Genome sequence and characterization of five bacteriophages infecting *Streptomyces coelicolor* and *Streptomyces venezuelae*: Alderaan, Coruscant, Dagobah, Endor1 and Endor2

Aël Hardy, Vikas Sharma, Larissa Kever and Julia Frunzke*

Institute of Bio- und Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich, 52425 Jülich, Germany

* Correspondence: j.frunzke@fz-juelich.de; Tel.: +49 2461 615430

Abstract: *Streptomyces* are well-known antibiotic producers, and are also characterized by a complex morphological differentiation. *Streptomyces*, like all bacteria, are confronted with the constant threat of phage predation, which in turn shapes bacterial evolution. However, despite significant sequencing efforts recently, relatively few phages infecting *Streptomyces* have been characterized compared to other genera. Here, we present the isolation and characterization of five novel *Streptomyces* phages. All five phages belong to the *Siphoviridae* family, based on their morphology as determined by transmission electron microscopy. Genome sequencing revealed that four of them were temperate phages, while one had a lytic lifestyle. Moreover, one of the newly sequenced phages shows very little homology to already described phages, highlighting the still largely untapped viral diversity. Altogether, this study expands the number of characterized phages of *Streptomyces* and sheds light on phage evolution and phage-host dynamics in *Streptomyces*.

Keywords: phage isolation; phage genomics; *Streptomyces*; *Siphoviridae*; actinobacteriophages, actinorhodin

1. Introduction

Streptomyces is a genus of Gram-positive bacteria belonging to the order of Actinobacteria, and exhibit a high GC-content (on average about 73 mol% G + C). *Streptomyces* are prolific producers of natural products with a wide range of biological activities. This repertoire of bioactive molecules has been harnessed for medical and agricultural purposes, as for example ⅔ of known antibiotics of microbial origin are produced by *Streptomyces* [1-3].

Another distinctive feature of *Streptomyces* is their complex developmental cycle. Unlike most bacteria—that divide by binary fission, *Streptomyces* development is instead centered on the formation of spores. Germinating spores first form a network of interconnected cells, called vegetative mycelium. The vegetative mycelium later serves as a basis for the coordinated erection of an aerial mycelium. This is followed by the segmentation of these aerial filaments into spores, which can then start a new

39 cycle [3-5].

40 Phages infecting *Streptomyces* were described at a quick pace in the 1970-1980s, but
41 most of them were not later sequenced [6-8]. The phage phiC31 represents a notable
42 exception to this trend, as it was used to develop crucial genetic tools for *Streptomyces*
43 before being sequenced in 1999 [9-11]. Phages R4, SV1, VP5 were also the subject of
44 numerous studies, but the latter was not sequenced [12–13].

45   *Streptomyces* peculiarities were studied in the context of phage infection. For
46 example, adsorption to mycelium of phage Pal6 was shown to differ depending on
47 the stage of development of *Streptomyces albus* [14]. In this instance, phage
48 adsorption was found to be maximal for germinating spores. Combined with the
49 observation that germinating spores showed an intense average metabolic activity,
50 this suggests that spore germination represents the most sensitive development
51 stage for phage infection.

52   Conversely, the recent years have seen a sustained effort into the isolation and
53 sequencing of *Streptomyces* phages, notably by the SEA-PHAGES (Science Education
54 Alliance-Phage Hunters Advancing Genomics and Evolutionary Science) program
55 in the USA (https://seaphages.org/) [15]. However, few of these phages were
56 extensively characterized.

57   Here, we report the isolation, characterization and genome analysis of five novel
58 *Streptomyces* phages. Two of them (Alderaan and Coruscant) were isolated using *S.*
59 *venezuelae*, the remaining three (Dagobah, Endor1 and Endor2) were isolated using
60 *S. coelicolor*. Observation with transmission electron microscopy showed that all five
61 phages belong to the *Siphoviridae* family. Lifestyle prediction with the complete
62 nucleotide sequences revealed that four (Alderaan, Dagobah, Endor1 and Endor2)
63 are temperate, while Coruscant is thought to be a lytic phage. Alderaan, Coruscant,
64 Endor1 and Endor2 show close relatedness to already described *Streptomyces* phages
65 – Endor1 and Endor2 being highly homologous to each other. In contrast, Dagobah
66 showed very little relatedness to any sequenced phage, highlighting the still
67 massively untapped viral diversity.

68 **2. Materials and Methods**

69 *2.1 Bacterial strains and growth conditions*

70   *Streptomyces venezuelae* ATCC 10712 [16] and *Streptomyces coelicolor* M600 [17]
71 and strain M145 [18] were used as main host strains in this study. Cultures were
72 started by inoculating spores from spore stocks stored in 20% glycerol at -20°C [19].
73 *S. venezuelae* was grown in liquid Glucose Yeast Malt extract (GYM) medium, while
74 *S. coelicolor* was grown in liquid Yeast Extract Malt Extract (YEME) medium. Unless
75 otherwise stated, cultivation was carried out at 30°C. For double agar overlays, GYM
76 agar was used for both species, with 0.5% and 1.5% agar for the top and bottom
77 layers, respectively.

78

79 *2.2 Phage isolation and propagation*

Phages were isolated from soil samples taken near the Forschungszentrum Jülich (Germany). Phages contained in soil samples were resuspended by incubation in sodium chloride/magnesium sulfate (SM) buffer (10 mM Tris-HCl pH 7.3, 100 mM NaCl, 10 mM $MgSO_4$, 2mM $CaCl_2$) for 2 hours. The samples were centrifuged at $5,000 \times g$ for 10 min to remove solid impurities. The supernatants were filtered through a 0.22-μm pore-size membrane filter to remove bacteria. For each sample, 1 ml of filtered supernatant was mixed with 3 ml of liquid medium inoculated with 10^7 Streptomyces spores.

After overnight incubation, the culture supernatant was collected by centrifugation at $5,000 \times g$ for 10 min and filtered through a 0.22-μm pore-size membrane filter. Serial dilutions of the filtrate were then spotted on a bacterial lawn propagated by mixing 200 μl of *Streptomyces* overnight culture with 4 ml top agar, according to a modified version of the double agar overlay method [20]. Plaques were visualized after overnight incubation at 30°C.

Purification of the phage samples was carried out by restreaking single plaques twice [20]. Phage amplification was achieved by mixing 100 μl of the purified phage lysate into top agar to obtain confluent lysis on the plate. After overnight incubation, 5 ml of SM buffer were used to soak the plates and resuspend phages. The resulting phage lysate was centrifuged, and the supernatant was filtered to obtain the high-titer phage solution used for downstream processes.

To assess presence of actinorhodin, the plates were inverted and exposed to ammonia fumes for 15 min by placing 5 ml of 20% ammonium hydroxide solution on the inner surface of the lid.

2.3 Electron microscopy observation of phage virions

For electron microscopy, 5 μl of purified phage suspension were deposited on a glow-discharged formvar carbon-coated nickel grids (200 mesh; Maxtaform; Plano, Wetzlar, Germany) and stained with 0.5% (wt/vol) uranyl acetate. After air drying, the sample was observed with a TEM LEO 906 (Carl Zeiss, Oberkochen, Germany) at an acceleration voltage of 60 kV.

2.4 Phage infection curves

Infection in shake flasks (*S. venezuelae* phages): 70 ml GYM medium were inoculated with 10^5 spores and incubated at 30°C for 6-8 hours to allow spore germination. Phages were then added at the corresponding multiplicity of infection (MOI). OD_{450} was measured over time to assess bacterial growth. In parallel, the filtered supernatants of the cultures were collected at the same time points. 3 μl of these supernatants were spotted on a *Streptomyces venezuelae* lawn (inoculated to an $OD_{450}=0.4$) at the end of the experiment to estimate the phage titer.

