



OPEN

Complete genomes and comparative analyses of *Streptomyces* phages that influence secondary metabolism and sporulation

Sarah Kronheim¹, Ethan Solomon¹, Louis Ho¹, Michelle Glossop¹, Alan R. Davidson^{1,2} & Karen L. Maxwell¹✉

Bacteria in the genus *Streptomyces* are found ubiquitously in nature and are known for the number and diversity of specialized metabolites they produce, as well as their complex developmental lifecycle. Studies of the viruses that prey on *Streptomyces*, known as phages, have aided the development of tools for genetic manipulation of these bacteria, as well as contributing to a deeper understanding of *Streptomyces* and their behaviours in the environment. Here, we present the genomic and biological characterization of twelve *Streptomyces* phages. Genome analyses reveal that these phages are closely related genetically, while experimental approaches show that they have broad overlapping host ranges, infect early in the *Streptomyces* lifecycle, and induce secondary metabolite production and sporulation in some *Streptomyces* species. This work expands the group of characterized *Streptomyces* phages and improves our understanding of *Streptomyces* phage-host dynamics.

Streptomyces are Gram-positive bacteria that are ubiquitous in nature. These bacteria live primarily in soil but are also found in diverse habitats including oceans and deserts¹. *Streptomyces* are well known for their prodigious production of secondary metabolites, with individual strains having the capacity to produce an estimated 20–40 different compounds^{2–4}. These metabolites are not required for growth of the bacteria under laboratory conditions but are thought to provide fitness advantage in the natural environment. Many secondary metabolites inhibit the growth of competing species in the environment, including bacteria, fungi, and parasites^{5,6}. Recent studies have also shown a role for secondary metabolites in defence against bacterial viruses known as bacteriophages (phages)^{7–10}.

Phages are widespread in nature, occupying all environments where bacteria are found. They are thought to outnumber their bacterial hosts by a factor of ten and their predation of bacteria plays important roles in shaping microbial communities^{11,12}. *Streptomyces* phages were isolated as far back as the 1970s, and phages such as ϕ C31, ϕ BT1, and SV1 have been used to develop important genetic tools that can be used to manipulate *Streptomyces*^{13,14}. Both phage infection and secondary metabolite production are linked to the *Streptomyces* lifecycle, and certain secondary metabolites have been shown to inhibit phage replication^{7–10,15–17}. However, the relationship between *Streptomyces* secondary metabolism and phage infection has not been well characterized.

In this work we report the characterization and genomic analysis of twelve *Streptomyces* phages. Comparative analyses show that these phages, isolated from diverse geographic locations, share closely related genome sequences and cluster with the previously characterized *Streptomyces* phages R4 and ϕ Joe. These phages display broad host ranges and can efficiently replicate in many *Streptomyces* species. Whole genome analyses of these phages show a variable accessory genome that likely contributes to host range and the differing abilities of these phages to induce secondary metabolite production and early sporulation.

¹Department of Biochemistry, University of Toronto, 661 University Avenue, Suite 1600, Toronto, ON M5G 1M1, Canada. ²Department of Molecular Genetics, University of Toronto, 661 University Avenue, Suite 1600, Toronto, ON M5G 1M1, Canada. ✉email: karen.maxwell@utoronto.ca

Results

Streptomyces phage sequencing and genome comparisons. Previous investigations in our lab characterizing chemical anti-phage defence⁷ used a collection of *Streptomyces coelicolor* phages (ϕ Scoe phages) isolated from geographically diverse soil samples. Negatively stained electron microscopy images showed that these phages belong to the order *Caudovirales*, with double-strand DNA genomes contained inside icosahedral heads and long, non-contractile tails. To gain insight into this phage collection, we determined their genome sequences using Illumina NextSeq 2000 short-read technology. Each genome was assembled into a single contig, with >86% of the reads mapped to each sequence. The genomes were annotated using automated gene predictions followed by careful manual inspection using our in-house phage database for functional annotations.

DNA-DNA comparisons performed using Pyani¹⁸ revealed that the twelve ϕ Scoe phage genomes shared pairwise nucleotide identities ranging from 73 to 91% across the complete genome sequences (Fig. 1, Supplementary Fig. S1a). The ϕ Scoe phage genome lengths varied in size from 48 to 51 kb, and each was found to contain between 75 and 84 predicted open reading frames (ORFs) (Table 1). Of these ORFs, 51 were found to be conserved among all twelve phages (Fig. 2, Supplementary Fig. S1b, Supplementary Table S1). The genome organization of these phages is similar to previously characterized *Streptomyces* phages, encoding the functional modules necessary for Siphophage particle formation (DNA packaging, capsid and tail assembly), cell lysis, DNA replication, as well as an immunity repressor and serine integrase that suggest a temperate lifestyle for these phages (Supplementary Table S2). Genome comparisons with previously characterized phages revealed striking similarity to *Streptomyces* phages R4¹⁹ and ϕ Joe²⁰, with average pairwise nucleotide identities of 71–78% and 73–89% respectively (Fig. 1, Supplementary Fig. S1). The R4-like group of phages appears to be the most abundant group of *Streptomyces* phages²¹, falling within the BD family at the Actinobacteriophage Database²².

Conserved *Streptomyces* phage proteins. Both R4 and ϕ Joe are temperate phages, which means that they can pursue either a lytic or lysogenic lifestyle following phage infection. Temperate phages require an immunity repressor to control the balance between lysis and lysogeny. They do so by downregulating promoters required for the lytic pathway. Each of the ϕ Scoe phages was found to encode a repressor protein between the head–tail–lysis genes and the putative early region genes. These repressor proteins share sequence identities of 64–90% with the repressor proteins found in *Streptomyces* phages R4 and ϕ Joe (Supplementary Fig. S2). Phage R4 has an atypical regulatory mechanism that silences transcription from the prophage, known as the repressor-

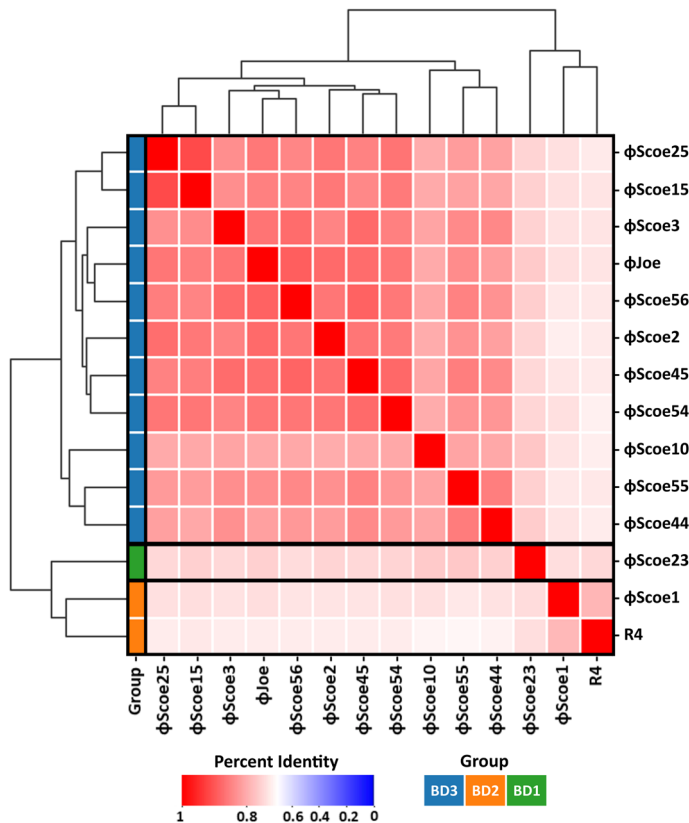


Figure 1. Heatmap of the average nucleotide identity (ANI) values for the twelve newly sequenced phage genomes with previously reported *Streptomyces* phages ϕ Joe and R4. The shading represents the percent sequence identity shared by the noted phages. The Actinophage Database BD subgroup each phage belongs to is noted, and phages in the same subgroup are boxed.

