from ghost particles. Methods Mol Biol 2009, 502:227-238.

45. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970, 227(5259):680-685.

doi:10.1186/1471-2164-12-198

Cite this article as: Lood and Collin: Characterization and genome sequencing of two Propionibacterium acnes phages displaying pseudolysogeny. BMC Genomics 2011 12:198.