Infection in microtiter plates (*S. coelicolor* phages): Growth experiments were performed in the BioLector® microcultivation system of m2p-labs (Aachen, Germany). Cultivation was performed as biological triplicates in 48-well

122 FlowerPlates (m2plabs, Germany) at 30 °C and a shaking frequency of 1200 rpm [21].
123 Backscatter was measured by scattered light with an excitation wavelength of 620
124 nm (filter module: λ_{Ex}/ λ_{Em}: 620 nm/ 620 nm, gain: 25) every 15 minutes. Each well
125 contained 1 ml YEME medium and was inoculated with 10^6 spores of *S. coelicolor*
126 M145. Phages were added after 7 h, and sampling was performed at the indicated
127 time points. Subsequently, 2 µl of the supernatants were spotted on a lawn of *S.*
128 *coelicolor* propagated on a double overlay of GYM agar inoculated at an initial
129 OD_{450}=0.4.
130

131 *2.5 Host range determination*

132 The host range of our phages was determined for the following *Streptomyces*
133 species: *S. rimosus* (DSM 40260), *S. scabiei* (DSM 41658), *S. griseus* (DSM 40236), *S.*
134 *platensis* (DSM 40041), *S. xanthochromogenes* (DSM 40111), *S. mirabilis* (DSM 40553),
135 *S. lividans* TK24 [22], *S. olivaceus* (DSM 41536) and *S. cyaneofuscatus* (DSM 40148). The
136 different *Streptomyces* species were grown in GYM medium, to which glass beads
137 were added to favor dispersed growth.
138 The host range was determined by spotting serial dilutions of phage solution on
139 lawns of the different *Streptomyces* species, in duplicates. A species was considered
140 sensitive to a given phage only if single plaques could be detected; we further
141 indicated if the phages are able to lyse a species (Table 1).

142 *2.6 DNA isolation*

143 For isolation of phage DNA, 1 µl of 20 mg/ml RNAse A and 1 U/µl DNAse
144 (Invitrogen) were added to 1 ml of the filtered lysates to limit contamination by host
145 nucleic acids. The suspension was incubated at 37 °C for 30 min. Then, EDTA,
146 proteinase K and SDS were added to the mixture at final concentrations of 50 mM
147 (EDTA and proteinase K) and 1% SDS (w/v), respectively. The digestion mixture
148 was incubated for 1 h at 56°C, before adding 250 µl of
149 phenol:chloroform:isopropanol. The content was thoroughly mixed before
150 centrifugation at 16,000 x g for 4 min.
151 The upper phase containing the DNA was carefully transferred to a clean
152 microcentrifuge tube and 2 volumes of 100% ethanol were added as well as sodium
153 acetate to a final concentration of 0.3 M. After centrifugation at 16,000 x g for 10 min,
154 the supernatant was discarded, and the pellet washed with 1 ml 70% ethanol.
155 Finally, the dried pellet was resuspended in 3 µl DNAse-free water and stored at
156 4°C until analyzed.
157
158
159

160 *2.7 DNA sequencing and genome assembly*

161 The DNA library was prepared using the NEBNext Ultra II DNA Library Prep
162 Kit for Illumina according to the manufacturer's instructions and shotgun-

163 sequenced using the Illumina MiSeq platform with a read length of 2 x 150bp
164 (Illumina). In total, 100,000 reads were subsampled for each phage sample, and *de*
165 *novo* assembly was performed with Newbler (GS De novo assembler; 454 Life
166 Sciences, Brandford, USA). Finally, contigs were manually curated with Consed
167 version 29.0 [23].
168

169 *2.8 Gene prediction and functional annotation*

170 Open reading frames (ORFs) in the phage genomes were identified with
171 Prodigal v2.6.3 [24] and functionally annotated using an automatic pipeline using
172 Prokka 1.11 [25]. The functional annotation was automatically improved and
173 curated with hidden Markov models (HMMs), and Blastp [26] searches against
174 different databases (Prokaryotic Virus Orthologous Groups (pVOGs)[27], viral
175 proteins and Conserved Domain Database CDD [28]), with the e-value cutoff 10^{-10}.
176 The annotated genomes were deposited in GenBank under the following
177 accession numbers: MT711975 (Alderaan), MT711976 (Coruscant), MT711977
178 (Dagobah), MT711978 (Endor1) and MT711979 (Endor2).
179 The ends of the phage genomes were determined with PhageTerm [29] using
180 default parameters. Phage lifecycle was predicted with PhageAI [30] using default
181 parameters.
182

183 *2.9 Genome comparison and classification*

184 To classify the unknown phage genomes at nucleotide level, 31 complete
185 reference actinophage genomes belong to different known clusters were
186 downloaded from the Actinobacteriophage Database [31]. The pairwise average
187 nucleotide identity (ANI) were calculated with the five unknown *Streptomyces*
188 phages including 31 reference genomes using the python program pyani 0.2.9 [32]
189 with ANIb method. The output average percentage identity matrix file generated
190 from pyani was used for clustering and displayed using the ComplexHeatmap
191 package in R [33]. Phage genome map with functional annotation was displayed
192 using the gggenes package in R.

193 *2.10 Protein domain-based classification*

194 An alternative approach was used to classify newly sequenced phages based on
195 conserved protein domains [28]. RPS-BLAST (Reverse PSI-BLAST) searches were
196 performed with e-value cutoff 0.001 against the Conserved Domain Database [28]
197 using the 2486 complete reference actinophages [31], including the newly sequenced
198 phage genomes. Identified Pfam protein domains output files from each phage
199 genome were merged and converted into a numerical presence-absence matrix. The
200 hierarchical clustering dendrogram was constructed with the help of the ward.2
201 method using the R platform. The resulting dendrogram was visualized using
202 ggtree [34].
203

3. Results

3.1 Phage isolation and virion morphology

Five novel phages infecting *Streptomyces* were isolated from soil samples close to the Forschungszentrum Jülich, Germany. The phages Alderaan and Coruscant were isolated using *Streptomyces venezuelae* ATCC 10712 and formed small, transparent and round plaques of approximately 2 mm of diameter (Figure 1A).

The phages Dagobah, Endor1 and Endor2 were isolated using *Streptomyces coelicolor* M600 as a host strain. Dagobah's plaques were very small (<1 mm) and were completely formed only after 2 days of incubation. Endor1 and Endor2 formed plaques of 2 mm in diameter with a distinct turbid zone in the center. Additionally, colored halos circling the plaques appeared after 3 days of incubation (Figure 1B). These halos were mostly brownish in the case of Dagobah, and reddish for Endor1 and Endor2. Exposure to ammonia fume resulted in an pronounced blue coloration around plaques, confirming that the halos surrounding plaques contained actinorhodin (Figure S1) [35].