Phage	Genome size (kb)	GC content (%)	Predicted ORFs	Lifestyle	Genbank accession
φScoe1	51	66.83	80	Temperate	OQ672741
φScoe2	49.4	65.77	79	Temperate	OQ672744
φScoe3	48.3	65.6	81	Temperate	OQ672747
φScoe10	51.1	65.23	80	Temperate	OQ672742
φScoe15	49.4	66.07	80	Temperate	OQ672743
φScoe23	49.8	65.84	75	Temperate	OQ672745
φScoe25	49.4	66.13	84	Temperate	OQ672746
φScoe44	48.8	65.08	80	Temperate	OQ672748
φScoe45	48.8	65.57	80	Temperate	OQ672749
φScoe54	48.7	65.83	78	Temperate	OQ672750
φScoe55	48.5	65.17	77	Temperate	OQ672751
φScoe56	48.4	65.47	78	Temperate	OQ672752

Table 1. Phage genome features.

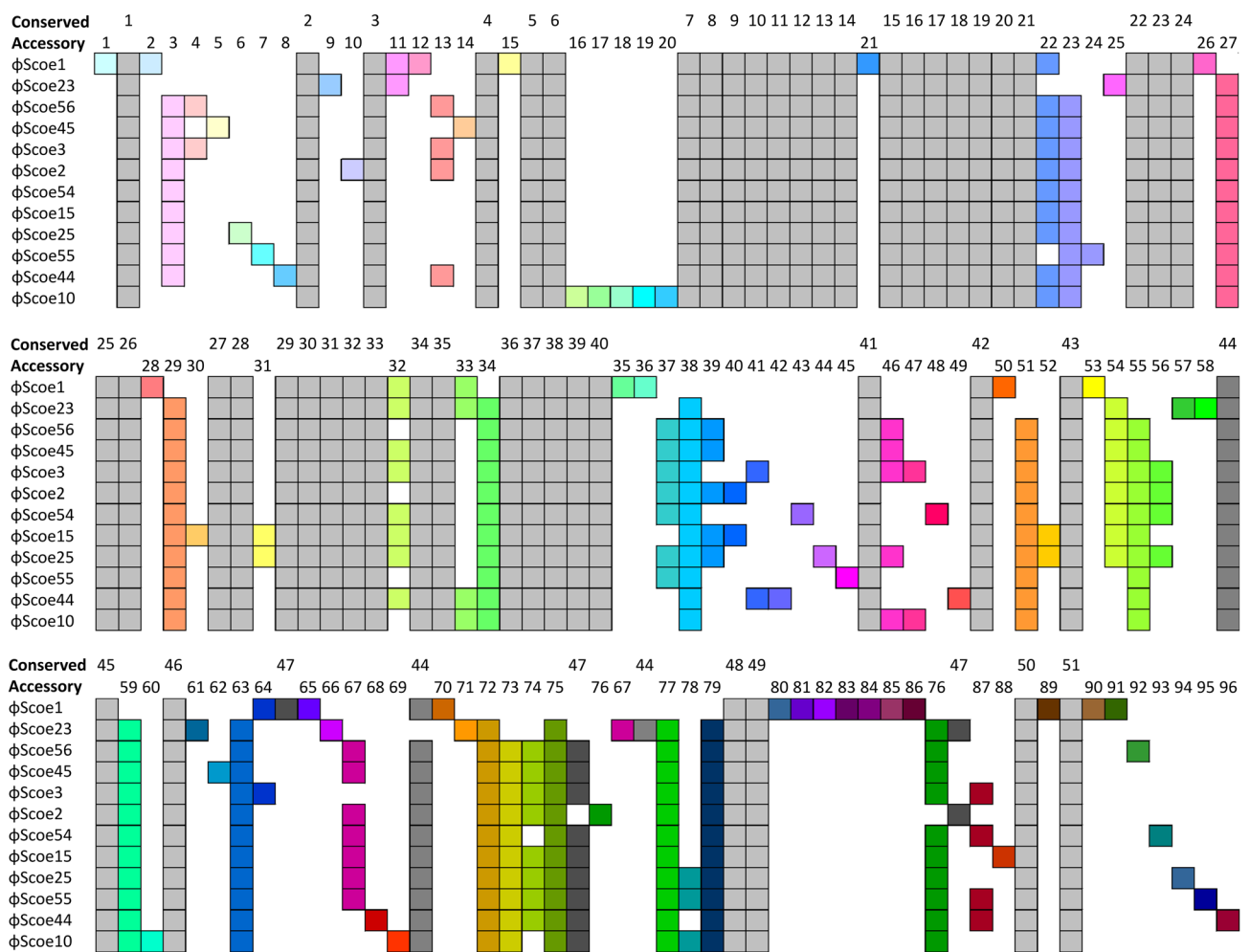


Figure 2. Conserved and accessory gene map. Alignment of the phage genomes shows that these phages are closely related. Each box represents a single open reading frame; conserved genes shown in grey, and the non-conserved “accessory genes” are shown in colour. Different vertical positions and colours for the accessory genes indicate that they encode distinct protein families. Darker grey boxes indicate genes that are conserved but appear in different places in these genomes. Genes are grouped if they share > 30% identity or share a specific confirmed biological function (e.g., receptor binding proteins are grouped by sequence identity while “serine integrase” is a group despite high sequence diversity). Genomes are organized by average nucleotide identity. See Supplementary Table S1 for details about specific gene annotations.

stoperator regulatory system²¹. This type of repressor regulates transcription initiation at an early lytic promoter, but also acts at many sites across the phage genome, blocking the movement of RNA polymerase and thereby helping to maintain lysogeny. Bioinformatic analyses revealed eight or more R4 stoperator binding sites in intergenic regions of the ϕ Scoe phages (Supplementary Table S3), suggesting that this mechanism of repression is conserved across this family of phages.

Most temperate phages encode an integrase protein whose activity mediates integration and excision of the prophage during the lysis-lysogeny decision. The integrases of many *Streptomyces* phages have been shown to belong to the large serine recombinase family²³. In keeping with this, the ϕ Scoe phages characterized in this work each encode a serine integrase. These proteins are highly variable in sequence across the twelve phages, with pairwise sequence identities ranging from 13 to 91% (Supplementary Fig. S3a). Two groups of phages have more closely related integrases: phages ϕ Scoe2, ϕ Scoe15, and ϕ Scoe55 share 75–91% identity with each other and the previously characterized ϕ Joe integrase, while ϕ Soce25, ϕ Scoe56, ϕ Scoe23, ϕ Soce45, and ϕ Scoe10 share 64–83% identity among their integrases. During phage integration, the serine integrase protein acts at specific sites: the *attP* site in the phage and the *attB* site in the bacterial chromosome. Phages that encode similar integrases and *attP* sites use the same bacterial *attB* sites for integration²⁴. To identify potential *attP* sites in the ϕ Scoe phages we queried their sequences with known *attP* sequences from previously characterized actinophages R4²⁵, ϕ Joe²⁰, ϕ C31²⁶, ϕ BT1²⁷, TG1²⁸, and SV1¹⁴ (Supplementary Fig. S3b). Sequences related to the ϕ Joe *attP* site were identified upstream of the integrase genes in phages ϕ Scoe2, ϕ Scoe15, and ϕ Scoe55 (Supplementary Fig. S3c). These four phages encode integrases that share high sequence identity and are more distantly related to the other serine integrases found in the ϕ Scoe group of phages. The *attP* sites were not identified for the remaining phages.