TEM observation of the phage particles revealed that all five phages exhibit an icosahedral capsid and a non-contractile tail (Figure 1C). Based on the morphology, the phages were classified as members of the *Siphoviridae* family.
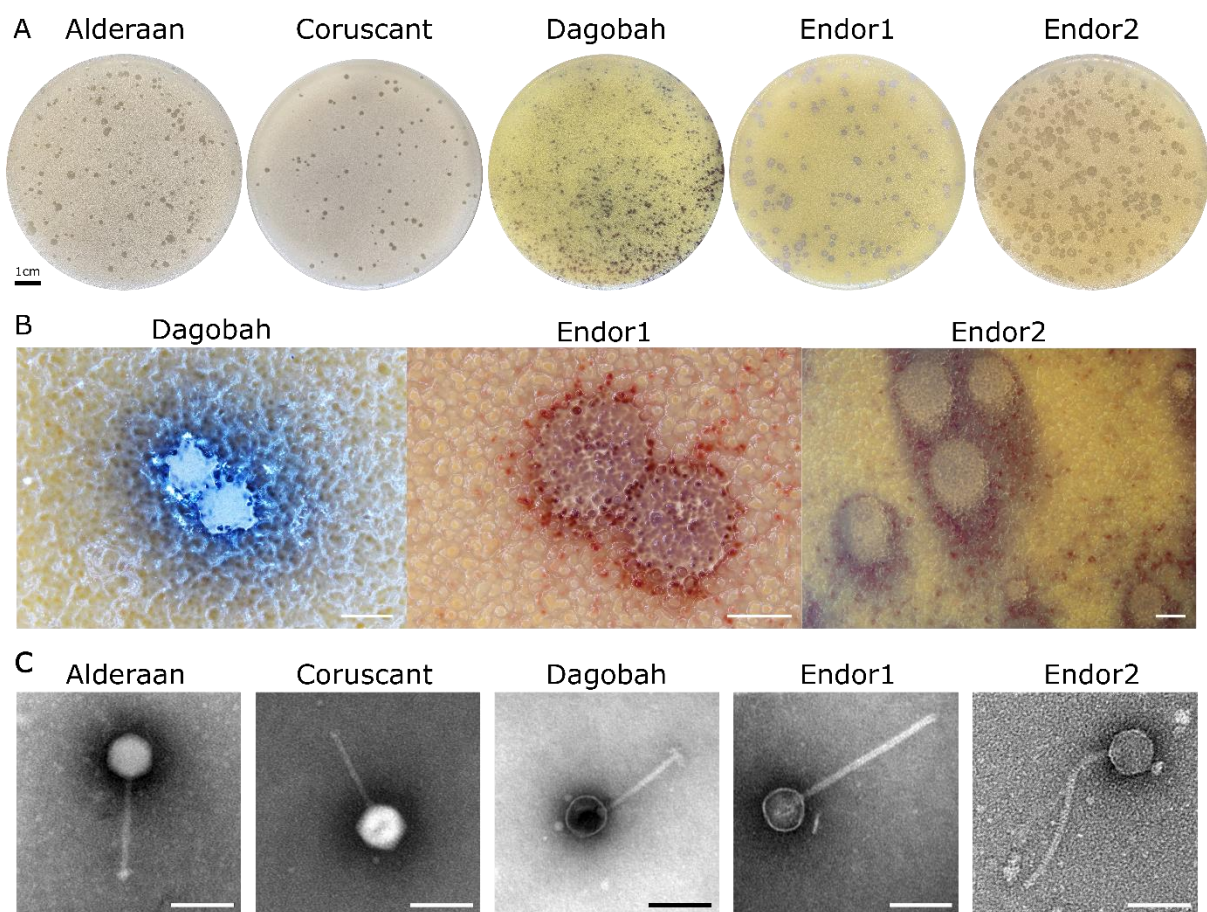
Figure 1. Morphology observation of five novel *Streptomyces* phages. (a) Plaque morphologies of the five phages. Double agar overlays were performed to infect *S.venezuelae* ATCC 10712 with the

226   phages Alderaan and Coruscant, and *S. coelicolor* M600 with the phages Dagobah, Endor1 and Endor2.
227   Plates were incubated overnight at 30°C and another day (3 days in the case of Dagobah) at room
228   temperature to reach full maturity of the bacterial lawn; (b) Close-ups of phage plaques imaged using
229   a stereomicroscope Nikon SMZ18. *S. coelicolor* M145 was infected by phages using GYM double agar
230   overlays. The plates were incubated at 30°C overnight and then kept at room temperature for two
231   (Endor1 and Endor2) or three days (Dagobah). Scale bar: 1 mm; (c) Transmission electron microscopy
232   (TEM) of phage isolates. The phage virions were stained with uranyl acetate. Scale bar: 150 nm.

233   *3.2 Infection curves and host-range determination*

234   Phage infection in liquid cultures was performed to assess infection dynamics.
235   Due to the complex developmental cycle of *Streptomyces*, standard one-step growth
236   curves could not be performed. We instead inoculated liquid cultures with spores of
237   *Streptomyces* and let them germinate for approximately 7 h before adding the phages
238   to a multiplicity-of-infection (MOI) from 0.1 to 10. For the *S. venezuelae* phages,
239   infection was performed in flasks and OD_{450} was used to estimate cell density. In
240   contrast, *S. coelicolor* was cultivated in microtiter plates, and cell growth was
241   monitored using continuous backscatter measurements. In both cases, phage titer
242   was measured over time to estimate the production of phage progeny.

243   Infection of *S. venezuelae* with Alderaan and Coruscant showed moderate lysis
244   for MOI 1, and distinct OD drops for the MOI 10, which was reduced to almost zero
245   after 24 h of infection (Figure 2A). Phage titers showed a significant increase after 16
246   h of infection and were markedly higher for MOI 10 than MOI 1.

247   As for the *S. coelicolor* phages (Figure 2B), infection with Dagobah caused a mild
248   growth delay, visible especially for the highest MOI (MOI 1). In parallel, the phage
249   titers grew moderately (10^2-fold increase between 0 and 48 h) or strongly (10^5-fold
250   increase between 0 and 48 h) for initially low (MOI 0.05) or high (MOI 1) MOIs,
251   respectively. In contrast, infection with Endor1 had a profound effect on bacterial
252   growth, as the highest MOIs (MOI 0.1 and 1) effectively suppressed growth. The
253   phage titers showed concordant behavior, with a strong increase from 16 hours and
254   a titer plateauing at a high level for MOI 0.1. Endor2 showed an intermediate effect:
255   the growth curves were significantly shifted, proportionally to the initial MOIs. At
256   low MOIs, the evolution of Endor1/2 titer was bell-shaped, with an initial increase
257   until 40 hours followed by a decline down to a virtually null titer at the end of the
258   experiment.

259   Furthermore, the backscatter started to decrease in the uninfected wells starting
260   from 50h, coinciding with the start of the production of blue-pigmented
261   actinorhodin. A similar drop was also observed in the samples infected with
262   Dagobah, Endor1 and the lowest MOI of Endor2.

263

264
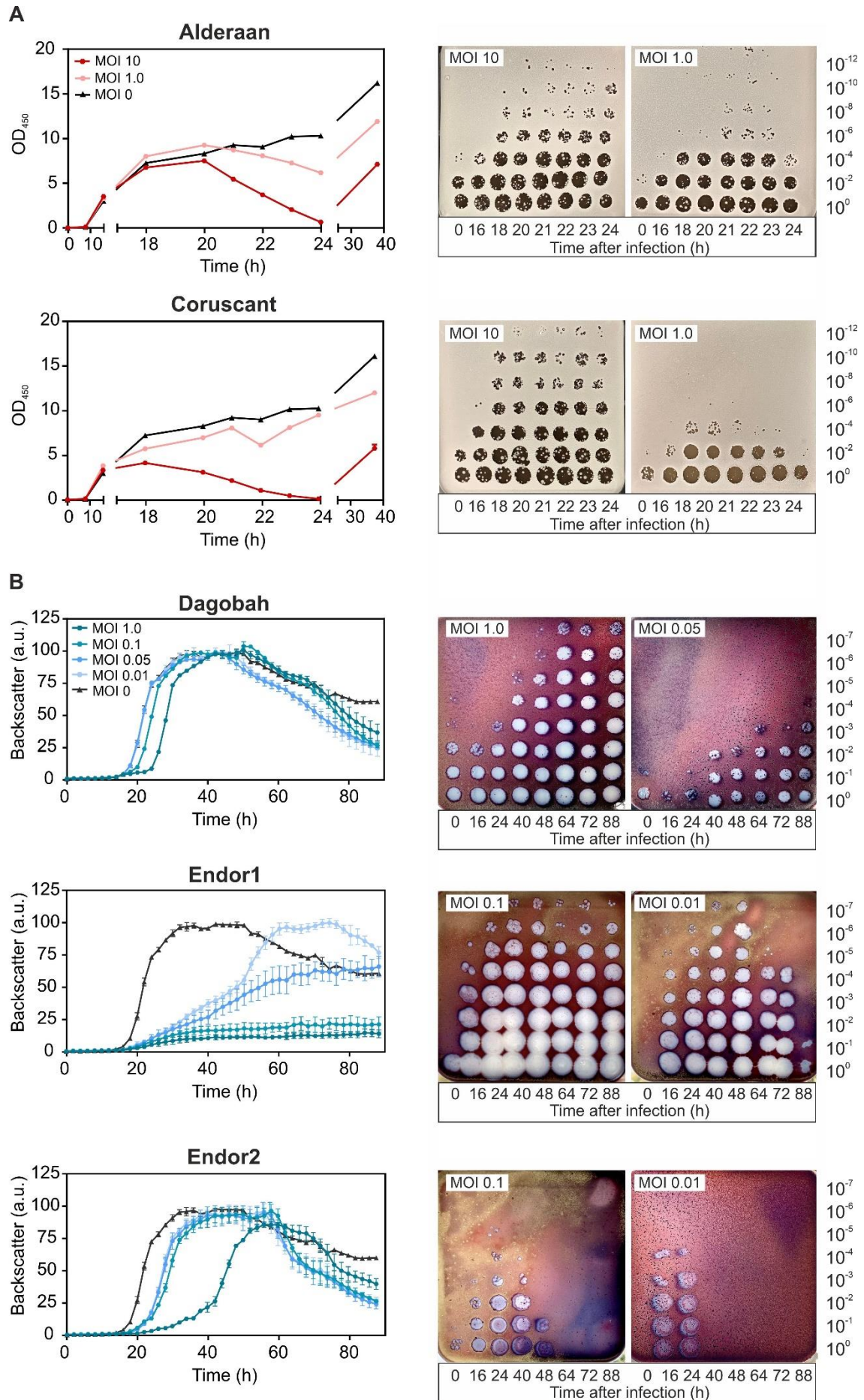
Figure 2. Infection curves of the five phages infecting *S. venezuelae* (a) and *S. coelicolor* (b). Spores of either *S. venezuelae* (10^5) or *S. coelicolor* M145 (10^6) were grown in GYM or YEME medium, respectively. After 6 to 8 hours, phages were added at the corresponding multiplicity of infection (MOI). OD_{450} or backscatter were measured over time (left panels), in parallel to phage titers (right panels).

Table 1. The host range of the five phages was assessed by spotting serial dilutions of these phages on lawns of different *Streptomyces* species propagated on GYM medium. The outcome of the spot assays is reported as follows: hosts used for phage isolation (grey), plaque formation (green), clearance of the bacterial lawn without visible plaques (yellow).

	Alderaan	Coruscant	Dagobah	Endor1	Endor2
S. venezuelae	grey	grey		green	green
S. coelicolor M600			grey	grey	grey
S. coelicolor M145			green	green	green
S. rimosus subsp. rimosus					
S. scabiei				yellow	
S. griseus					
S. platensis				yellow	yellow
S. xanthochromogenes				yellow	yellow
S. lividans	yellow		green		
S. olivaceus					green
S. cyaneofuscatus				green	green

While phages usually have a relatively narrow host range, some phages can sometimes infect many strains of the same species and even distinct species. We assessed the host-range of our phages by spotting them on lawns of different *Streptomyces* species (Table 1).

Interestingly, Endor1 and Endor2—but not Dagobah—formed plaques on *S. venezuelae*. *S. coelicolor* M145 showed the same sensitivity pattern than the M600 strain. M145 and M600 are both plasmid-free derivatives of A3(2) and mainly differ from each other in the length of their direct terminal repeats [17].

Beside *S. venezuelae* and *S. coelicolor*, *S. lividans* showed plaque formation by phage Dagobah. Endor1 and Endor2 also formed plaques on *S. olivaceus* and *S. cyanofuscatus*. Alderaan, Endor1 and Endor2 caused indefinite clearance of the bacterial lawn of several species, but higher dilutions did not reveal distinct, single plaques. For these species, the phage lysates could have inhibitory effects on growth or cause non-productive infection [36–37]. In summary, Endor1 and Endor2 showed the broadest host range, but overall, the five phages we isolated feature a relatively modest host range as they are only able to infect few other *Streptomyces* species.

3.3 Genome sequencing and genome features

293   All phages were sequenced using short-read technology (Illumina Mi-Seq). Each
294 genome could be assembled to a single contig, to which >80% of the reads could be
295 mapped confirming the purity of the samples.

296   The genome features of the five phages are summed up in Table 2. Briefly, they
297 show diverse genome sizes (39 to 133 kb), GC-contents (48 to 72%) and ORFs
298 numbers (51 to 290). The phage Coruscant differed from other phages, in that its
299 genome is significantly larger than the other phages and exhibits a markedly low
300 GC content (48%), in comparison to the one of its host (72%). The genomic ends were
301 predicted using PhageTerm, which detects biases in the number of reads to
302 determine DNA termini and phage packaging mechanisms [29]. Alderaan, Endor1
303 and Endor2 showed a headful packaging mechanism where the phage genomes
304 have a fixed start at the *pac* site, but the end of the genome is variable. In contrast,
305 phages Coruscant and Dagobah have direct terminal repeats (DTR). These DTR were
306 identified in the initial assembly by an approximately 2-fold increase in coverage
307 clearly delimitated at single base positions. Phage lifecycle was predicted using
308 PhageAI, which developed a lifecycle classifier based on machine learning and
309 natural language processing [30].

310

311   **Table 2. Basic genome features of the five phages.** Open reading frames (ORFs) were predicted using
312   Prokka [25] and were later manually curated. Protein domains encoded in ORFs were identified using
313   RPS-BLAST against the Conserved Domain Database (CDD). The type of genome ends was
314   determined using Phage Term [29]. The lifestyle of each phage was predicted by the machine-learning
315   based program PhageAI [30].

Phage name	Accession number	Reference host	Genome size (kb)	GC content (%)	ORF number	Genome termini class	Lifestyle prediction
Alderaan	MT711975	*Streptomyces venezuelae* ATCC 10712	39	72.1	51	Headful (*pac*)	Temperate
Coruscant	MT711976	*Streptomyces venezuelae* ATCC 10712	133 (12kb DTR)	48.4	290	DTR (long)	Virulent
Dagobah	MT711977	*Streptomyces coelicolor* M600	47 kb (1kb DTR)	68.9	93	DTR (short)	Temperate
Endor1	MT711978	*Streptomyces coelicolor* M600	49	65.8	75	Headful (*pac*)	Temperate
Endor2	MT711979	*Streptomyces coelicolor* M600	48	65.1	75	Headful (*pac*)	Temperate

316

317   Phage genes involved in the same function are usually clustered together,
318 forming functional modules (Figure 3) [38–39]. These modules fulfil the basic
319 functions necessary for production of progeny phages, including DNA/RNA
320 metabolism, DNA replication and repair, DNA packaging, virion structure and