The genes that are required for the assembly of the head and long, non-contractile tail of the siphophage virion are typically found in a cluster with a conserved gene order. As previously described²⁹, we combined analysis by HHpred³⁰ and gene position to annotate genes encoding morphogenetic functions, such as the large and small terminases, the capsid and portal proteins, head–tail joining proteins, and tail tube protein, that are conserved among most long-tailed phages (Supplementary Table S2). These functions were named according to Casjens et al.²⁹. Proteins comprising the distal end of the tail, referred to as the baseplate or tail tip complex, vary among different types of long-tailed phages. In this region, the ϕ Scoe phages resemble other siphophages that infect Gram-positive species, such as *Lactococcus* phage TP901-1. Accordingly, the first conserved gene in the baseplate region encodes the Distal Tip (Dit) protein and the second gene encodes a protein that structurally resembles the baseplate hub protein (BH2) of myophages³¹ and is often referred to as Tal in the Gram-positive siphophage literature³². Proteins encoded by genes CG21, CG22, and AG22 are predicted by HHpred to resemble receptor binding proteins found in other siphophages; thus, these proteins have been assigned that function (Supplementary Fig. S1). These receptor binding proteins share average pairwise sequence identities of 34–70%. The greatest variability observed in the morphogenetic region was noted in the proteins that comprise the tail tip (Fig. 3). The variability in the tail tip proteins may contribute to the bacterial host range differences noted for these phages.

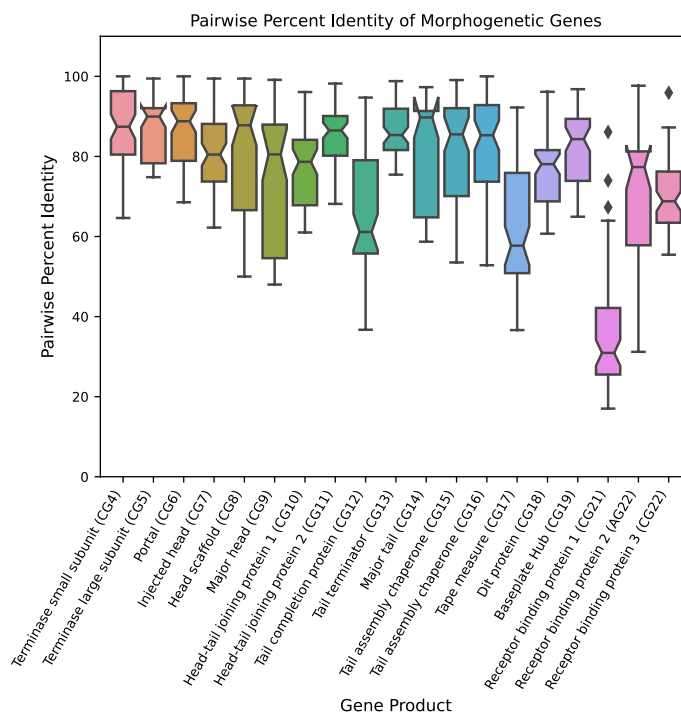


Figure 3. Morphogenetic gene conservation. Box-and-whisker plot of pairwise percent sequence identities for each annotated morphogenetic gene that is conserved among at least ten phages. Receptor binding protein 2 is conserved in all phages except ϕ Scoe23 and ϕ Scoe55. The box bounds the IQR divided by the median, and the whiskers extend to $1.5 \times$ IQR beyond the box. Outliers are shown as individual points.

Non-conserved phage proteins. In addition to the highly conserved genes required for replication and assembly of the phage particle, the ϕ Scoe phages encode many genes that are not conserved among all members of this group, which we refer to as accessory genes (AGs). Between the twelve phages there are a total of 96 different accessory genes that are present in at least one phage. Eleven accessory genes are present in every phage except ϕ Scoe1. Nine of these genes are of unknown function, and the remaining two encode for Lsr2 and Ocr (Overcome classical restriction) proteins (Fig. 2, Supplementary Table S1). Lsr2 is a nucleoid-associated protein that binds AT-rich regions of DNA; it plays a role in silencing prophages in *Streptomyces venezuelae*³³ and *Corynebacterium glutamicum*³⁴. As the ϕ Scoe phages encode repressor proteins that maintain lysogeny, the Lsr2 proteins they encode may play a role in microbial warfare by inhibiting the host encoded Lsr2 proteins. Lsr2 also influences the dynamics of the replication machinery and the duration of DNA synthesis in mycobacteria³⁵, suggesting a possible role for the phage Lsr2 proteins in disrupting host processes to subvert them in a way that benefits the phage. Ocr is a protein found in phage T7 that counters restriction-modification and BREX anti-phage defences^{36,37} and likely plays a similar role in these phages, perhaps to counter the BREX phage growth limitation (Pgl) system³⁸ found in *S. coelicolor*.

Fifty-nine genes are found in only one phage (Fig. 2, Supplementary Table S1). Of these genes, only two have predicted functions; ϕ Scoe23 encodes a DNA polymerase III (AG58), and ϕ Scoe45 contains an endonuclease (AG62) in addition to the HNH-endonuclease that all twelve phages share. Phage ϕ Scoe10 contains a cluster of five accessory genes in the morphogenetic region, between the portal protein and capsid maturation protease (Fig. 2). Four of these genes are annotated as hypothetical proteins, while AG17 shares 62% sequence identity with a tail fibre protein from an *Agrobacterium* phage, suggesting a potential role in bacterial host range. Most accessory genes in this group of phages are located downstream of the morphogenetic region. These genes share synteny, but there are several local rearrangements in this region (Fig. 2). For example, AG67 in ϕ Scoe23 and AG76 in ϕ Scoe2 are present in different relative locations than in the remaining phages that encode these proteins. This pattern is observed for conserved genes as well. For example, conserved gene CG47 appears at different locations in phages ϕ Scoe1, ϕ Scoe2, and ϕ Scoe23 as compared to the rest of the phages.

Phage infection induces secondary metabolite production and early sporulation. While the ϕ Scoe group of phages share many genes across the full length of the genome, the variable accessory gene complements suggested that they may have different host ranges. To address this question, the twelve phages were propagated to high titer in *S. coelicolor* and their plating efficiencies were assessed on a collection of 21 *Streptomyces* species. Each phage showed a unique infection profile (Fig. 4a), with all phages able to efficiently replicate in more than half of the species tested. In some cases, we observed clearings that did not resolve into plaques (grey shading). This suggests that the phages infect but are unable to proceed through the replication cycle, perhaps because the strain lacks a host factor that the phage needs to replicate, or because it encodes a defence system