321 assembly (tail and capsid), regulation, lysogeny (in the case of temperate phages)
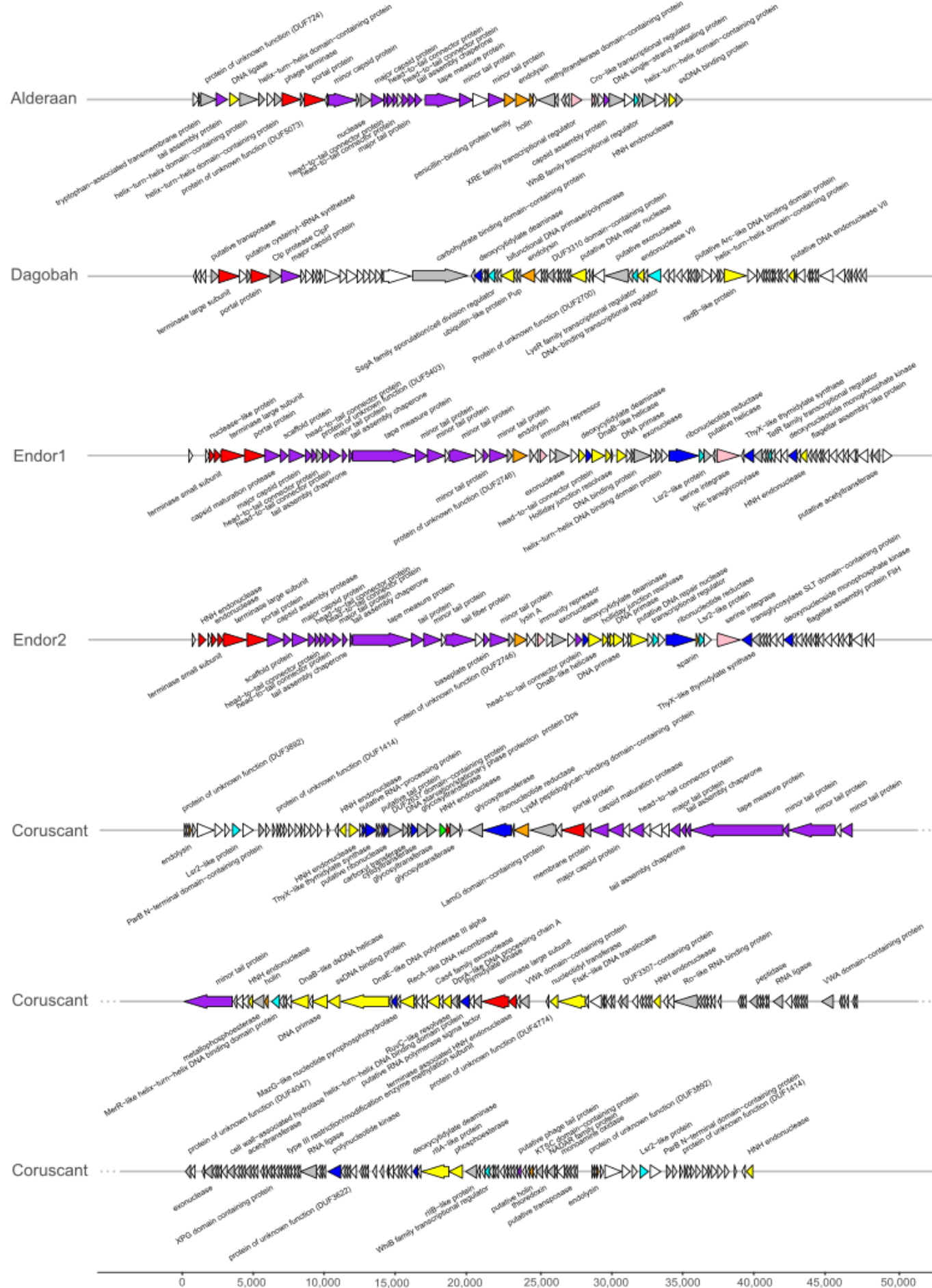322 and lysis.
323   Interestingly, Coruscant's large genomes is paralleled by a high genome
324 complexity. It contains no less than 41 copies of tRNAs, covering 19 different amino
325 acids - all standard amino acids except valine. Coruscant has also a relatively high
326 fraction of coding sequences for which no function could be predicted (155
327 hypothetical proteins out 290 CDS compared to 16/51 for Alderaan).
328   The phages were also found to encode homologs of bacterial regulators that are
329 typically used by *Streptomyces* to control sporulation and overall development. For
330 example, *whiB* (found in Alderaan, and Coruscant) and *ssgA* (found in Dagobah) are
331 both essential for sporulation of *Streptomyces* [40–41]. Three phages (Coruscant,
332 Endor1 and Endor2) also encode Lsr2-like proteins, which are nucleoid-associated
333 proteins functioning as xenogeneic silencing proteins and are conserved throughout
334 Actinobacteria [42].

335

337 **Figure 3. Genome map of the five Streptomyces phages**. Open reading frames (ORFs) were
338 identified with Prodigal and functionally annotated using an automatic pipeline based on Prokka
339 [25]. The functional annotation was automatically improved and curated using hidden Markov
340 models (HMMs), and Blastp searches [26] against different databases (Prokaryotic Virus Orthologous
341 Groups (pVOGs) [27], viral proteins and Conserved Domain Database (CDD)) [28]. Genome maps
342 were created using the R package gggenes.

343 *3.4 Average nucleotide identity (ANI) analysis*

344 We established the sequence relationship between the newly sequenced
345 *Streptomyces* phages and the selected genomes from the representative group
346 members of actinophages. The Average nucleotide identity (ANI) based clustering
347 dendrogram analysis showed four (Endor1, Endor2, Alderaan, and Coruscant) out
348 of five phage genomes clustered confidently with the members of already known
349 clusters (Endor1/Endor2: BD, Coruscant: BE, and Alderaan: BC) (Figure 4).
350 However, one of the phage genomes (Dagobah) does not share sufficient similarity
351 and was therefore clustered as an unresolved group. As expected, the overall
352 analysis showed that except Dagobah, all four phages show close relatedness to
353 *Streptomyces* phages.

354
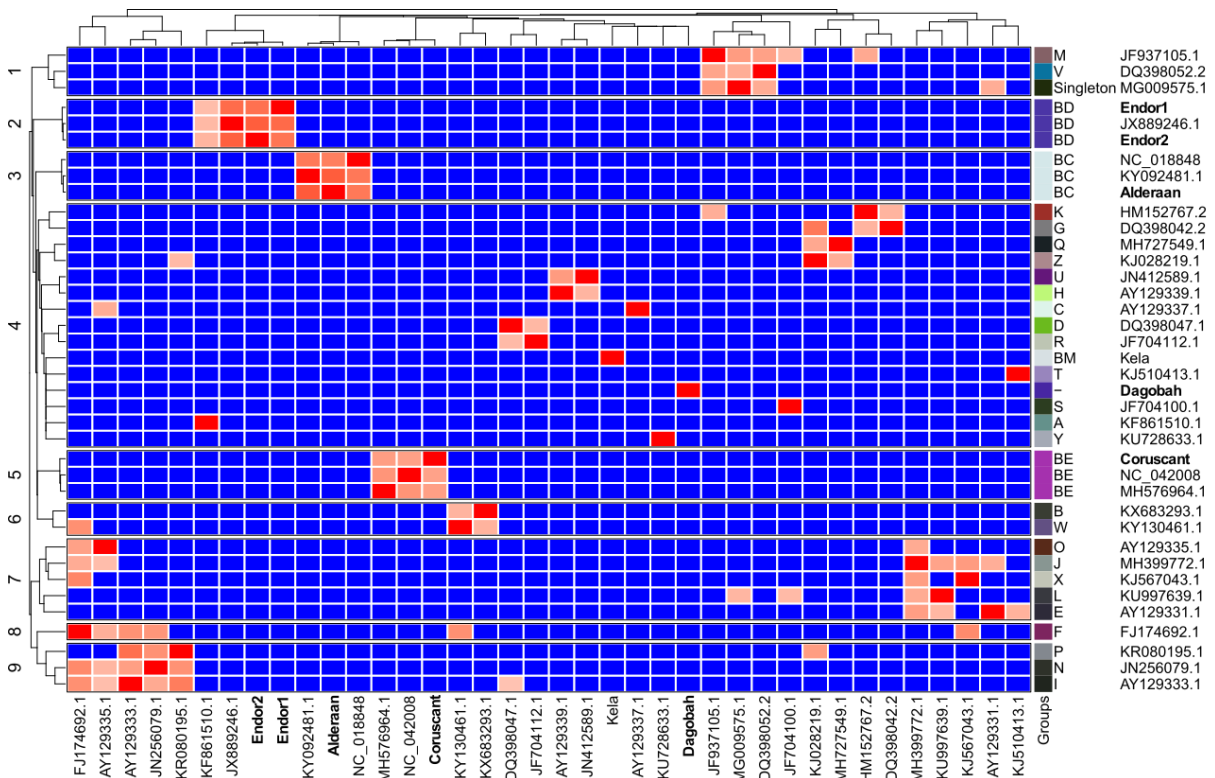
356 **Figure 4. Average nucleotide-based dendrogram analysis using 38 actinophage genomes.** These 38
357 genomes include 31 genomes downloaded from the Actinophage Database (https://phagesdb.org/),
358 two genomes from NCBI based on close relatedness, and the five newly sequenced phages. The group
359 of each phage, as defined by the Actinophage Database, is indicated.