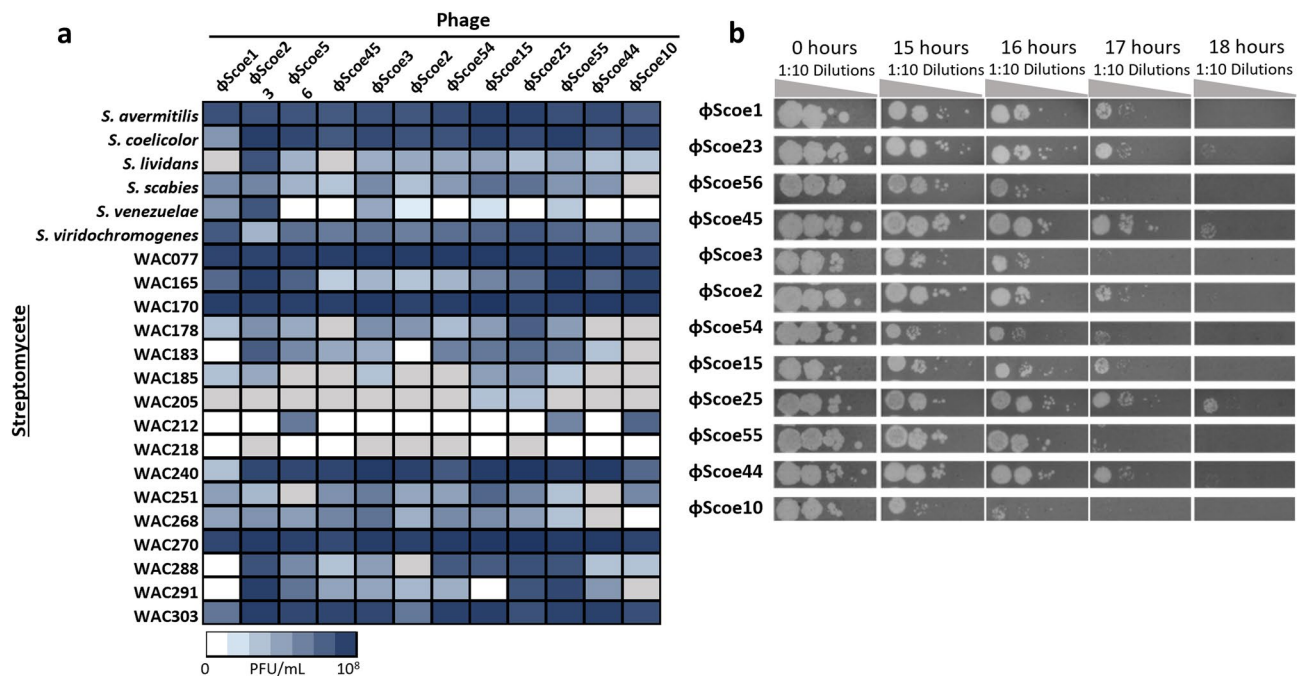


Figure 4. Infection profiles for the ϕ Scoe phages. (a) Host range of the ϕ Scoe phages was assessed using six well characterized *Streptomyces* species and 16 strains from the Wright Actinomycete Collection (WAC). Boxes are colored based on the apparent phage titer on each strain, from 0 to 10^8 PFU/mL. Grey boxes indicate clearance of the bacterial lawn without visible plaques. (b) Infection of *S. coelicolor* over time. Ten-fold serial dilutions of each phage (starting titer of 3×10^4 PFU/mL) were plated on a lawn of *S. coelicolor* spores after 0, 15, 16, 17, or 18 h of growth at 30 °C.

that targets the phage intracellularly. As we previously determined⁷, some species are highly resistant to phage infection, while others are very sensitive. No patterns between phage gene conservation and host range could be discerned. However, we noted that some *Streptomyces* species appear to upregulate secondary metabolite production in the presence of phage challenge as shown by the production of coloured compounds (Figs. 5, 6).

Secondary metabolite production³⁹ and protein-based anti-phage defences^{40–42} are upregulated at high cell density. *Streptomyces* have a complex developmental lifecycle, in which they grow from spores into a branching network of young mycelia. The young mycelia grow into mature mycelia over time, moving into exponential growth approximately 12–14 h post-inoculation^{43,44}. The cells move into a transitional phase after approximately 18 h of growth and undergo a developmental switch, after which they begin producing specialized metabolites and aerial hyphae which mature into spores^{45,46}. To determine how this developmental process affected the ϕ Scoe phages, we determined how efficiently they could initiate infection and replicate in *S. coelicolor* over time. The phage lysates were diluted to approximately 3×10^6 plaque forming units (PFU) per mL and plated on cells that had been grown for various amounts of time following spore germination. Most of the phages maintained normal plating efficiency until ~16 h, with some (e.g. ϕ Scoe2 and ϕ Scoe45) still able to plate on cells that had been grown for 18 h before the phages were applied (Fig. 4b). By contrast, ϕ Scoe56 formed plaques poorly at 15 h, suggesting that this phage was more sensitive to some cellular factor that is differentially expressed during the *Streptomyces* lifecycle. None of the phages were able to efficiently replicate in mature mycelia, which is consistent with previous work showing that *Streptomyces* phages replicate best in exponentially growing cells^{17,45,46}.

We also investigated the effects of ϕ Scoe phage infection on bacterial growth. In this assay ~100 phages were applied to a DNB-agar plate adjacent to a spot containing ~1000 *S. coelicolor* spores, so that the two samples were partially overlapping (Fig. 5a). These cultures were incubated at 30 °C and bacterial growth was monitored using a high-resolution scanner, with images captured every 12 h for a 120 h incubation period. Bacterial growth was initially inhibited by phage infection where the bacterial spores and phages overlapped (Fig. 5b). At 24–36 h bacterial growth began to appear where the initial phage infection had taken place, radiating out from the larger uninfected colony into the zones of clearing. The cells that grew in these areas, which had initially been