360

361

362   *3.5 Protein domain-based analysis*

363   Sequence relationship between the phage genomes is most commonly
364   determined with the help of genome-wide similarity or average nucleotide identity-
365   based analysis. However, a traditional method such as phylogeny with single genes
366   is challenging because of the high variability and lack of universal genes across the
367   phage genomes. Thus, we used additional phyletic-based analysis to establish a
368   sequence relationship between the phage genomes. The hierarchical clustering
369   dendrogram based on the identified 703 Pfam domains presence-absence matrix
370   confidently clusters newly sequenced phages with known actinophages (Figure 5).
371   In comparison to ANI-based analysis, hierarchical clustering showed congruent
372   topology for the four newly sequenced *Streptomyces* phage genomes (Endor1 and
373   Endor2: BD cluster, Alderaan: BC cluster, and Coruscant: BE cluster)
374   (Supplementary Figures S2-S5). It also resolved polytomy between the unresolved
375   groups and showed that Dagobah comes under the singleton group, consisting of
376   highly divergent phages. Moreover, a high level of congruence was observed
377   between already known groups and the groups identified by our hierarchical
378   clustering. Thus, our results strongly suggest domain-based phyletic or hierarchical
379   clustering analysis as an alternate way of classification of newly sequenced phage
380   genomes.
381
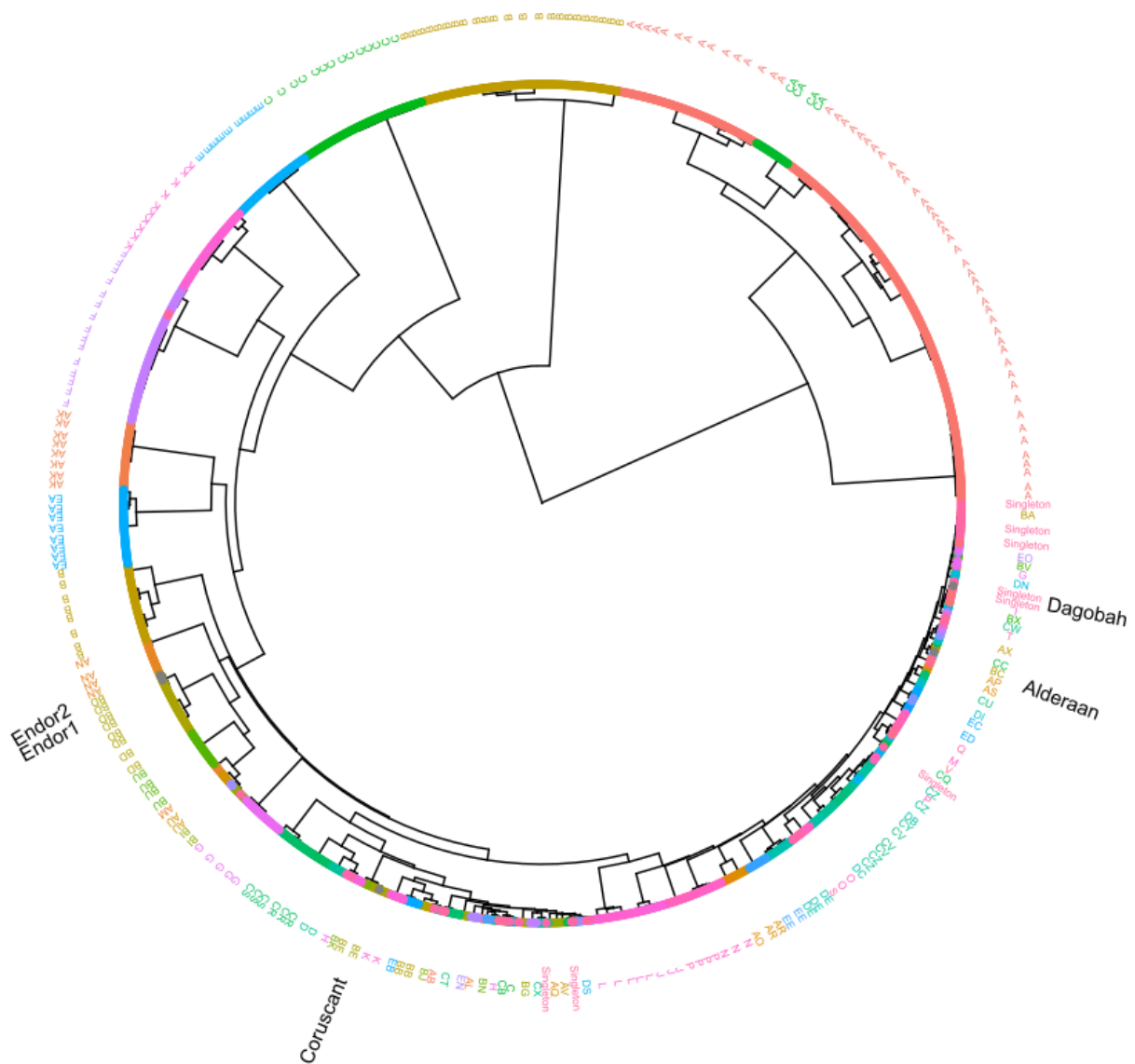
Figure 5. Protein domain-based hierarchical clustering. The dendrogram was constructed based on the presence-absence matrix of the > 700 Pfam protein domains identified from 2486 actinophage genomes. Phages are color-coded according to known groups from the Actinobacteriophage Database (https://phagesdb.org/) [31]. The position of the five new phage genomes is indicated as black text.

4. Discussion

In this study, we report the isolation and characterization of five novel *Streptomyces* phages. Alderaan and Coruscant were isolated using *S. venezueale*, while *S. coelicolor* was the host used for isolation of Dagobah, Endor1 and Endor2.

Four for of them (Alderaan, Dagobah, Endor1 and Endor2) were predicted to be temperate phages. In contrast, Coruscant was predicted to feature a lytic lifestyle, and exhibits a large genome (superior to 130 kb) with a low GC content (48%), in comparison to the 72% of its *Streptomyces* host. Coruscant also encodes 41 copies of

399 tRNA genes, spanning 19 of the 20 standard amino acids. This large tRNA gene
400 repertoire could be used to optimize gene expression in hosts that have differing
401 codon usage patterns or to counteract potential tRNA-based degradation defense
402 systems [43]. Altogether, the combination of a low GC content and a substantial
403 tRNA equipment suggests a recent adaptation of the phage Coruscant to
404 *Streptomyces*.

405   ANI and hierarchical clustering analysis revealed that Alderaan, Coruscant and
406 Endor1/Endor2 belong to clusters BC, BE and BD defined by PhagesDB [31],
407 respectively. In contrast, Dagobah showed very little homology with described
408 phages, and was thus considered as a singleton. This finding highlights the largely
409 untapped phage diversity, making the isolation of entirely "novel" phages still
410 possible.

411   *Streptomyces* are characterized by their complex lifestyle and cellular
412 differentiation. Interestingly, the isolated actinophages also encode homologs of
413 SsgA, WhiB and Lsr2 proteins—regulatory proteins typically encoded by their hosts.
414 The *ssgA* gene product was previously shown to be necessary for proper sporulation
415 of *Streptomyces coelicolor* [41]; *whiB* is also essential for sporulation of *Streptomyces*
416 and was already reported to be found in several actinophages [44-46]. Interestingly,
417 the WhiB-like protein of mycobacteriophage TM4, WhiB$_{TM4}$, was shown to inhibit
418 the transcription of Mycobacterium *whiB2*. Expression of WhiB$_{TM4}$ in *M. smegmatis*
419 led to hindered septation resembling a WhiB2 knockout phenotype, highlighting
420 how phage can interfere with their host's development [45].

421   Lsr2-like proteins are nucleoproteins conserved in actinobacteria. In
422 *Streptomyces*, they were recently shown to silence cryptic specialized metabolic
423 clusters [47]. The first example of a phage-encoded Lsr2-like protein is the prophage-
424 encoded Lsr2-like protein CgpS in *Corynebacterium glutamicum*. CgpS was shown to
425 maintain the lysogenic state of the prophage on which it resides. Further
426 bioinformatic searches revealed that Lsr2-like proteins are abundant in
427 actinophages, with almost 20% of *Streptomyces* phages encoding such proteins [42].
428 However, their role in the coordination of the phage life cycle remains still unclear.
429 Altogether, these observations suggest that phages manipulate their host
430 development, by interfering with central processes such as sporulation and
431 antibiotic production.

432   More generally, the specificities of *Streptomyces* – especially its morphological
433 complexity – impact the phage isolation and characterization process. For example,
434 the mycelial nature of streptomycetes complicates quantitative studies. The notion
435 of MOI loses a lot of its significance once mycelium has formed, as the network
436 structure originating from one spore has greatly increased phage adsorption but
437 would still be counted as one CFU [14,48]. Furthermore, the formation of clumps,
438 although mitigated by the addition of glass beads or increase of osmotic pressure
439 [49], makes accurate monitoring of cell growth (based on optical density or
440 backscatter) difficult.