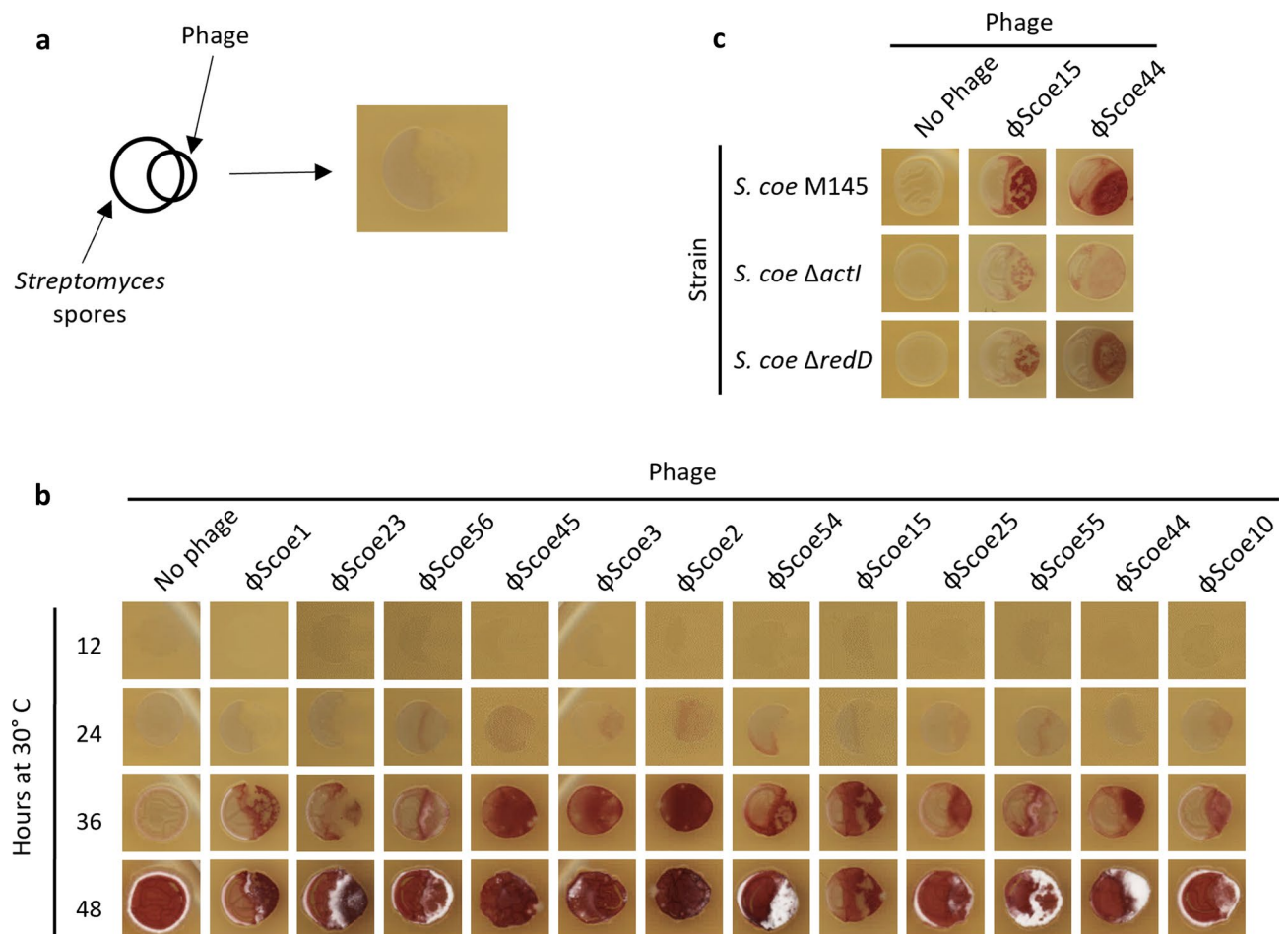


Figure 5. Phenotypic changes in *S. coelicolor* following phage infection. (a) *S. coelicolor* spore stocks were applied to agar plates and phage lysates were deposited so that they partially overlapped the spores to allow for infection to proceed from one side of the bacterial colony. (b) Images of $\sim 10^4$ spores of *S. coelicolor* M145 spotted alone and after infection of one side of the bacterial colony by each phage at 12, 24, 36, and 48 h of growth. (c) Images of *S. coelicolor* M145, $\Delta actI$, and $\Delta redD$ after 36 h of growth, alone or when challenged by phages ϕ Scoe15 and ϕ Scoe44. Images were captured using an Epson Canada Perfection V850 Pro Scanner.

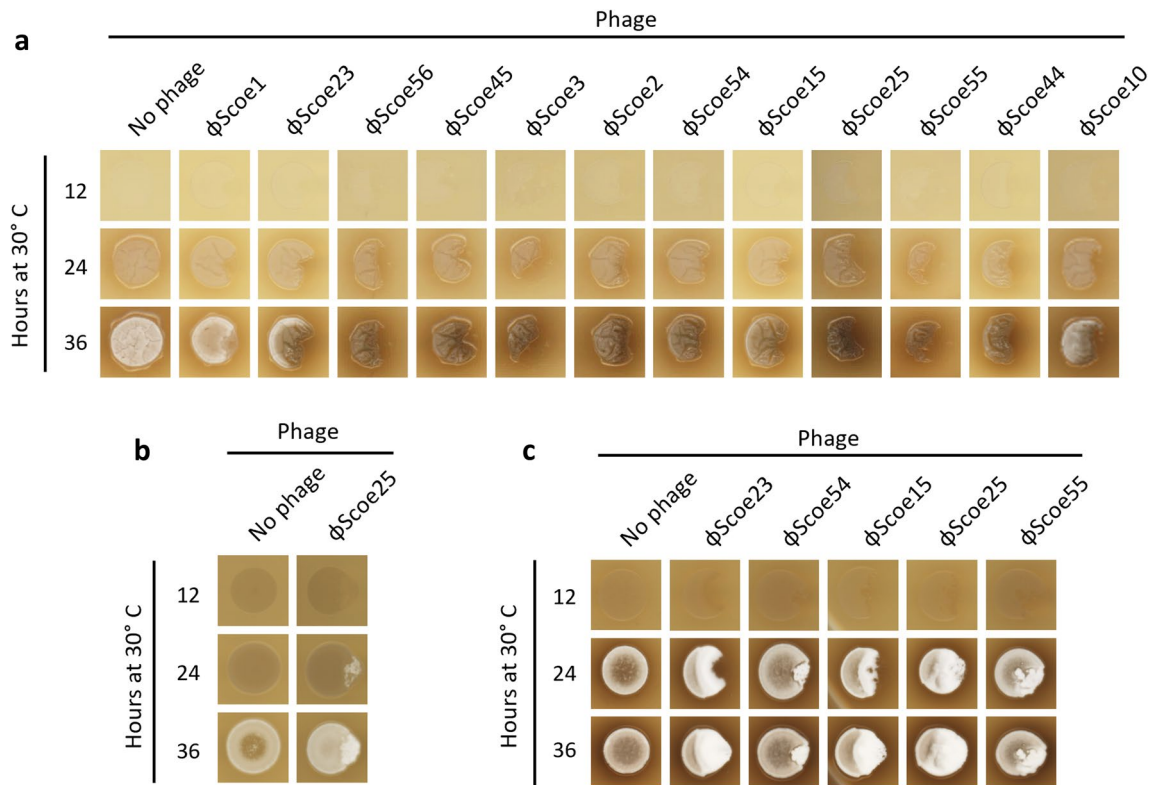


Figure 6. Infection phenotypes of other *Streptomyces*. (a) Images of WAC170 colonies following challenge by the ϕ Scoe phages were recorded at 12, 24, and 36 h of growth. (b) Images of WAC178 following infection by phage ϕ Scoe25 taken after 12, 24, and 36 h of growth. (c) Images of WAC288 spotted alone and following infection by phages ϕ Scoe23, ϕ Scoe54, ϕ Scoe15, ϕ Scoe25, and ϕ Scoe55 taken after 12, 24, and 36 h of growth. All images were captured using an Epson Canada Perfection V850 Pro Scanner.

completely lysed by the activity of the phages, appeared to be resistant to phage infection—this could be either through mutations in the bacteria that provide resistance, through the formation of lysogens, or perhaps through a developmentally mediated resistance mechanism. To determine if bacterial mutations or lysogen formation was the reason for the observed resistance, we isolated colonies from these resistant cells and prepared spore stocks to allow us to test for resistance to further phage infection. Following germination of the spores, no phage resistance was observed and PCR-based screening of the previously phage resistant cells showed that they did not contain prophages. These data suggest that resistance that arises following phage challenge is through some developmentally regulated factor.

The infection and replication of the ϕ Scoe phages in *S. coelicolor* colonies triggered the production of a red-coloured secondary metabolite in response to the phage challenge (Fig. 5b). *S. coelicolor* is known to make two red metabolites, actinorhodin⁴⁷ (red or blue depending on pH) and undecylprodigiosin⁴⁸. To determine which of these metabolites was being induced by phage infection, we monitored growth of *S. coelicolor* strains that harboured mutations in *actI*⁴⁹ and *redD*⁵⁰, which do not produce actinorhodin or undecylprodigiosin, respectively. Infection of wild type *S. coelicolor* by all ϕ Scoe phages induced high levels of red metabolite at 36 h (Fig. 5b, c). This phenotype was greatly abrogated in the *actI* mutant, suggesting that actinorhodin is the major contributor to this phenotype (Fig. 5c). This result is consistent with recent work that showed several other *Streptomyces* phages induced actinorhodin production upon phage infection⁵¹. Infection of *S. coelicolor* by several of the ϕ Scoe phages also appeared to trigger early sporulation in infected colonies as shown by the appearance of a fuzzy white colony morphology at later timepoints (Fig. 5b).

To determine if these metabolite and sporulation effects are a general response to phage infection, we challenged 15 other *Streptomyces* species with phages that were able to plate to high titer and looked for visible cellular responses. Phage infection induced production of a dark brown metabolite in WAC170 (Fig. 6a). Notably, induction of this metabolite appears to be phage-specific; although ϕ Scoe1 visibly infected and killed a portion of the bacterial colony it did not induce production of this secondary metabolite. Strains WAC178 and WAC288 both showed signs of early sporulation in response to phage infection (Fig. 6b,c). As *Streptomyces* phages replicate only in exponentially growing cells^{17,45,46}, this hastening of sporulation may be an adaptive response to allow cells within the colony to become phage resistant and limit the phage epidemic.

Discussion

S. coelicolor has long been used as the model organism for understanding *Streptomyces* biology. Studies of its lifecycle provided critical insight into morphological development pathways and the production of secondary metabolites. In this work we present the genome sequences of twelve *S. coelicolor* phages and show that their activities affect both sporulation and secondary metabolite production. Despite the fact that these phages were isolated from geographically diverse sources spanning multiple continents⁷, they showed remarkable sequence conservation across their complete genomes, as well as to previously characterized phages^{19–21}. This is likely a result of the selective pressure imposed by the bacterial strain on which these phages were initially isolated. The remarkable similarity of these genomes highlights the common worldwide gene pool that phages draw from.

Sequence analyses revealed that this group of phages fall within a single cluster and share high genetic similarity with a variety of other phages infecting *Streptomyces*. The variability of plating behaviours, particularly with respect to secondary metabolite production and the early sporulation effects, is surprising in light of this high degree of sequence conservation. These observed effects on the bacterial host following infection are likely the result of the highly variable accessory gene complement carried by these phages; some of the accessory genes are shared among many of the phages, while others are unique to one. The variable accessory gene content of these phages makes it difficult to predict which ones are contributing to these effects.

While only visible phenotypic changes in bacterial colony morphologies—production of coloured secondary metabolites and early sporulation—following phage challenge were monitored in this study, there are likely many other changes induced by this group of phages. Many *Streptomyces* species, including *S. coelicolor*, produce a wide variety of uncolored metabolites; any changes in production of these compounds in response to phage infection would not have been detected in our assays. However, we believe that induction of secondary metabolites upon phage challenge is likely a frequently occurring response to phage infection in *Streptomyces*. Phage infection may act as a global inducer of specialized metabolism in *Streptomyces*, or it may induce specific metabolites. The reasons for the increased production of secondary metabolites are not yet known, but they may act as a defence against other incoming phages or serve as a warning to surrounding cells of phage infection. More detailed studies will be required to determine what metabolites are induced and the function of these metabolites in the context of phage infection. As secondary metabolite production is tightly linked to sporulation in *Streptomyces*, these effects may also be linked to the acceleration of sporulation observed in *S. coelicolor*, WAC178, and WAC288 in response to phage infection. Our results, in combination with recent work showing induction of chloramphenicol and accelerated sporulation in *S. venezuelae* in response to phage infection¹⁷, suggest a variety of developmental changes are triggered in *Streptomyces* in response to phage infection. Understanding the complex interplay between *Streptomyces* and the phages that infect them has the potential to deepen our understanding of *Streptomyces* bacteria, uncover novel secondary metabolites, and improve our understanding of the role of secondary metabolites in the phage-host evolutionary arms race.

Materials and methods

Phage isolation and maintenance. Phages were isolated from soil samples using overnight enrichment with *Streptomyces coelicolor* M145 (ϕ Scoe45 and ϕ Soce54), WAC212 (ϕ Scoe10), or *Streptomyces avermitilis* SUKA22 (remaining phages) as reported previously⁷. Stocks were maintained by propagation on *S. avermitilis* (ϕ Scoe1) and *S. coelicolor* (all remaining phages) in Difco nutrient broth supplemented with 4 mM CaCl₂ and 0.5% glucose (DNB medium). Phage lysates were stored at 4 °C.

DNA isolation. For each phage, 100 mL of DNB medium was inoculated with 10⁹ spores of *S. coelicolor* and 10⁷ plaque forming units (PFU) of the relevant phage. Cultures were grown at 30 °C for 16 h after which time the culture supernatant was collected. Phages were precipitated via overnight incubation at 4 °C with 20% PEG 8000 and precipitate was collected by centrifugation at 11,000×g for 10 min at 4 °C. Phages were resuspended in 10 mM Tris pH 7.5, 10 mM MgSO₄, 70 mM NaCl, and 1 mM CaCl₂, and DNA was extracted using phenol-chloroform as previously described in Bondy-Denomy et al.⁵².

Genome sequencing and assembly. Phage genomes were sent to the Microbial Genome Sequencing Center in Pittsburgh for sequencing via NextSeq2000. Paired end fastQ files with reads of 30–150 bp long were obtained, and 50,000 reads from each phage was subsampled and assembled using Newbler (GS De novo assembler). Remaining reads were added to the assembly and the resulting single contigs for each phage were then manually curated using Consed⁵³. Bowtie2 was used to determine the percent of reads that aligned to each assembly⁵⁴.

Genome annotation and comparison. Open reading frames (ORFs) were identified and annotated using DNA Master⁵⁵, Glimmer⁵⁶, and GeneMark⁵⁷ software, BLAST searches⁵⁸ and visual inspection. The Phage Annotation Toolkit, our in-house annotation database was used to further annotate morphogenetic phage genes. The annotated genomes were deposited in GenBank (Table 1).

Alignment of the genomes was performed using Clinker⁵⁹ to determine shared percent identity of each ORF between these phages with a cut-off of 30% identity. Conserved genes are present in every phage and are defined as genes that either share > 30% identity or are annotated with the same specific function through BLAST search (i.e. serine integrase). Phage lifecycle was determined by visual inspection and identification of serine integrases in all phages and attachment sites of several phages using BLAST.

Phage genomes were assessed for the presence of known stoperator sites via MEME suite⁶⁰. The hits were then analyzed for location and summary table was generated to show intergenic hits of each known stoperator

sequence in each ϕ Scoe phage. Several regions where characterized stoperators are located were aligned to equivalent regions in each ϕ Scoe phage and a percent identity matrix was generated via Clustal Omega⁶¹.

Pairwise average nucleotide identities (ANI) were calculated with the twelve ϕ Scoe phages and characterized *Streptomyces* phages R4 and ϕ Joe using pyani 0.2.9¹⁸ with ANIb method. The output average nucleotide identity matrix was used for clustering and display via Seaborn clustermaps^{62,63}.

To find phage attachment sites, intergenic DNA flanking the integrase of each phage was BLASTed against published attP sites from 6 characterized *Streptomyces* phages. Significant hits were examined manually and matched up to the characterized attP site, and sequences with >70% identity over the entire attP site were reported.

Host range and efficiency of plating. Phages were propagated on *S. avermitilis* (ϕ Scoe1) or *S. coelicolor* (remaining phages). Ten-fold serial dilutions were plated on DNB plates overlaid with 0.5% (w/v) top agar containing approximately 10^6 spores of one of six known Streptomycetes or 16 strains from the Wright Actinomycete Collection. The titre of each phage on each strain was recorded, as was the appearance of clearings that did not resolve into individual plaques. Experiments were performed with three biological replicates.