441   *S. coelicolor* was established as a model system for the *Streptomyces* genus partly
442 because of its prolific pigment production [50]. Interestingly, we observed colored

443   halos around the plaques formed by the *S. coelicolor* phages. Exposure to ammonia
444   fume confirmed that these colored halos contain actinorhodin. This observation
445   suggests that *Streptomyces* release metabolites in reaction to phage predation, some
446   of which may potentially have anti-phage properties as it was shown recently with
447   anthracyclines in *Streptomyces peucetius* [51].

448   Understanding the processes governing phage infection has the potential to
449   illuminate the basic physiology of their hosts. Therefore, phages can serve as a basis
450   to study *Streptomyces*' specific traits—its complex reproduction cycle and abundant
451   production of secondary metabolites—in the context of phage infection.

452

453

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: Close-ups of phage plaques imaged using a Nikon SMZ18 stereomicroscope, before (upper row) and after (lower row) exposure to ammonia fumes. Figure S2: Subclade dendrogram with *Streptomyces* phage Alderaan and its closely related actinophages. Figure S3: Subclade dendrogram with *Streptomyces* phage Coruscant and its closely related actinophages. Figure S4: Subclade dendrogram with *Streptomyces* phage Dagobah and its closely related actinophages. Figure S5: Subclade dendrogram with *Streptomyces* phages Endor1 and Endor2 and their closely related actinophages

Supplementary Table 1: List of the functional annotation of proteins ORFs within phage genomes

Author Contributions: Conceptualization, A.H., and J.F.; methodology, All; validation, All; formal analysis, All; investigation, A.H. and L.K.; resources, V.S. and J.F.; data curation, A.H. and V.S.; writing—original draft preparation, A.H. and V.S.; writing—review and editing, All; visualization, A.H.,V.S. and L.K.; supervision, J.F.; project administration, J.F.; funding acquisition, J.F. All authors have read and agreed to the published version of the manuscript

Conflicts of Interest: The authors declare no conflict of interest.

References

491   1. M. J. Bibb, « Understanding and manipulating antibiotic production in actinomycetes », *Biochem Soc Trans*,
493   vol. 41, nᵒ 6, p. 1355-1364, 2013, doi: 10.1042/BST20130214.

494   2. D. A. Hopwood, *Streptomyces in Nature and Medicine: The Antibiotic Makers*. Oxford, New York: Oxford
495   University Press, 2007.

496   3. T. Keiser, M. J. Bibb, M. J. Buttner, K. F. Chater, et D. A. Hopwood, *Practical Streptomyces Genetics*. Norwich:
497   The John Innes Foundation, 2000.

498   4. M. A. Elliot, M. J. Buttner, J. R. Nodwell, « 24 Multicellular Development in *Streptomyces* », *Myxobacteria*, p.
499   419-438, 2008, doi: 10.1128/9781555815677.ch24.

500   5. J. R. McCormick, K. Flärdh, « Signals and regulators that govern *Streptomyces* development », *FEMS
501   Microbiol Rev*, vol. 36, nᵒ 1, p. 206-231, 2012, doi: 10.1111/j.1574-6976.2011.00317.x.

502   6. J. Anne, W. Wohlleben, H. J. Burkardt, R. Springer, A. Pohler, « Morphological and Molecular
503   Characterization of Several Actinophages Isolated from Soil Which Lyse *Streptomyces cattleya* or *S.
504   venezuelae* », *Microbiology*, vol. 130, nᵒ 10, p. 2639-2649, 1984, doi: 10.1099/00221287-130-10-2639.

505   7. S. Donadio, R. Paladino, I. Costanzi, P. Sparapani, W. Schreil, M. Iaccarino, « Characterization of
506   bacteriophages infecting *Streptomyces erythreus* and properties of phage-resistant mutants. », *J Bacteriol*, vol.
507   166, nᵒ 3, p. 1055-1060, 1986.

508   8. J. E. Dowding, « Characterization of a Bacteriophage Virulent for *Streptomyces coelicolor* A3 (2) |
509   Microbiology Society », *Journal of General Microbiology*, vol. 76, p. 163-176, 1973.

510   9. N. D. Lomovskaya, N. M. Mkrtumian, N. L. Gostimskaya, V. N. Danilenko, « Characterization of
511   Temperate Actinophage oC31 Isolated from *Streptomyces coelicolor* A3(2) », *J. VIROL.*, vol. 9, p. 5, 1972.

512   10. M. C. M. Smith, R. N. Burns, S. E. Wilson, M. A. Gregory, « The complete genome sequence of the
513   Streptomyces temperate phage φC31: evolutionary relationships to other viruses », *Nucleic Acids Res*, vol.
514   27, nᵒ 10, p. 2145-2155, 1999, doi: 10.1093/nar/27.10.2145.

515   11. M. C. M. Smith *et al.*, « Evolutionary Relationships among Actinophages and a Putative Adaptation for
516   Growth in *Streptomyces* spp. », *Journal of Bacteriology*, vol. 195, nᵒ 21, p. 4924-4935, 2013, doi:
517   10.1128/JB.00618-13.

518   12. J. Burke, D. Schneider, J. Westpheling, « Generalized transduction in *Streptomyces coelicolor* », *PNAS*, vol.
519   98, nᵒ 11, p. 6289-6294, 2001, doi: 10.1073/pnas.101589398.

520   13. N. D. Lomovskaya, K. F. Chater, N. M. Mkrtumian, « Genetics and molecular biology of *Streptomyces
521   bacteriophages.* », *Microbiology and Molecular Biology Reviews*, vol. 44, nᵒ 2, p. 206-229, 1980.

522   14. A. Rosner, R. Gustein, « Adsorption of actinophage Pal 6 to developing mycelium of *Streptomyces* », 1980.

523   15. T. C. Jordan *et al.*, « A Broadly Implementable Research Course in Phage Discovery and Genomics for First-
524   Year Undergraduate Students », *mBio*, vol. 5, nᵒ 1, 2014, doi: 10.1128/mBio.01051-13.

525   16. S. T. Pullan, G. Chandra, M. J. Bibb, M. Merrick, « Genome-wide analysis of the role of GlnR in *Streptomyces
526   venezuelae* provides new insights into global nitrogen regulation in actinomycetes », *BMC Genomics*, vol. 12,
527   p. 175, 2011, doi: 10.1186/1471-2164-12-175.

528   17. D. Weaver *et al.*, « Genome plasticity in *Streptomyces*: identification of 1 Mb TIRs in the *S. coelicolor* A3(2)
529   chromosome: Identification of 1 Mb TIRs in *S. coelicolor* A3(2) », *Molecular Microbiology*, vol. 51, nᵒ 6, p.
530   1535-1550, 2004, doi: 10.1111/j.1365-2958.2003.03920.x.

531   18. S. D. Bentley *et al.*, « Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2) »,
532   *Nature*, vol. 417, nᵒ 6885, p. 141-147, 2002, doi: 10.1038/417141a.

533   19. M. D. Shepherd, M. K. Kharel, M. A. Bosserman, J. Rohr, « Laboratory Maintenance of *Streptomyces
534   species* », *Curr Protoc Microbiol*, vol. CHAPTER, p. Unit-10E.1, 2010, doi:
535   10.1002/9780471729259.mc10e01s18.

536   20. K. M. Kauffman M. F. Polz, « Streamlining standard bacteriophage methods for higher throughput »,
537   *MethodsX*, vol. 5, p. 159-172, 2018, doi: 10.1016/j.mex.2018.01.007.

538   21. F. Kensy, E. Zang, C. Faulhammer, R.-K. Tan, J. Büchs, « Validation of a high-throughput fermentation
539   system based on online monitoring of biomass and fluorescence in continuously shaken microtiter plates »,
540   *Microb. Cell Fact.*, vol. 8, p. 31, 2009, doi: 10.1186/1475-2859-8-31.

541   22. C. Rückert *et al.*, « Complete genome sequence of *Streptomyces lividans* TK24 », *J. Biotechnol.*, vol. 199, p.
542   21-22, 2015, doi: 10.1016/j.jbiotec.2015.02.004.