Infection time-course. All phages were propagated on *S. coelicolor*, then diluted to 3×10^6 PFU/mL in liquid DNB. Plates and serial dilutions were prepared as described for host range with DNB top agar containing 3×10^7 *S. coelicolor* spores. Dilutions were spotted on fresh overlays as time 0, following which the plates were placed at 30 °C. Fresh phage dilutions were prepared the following day and spotted at 15, 16, 17, and 18 h post overlay with incubation at 30 °C between each timepoint. The resulting plaques were counted to determine phage titre at each hour and the experiment was performed in triplicate.

Phage infection phenotypes. To investigate phenotypic changes in *Streptomyces* upon phage infection, preliminary screening was performed using a plate pinner. Spores of 16 *Streptomyces* species were diluted to $\sim 10^6$ spore forming units per mL (sfu/mL) and aliquoted into wells of a 96 well plate. Phages were diluted to $\sim 10^7$ plaque forming units per mL (PFU/mL) and tenfold serial dilutions of each were aliquoted into descending wells of a 96 well plate, with no phage added to the bottom well of each column. ISP-2 agar supplemented with 4 mM CaCl₂ and made with either all MiliQ water or half MiliQ and half tap water (ISP-2 and ISP-2 tap), and MYM agar supplemented with 4 mM CaCl₂ and made with MiliQ water only (MYM MiliQ) were poured into Singer PlusPlates and allowed to dry. Approximately 0.5 mL each of spores were spotted onto a plate containing each medium at 384 spot density using a Singer ROTOR HDA pinning robot and Singer 96 density RePads and allowed to dry. Approximately 0.5 mL of each phage dilution was then spotted on top of each spore spot using the Singer ROTOR HDA and RePads. Plates were allowed to dry, and then placed facedown onto the glass of an Epson Canada Perfection V850 Pro Scanner—B11B224201. Pictures were taken using the scanner at 6 h increments for 120 h, controlled by Raspberry Pi and code adapted from that described in French et al.⁶⁴. Species used include *S. coelicolor*, *S. avermitilis*, *S. venezuelae*, WAC077, WAC170, WAC178, WAC185, WAC205, WAC212, WAC218, WAC240, WAC251, WAC268, WAC270, WAC288, WAC291, and WAC303. The experiment was performed with four technical replicates per pinned plate and observed phenotypic changes were confirmed as below.

To confirm the phenotypic changes observed using the scanner, ~ 1000 spores of *Streptomyces* strains of interest were plated on the medium on which phenotypic changes were observed. Spore spots were allowed to dry and then ~ 500 phages of interest were spotted onto one edge of each spore spot. Plates were allowed to dry and placed face-down on the glass of an Epson Canada Perfection V850 Pro Scanner—B11B224201 and pictures were taken as described above, with a focus on the timepoints of interest found using the large-scale method above. Pictures were taken of each strain every 12 h for 36 h total. Each phage was plated on the strains it was best able to replicate in and for which there was a noted phenotype. All phages were plated on *S. coelicolor* and WAC170; phage ϕ Scoe25 was plated on WAC178; and phages ϕ Scoe23, ϕ Scoe54, ϕ Scoe15, ϕ Scoe25, and ϕ Scoe55 were plated on WAC288. Each strain was also plated in the absence of phage. *S. coelicolor* was plated on ISP-2 tap medium; WAC288 and WAC170 were plated on ISP-2 medium; and WAC178 was plated on MYM MiliQ medium.

To determine the identity of the colored compounds produced in *S. coelicolor* upon phage infection, wild-type, $\Delta actI$, and $\Delta redD$ *S. coelicolor* spores were plated on ISP-2 tap medium and challenged with phages ϕ Scoe15 and ϕ Scoe44 as described above. Experiments were performed in triplicate.

Data availability

All data generated or analysed during this study are included in this published article and its Supplementary Information files. Phage genomes are available in GenBank with accession numbers as shown in Table 1.

Received: 20 April 2023; Accepted: 13 June 2023

Published online: 17 June 2023

References

- Cheng, K. et al. Population genetic analysis of *Streptomyces albidoflavus* reveals habitat barriers to homologous recombination in the diversification of streptomycetes. *Appl. Environ. Microbiol.* **81**, 966–975 (2015).
- Bentley, S. D. et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 7 (2002).
- Ikeda, H. et al. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat. Biotechnol.* **21**, 526–531 (2003).