543   23. D. Gordon P. Green, « Consed: a graphical editor for next-generation sequencing », *Bioinformatics*, vol. 29,
544   nᵒ 22, p. 2936-2937, 2013, doi: 10.1093/bioinformatics/btt515.

545 24. D. Hyatt, G.-L. Chen, P. F. LoCascio, M. L. Land, F. W. Larimer, L. J. Hauser, « Prodigal: prokaryotic gene
546   recognition and translation initiation site identification », *BMC Bioinformatics*, vol. 11, n° 1, p. 119, 2010, doi:
547   10.1186/1471-2105-11-119.
548 25. T. Seemann, « Prokka: rapid prokaryotic genome annotation », *Bioinformatics*, vol. 30, n° 14, p. 2068-2069,
549   2014, doi: 10.1093/bioinformatics/btu153.
550 26. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, « Basic local alignment search tool », *Journal
551   of Molecular Biology*, vol. 215, n° 3, p. 403-410, 1990, doi: 10.1016/S0022-2836(05)80360-2.
552 27. A. L. Grazziotin, E. V. Koonin, D. M. Kristensen, « Prokaryotic Virus Orthologous Groups (pVOGs): a
553   resource for comparative genomics and protein family annotation », *Nucleic Acids Res*, vol. 45, n° D1, p.
554   D491-D498, 2017, doi: 10.1093/nar/gkw975.
555 28. A. Marchler-Bauer *et al.*, « CDD: NCBI's conserved domain database », *Nucleic Acids Res*, vol. 43, n° D1, p.
556   D222-D226,. 2015, doi: 10.1093/nar/gku1221.
557 29. J. R. Garneau, F. Depardieu, L.-C. Fortier, D. Bikard, M. Monot, « PhageTerm: a tool for fast and accurate
558   determination of phage termini and packaging mechanism using next-generation sequencing data »,
559   *Scientific Reports*, vol. 7, n° 1, p. 8292, 2017, doi: 10.1038/s41598-017-07910-5.
560 30. P. Tynecki, A. Guziński, J. Kazimierczak, M. Jadczuk, J. Dastych, A. Onisko, « PhageAI - Bacteriophage Life
561   Cycle Recognition with Machine Learning and Natural Language Processing », Bioinformatics, preprint,
562   2020. doi: 10.1101/2020.07.11.198606.
563 31. D. A. Russell G. F. Hatfull, « PhagesDB: the actinobacteriophage database », *Bioinformatics*, vol. 33, n° 5, p.
564   784-786, 2017, doi: 10.1093/bioinformatics/btw711.
565 32. L. Pritchard, *widdowquinn/pyani*. 2020.
566 33. Z. Gu, R. Eils, M. Schlesner, « Complex heatmaps reveal patterns and correlations in multidimensional
567   genomic data », *Bioinformatics*, vol. 32, n° 18, p. 2847-2849, 2016, doi: 10.1093/bioinformatics/btw313.
568 34. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T.-Y. Lam, « ggtree: an r package for visualization and annotation
569   of phylogenetic trees with their covariates and other associated data », *Methods in Ecology and Evolution*,
570   vol. 8, n° 1, p. 28-36, 2017, doi: 10.1111/2041-210X.12628.
571 35. B. A. M. Rudd, D. A. Hopwood, « Genetics of Actinorhodin Biosynthesis by *Streptomyces coelicolor* A3(2) »,
572   *Microbiology*, vol. 114, n° 1, p. 35-43, 1979, doi: 10.1099/00221287-114-1-35.
573 36. S. T. Abedon, « Lysis from without », *Bacteriophage*, vol. 1, n° 1, p. 46-49, 2011, doi: 10.4161/bact.1.1.13980.
574 37. S. T. Abedon, « Detection of Bacteriophages: Phage Plaques », in *Bacteriophages: Biology, Technology, Therapy*,
575   D. R. Harper, S. T. Abedon, B. H. Burrowes, M. L. McConville, Éd. Cham: Springer International Publishing,
576   2018, p. 1-32.
577 38. D. Botstein, « A Theory of Modular Evolution for Bacteriophages* », *Annals of the New York Academy of
578   Sciences*, vol. 354, n° 1, p. 484-491, 1980, doi: 10.1111/j.1749-6632.1980.tb27987.x.
579 39. H. Brüssow, F. Desiere, « Comparative phage genomics and the evolution of *Siphoviridae*: insights from
580   dairy phages », *Molecular Microbiology*, vol. 39, n° 2, p. 213-223, 2001, doi: 10.1046/j.1365-2958.2001.02228.x.
581 40. V. Molle, W. J. Palframan, K. C. Findlay, M. J. Buttner, « WhiD and WhiB, homologous proteins required
582   for different stages of sporulation in *Streptomyces coelicolor* A3(2) », *J. Bacteriol.*, vol. 182, n° 5, p. 1286-1295,
583   2000, doi: 10.1128/jb.182.5.1286-1295.2000.
584 41. G. P. van Wezel, J. van der Meulen, S. Kawamoto, R. G. M. Luiten, H. K. Koerten, B. Kraal, « *ssgA* Is Essential
585   for Sporulation of *Streptomyces coelicolor* A3(2) and Affects Hyphal Development by Stimulating Septum
586   Formation », *Journal of Bacteriology*, vol. 182, n° 20, p. 5653-5662, 2000, doi: 10.1128/JB.182.20.5653-5662.2000.
587 42. E. Pfeifer, M. Hünnefeld, O. Popa, J. Frunzke, « Impact of Xenogeneic Silencing on Phage–Host
588   Interactions », *Journal of Molecular Biology*, 2019, doi: 10.1016/j.jmb.2019.02.011.
589 43. P. Hyman, S. T. Abedon, « Chapter 7 - Bacteriophage Host Range and Bacterial Resistance », in *Advances in
590   Applied Microbiology*, vol. 70, Academic Press, 2010, p. 217-248.
591 44. P. Morris, L. J. Marinelli, D. Jacobs-Sera, R. W. Hendrix, G. F. Hatfull, « Genomic Characterization of
592   Mycobacteriophage Giles: Evidence for Phage Acquisition of Host DNA by Illegitimate Recombination »,
593   *J Bacteriol*, vol. 190, n° 6, p. 2172-2182, 2008, doi: 10.1128/JB.01657-07.
594 45. J. Rybniker *et al.*, « Insights into the function of the WhiB-like protein of mycobacteriophage TM4--a
595   transcriptional inhibitor of WhiB2 », *Mol. Microbiol.*, vol. 77, n° 3, p. 642-657, 2010, doi: 10.1111/j.1365-
596   2958.2010.07235.x.
597 46. W. Van Dessel *et al.*, « Complete genomic nucleotide sequence and analysis of the temperate bacteriophage
598   VWB », *Virology*, vol. 331, n° 2, p. 325-337, 2005, doi: 10.1016/j.virol.2004.10.028.

599 47. E. J. Gehrke *et al.*, « Silencing cryptic specialized metabolism in *Streptomyces* by the nucleoid-associated
600 protein Lsr2 », *eLife*, vol. 8, p. e47691, 2019, doi: 10.7554/eLife.47691.
601 48. C. M. Gilmour, E. C. Noller, B. Watkins, « STUDIES ON *STREPTOMYCES* PHAGE: I. Growth
602 Characteristics of the *Streptomyces griseus* Host-Phage System », *Journal of Bacteriology*, vol. 78, nº 2, p.
603 186-192, 1959.
604 49. L. D. Nguyen *et al.*, « Cultivation System Using Glass Beads Immersed in Liquid Medium Facilitates Studies
605 of *Streptomyces* Differentiation », *Appl. Environ. Microbiol.*, vol. 71, nº 6, p. 2848-2852, 2005, doi:
606 10.1128/AEM.71.6.2848-2852.2005.
607 50. K. Chater, « David Hopwood and the emergence of *Streptomyces* genetics », *Int. Microbiol.*, vol. 2, nº 2, p.
608 61-68, 1999.
609 51. S. Kronheim *et al.*, « A chemical defence against phage infection », *Nature*, vol. 564, nº 7735, p. 283, 2018,
610 doi: 10.1038/s41586-018-0767-x.

611