4. Ohnishi, Y. *et al.* Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. *J. Bacteriol.* **190**, 4050–4060 (2008).
5. Bérdy, J. Bioactive microbial metabolites: A personal view. *J. Antibiot.* **58**, 1–26 (2005).
6. Watve, M., Tickoo, R., Jog, M. & Bhole, B. How many antibiotics are produced by the genus *Streptomyces*?. *Arch. Microbiol.* **176**, 386–390 (2001).
7. Kronheim, S. *et al.* A chemical defence against phage infection. *Nature* **564**, 283–286 (2018).
8. Jiang, Z., Wei, J., Liang, Y., Peng, N. & Li, Y. Aminoglycoside antibiotics inhibit mycobacteriophage infection. *Antibiotics* **9**, 714 (2020).
9. Zuo, P., Yu, P. & Alvarez, P. J. J. Aminoglycosides antagonize bacteriophage proliferation, attenuating phage suppression of bacterial growth, biofilm formation, and antibiotic resistance. *Appl. Environ. Microbiol.* **87**, e00468–e521 (2021).
10. Kever, L. *et al.* Aminoglycoside antibiotics inhibit phage infection by blocking an early step of the infection cycle. *MBio* **13**, e00783–e822 (2022).
11. Wommack, K. E. & Colwell, R. R. Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* **64**, 69–114 (2000).
12. Suttle, C. A. Marine viruses—Major players in the global ecosystem. *Nat. Rev. Microbiol.* **5**, 801–812 (2007).
13. Baltz, R. H. *Streptomyces* temperate bacteriophage integration systems for stable genetic engineering of actinomycetes (and other organisms). *J. Ind. Microbiol. Biotechnol.* **39**, 661–672 (2012).
14. Fayed, B., Younger, E., Taylor, G. & Smith, M. C. M. A novel *Streptomyces* spp. integration vector derived from the *S. venezuelae* phage, SV1. *BMC Biotechnol.* **14**, 51 (2014).
15. Yagüe, P., Lopez-Garcia, M. T., Rioseras, B., Sanchez, J. & Manteca, A. New insights on the development of *Streptomyces* and their relationships with secondary metabolite production. *Curr. Trends Microbiol.* **8**, 65–73 (2012).
16. *Practical Streptomyces genetics*. (Innes, 2000).
17. Luthe, T. *et al.* *Streptomyces* development is involved in the efficient containment of viral infections. *microLife* **4**, uqad002 (2023).
18. Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G. & Toth, I. K. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal. Methods* **8**, 12–24 (2016).
19. Chater, K. F. & Carter, A. T. A new, wide host-range, temperate bacteriophage (R4) of *Streptomyces* and its interaction with some restriction-modification systems. *J. Gen. Microbiol.* **115**, 431–442 (1979).
20. Fogg, P. C. M., Haley, J. A., Stark, W. M. & Smith, M. C. M. Genome integration and excision by a new *Streptomyces* bacteriophage, ϕ Joe. *Appl. Environ. Microbiol.* **83**, e02767–e2816 (2017).
21. Smith, M. C. M. *et al.* Evolutionary relationships among actinophages and a putative adaptation for growth in *Streptomyces* spp. *J. Bacteriol.* **195**, 4924–4935 (2013).
22. Russell, D. A. & Hatfull, G. F. PhagesDB: the actinobacteriophage database. *Bioinformatics* **33**, 784–786 (2017).
23. Smith, M. C. M., Brown, W. R. A., McEwan, A. R. & Rowley, P. A. Site-specific recombination by ϕ C31 integrase and other large serine recombinases. *Biochem. Soc. Trans.* **38**, 388–394 (2010).
24. Smith, M. C. M. Phage-encoded serine integrases and other large serine recombinases. *Microbiol. Spectr.* **3**, 3.4.06 (2015).
25. Shirai, M., Nara, H., Sato, A., Aida, T. & Takahashi, H. Site-specific integration of the actinophage R4 genome into the chromosome of *Streptomyces parvulus* upon lysogenization. *J. Bacteriol.* **173**, 4237–4239 (1991).
26. Gupta, M., Till, R. & Smith, M. C. M. Sequences in attB that affect the ability of ϕ C31 integrase to synapse and to activate DNA cleavage. *Nucleic Acids Res.* **35**, 3407–3419 (2007).
27. Gregory, M. A., Till, R. & Smith, M. C. M. Integration site for *Streptomyces* phage ϕ BT1 and development of site-specific integrating vectors. *J. Bacteriol.* **185**, 4 (2003).
28. Morita, K. *et al.* The site-specific recombination system of actinophage TG1. *FEMS Microbiol. Lett.* **297**, 234–240 (2009).
29. Casjens, S. R., Davidson, A. R. & Grose, J. H. The small genome, virulent, non-contractile tailed bacteriophages that infect Enterobacteriales hosts. *Virology* **573**, 151–166 (2022).
30. Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–W248 (2005).
31. Büttner, C. R., Wu, Y., Maxwell, K. L. & Davidson, A. R. Baseplate assembly of phage Mu: Defining the conserved core components of contractile-tailed phages and related bacterial systems. *Proc. Natl. Acad. Sci. U S A* **113**, 10174–10179 (2016).
32. Velesler, D. *et al.* Structure of the phage TP901–1 1.8 MDA baseplate suggests an alternative host adhesion mechanism. *Proc. Natl. Acad. Sci. USA* **109**, 8954–8958 (2012).
33. Gehrke, E. J. *et al.* Silencing cryptic specialized metabolism in *Streptomyces* by the nucleoid-associated protein Lsr2. *Elife* **8**, e47691 (2019).
34. Pfeifer, E. *et al.* Silencing of cryptic prophages in *Corynebacterium glutamicum*. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkw692> (2016).
35. Kołodziej, M. *et al.* Lsr2, a nucleoid-associated protein influencing mycobacterial cell cycle. *Sci. Rep.* **11**, 2910 (2021).
36. Roberts, G. A. *et al.* Exploring the DNA mimicry of the Ocr protein of phage T7. *Nucleic Acids Res.* **40**, 8129–8143 (2012).
37. Isaev, A. *et al.* Phage T7 DNA mimic protein Ocr is a potent inhibitor of BREX defence. *Nucleic Acids Res.* **48**, 5397–5406 (2020).
38. Sumbly, P. & Smith, M. C. M. Genetics of the phage growth limitation (Pgl) system of *Streptomyces coelicolor* A3(2): Phage growth limitation phenotype in *S. coelicolor*. *Mol. Microbiol.* **44**, 489–500 (2002).
39. Sánchez, L. & Braña, A. F. Cell density influences antibiotic biosynthesis in *Streptomyces clavuligerus*. *Microbiology* **142**, 1209–1220 (1996).
40. Patterson, A. G. *et al.* Quorum sensing controls adaptive immunity through the regulation of multiple CRISPR-Cas systems. *Mol. Cell* **64**, 1102–1108 (2016).
41. Høyland-Kroghsbo, N. M., Mærkedahl, R. B. & Svenningsen, S. L. A quorum-sensing-induced bacteriophage defense mechanism. *MBio* **4**, e00362–e412 (2013).
42. O'Hara, B. J., Alam, M. & Ng, W.-L. The vibrio cholerae seventh pandemic Islands act in tandem to defend against a circulating phage. *PLoS Genet.* **18**, e1010250 (2022).
43. Li, S., Wang, W., Li, X., Fan, K. & Yang, K. Genome-wide identification and characterization of reference genes with different transcript abundances for *Streptomyces coelicolor*. *Sci. Rep.* **5**, 15840 (2015).
44. Huang, J., Lih, C.-J., Pan, K.-H. & Cohen, S. N. Global analysis of growth phase responsive gene expression and regulation of antibiotic biosynthetic pathways in *Streptomyces coelicolor* using DNA microarrays. *Genes Dev.* **15**, 3183–3192 (2001).
45. Reyes-Gavilán, C. G., Aparicio, J. F., Barbés, C., Carlos, H. & Sánchez, J. Adsorption capability determines phage plaque size on *Streptomyces antibioticus* producing endodeoxyribonuclease. *FEMS Microbiol. Lett.* **56**, 301–305 (1988).
46. Rosner, A. & Gutstein, R. Adsorption of actinophage Pal 6 to developing mycelium of *Streptomyces albus*. *Can. J. Microbiol.* **27**, 254–257 (1981).
47. Bystrykh, L. V. *et al.* Production of actinorhodin-related 'blue pigments' by *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **178**, 2238–2244 (1996).
48. Tsao, S.-W., Rudd, B. A. M., He, X.-G., Chang, C.-J. & Floss, H. G. Identification of a red pigment from *Streptomyces coelicolor* A3(2) as a mixture of prodigiosin derivatives. *J. Antibiot.* **38**, 128–131 (1985).
49. Lu, Z., Xie, P. & Qin, Z. Promotion of markerless deletion of the actinorhodin biosynthetic gene cluster in *Streptomyces coelicolor*. **5**.

50. Floriano, B. & Bibb, M. *afsR* is a pleiotropic but conditionally required regulatory gene for antibiotic production in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.* **21**, 385–396 (1996).
51. Hardy, A., Sharma, V., Kever, L. & Frunzke, J. Genome sequence and characterization of five bacteriophages infecting *Streptomyces coelicolor* and *Streptomyces venezuelae*: Alderaan, Coruscant, Dagobah, Endor1 and Endor2. *Viruses* **12**, 1065 (2020).
52. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429–432 (2013).
53. Gordon, D. & Green, P. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* **29**, 2936–2937 (2013).
54. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).
55. Lawrence, J. DNA Master. (2007).
56. Delcher, A. L., Bratke, K. A., Powers, E. C. & Salzberg, S. L. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **23**, 673–679 (2007).
57. Besemer, J. & Borodovsky, M. GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res.* **33**, W451–W454 (2005).
58. Sayers, E. W. *et al.* Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **50**, D20–D26 (2022).
59. Gilchrist, C. L. M. & Chooi, Y.-H. clinker & clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics* **37**, 2473–2475 (2021).
60. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME suite. *Nucleic Acids Res.* **43**, W39–W49 (2015).
61. Goujon, M. *et al.* A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* **38**, W695–W699 (2010).
62. Waskom, M. seaborn: Statistical data visualization. *JOSS* **6**, 3021 (2021).
63. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
64. French, S. *et al.* A robust platform for chemical genomics in bacterial systems. *MBoC* **27**, 1015–1025 (2016).

Acknowledgements

This work was supported by a Natural Sciences and Engineering Research Council Discovery Grant (RGPIN-2017-06898) and the John C. Polanyi Award (JCP 538013-2020) to K.L.M.

Author contributions

S.K. performed and analyzed most of the biological experiments, performed the bioinformatic analyses and comparisons, led genome sequencing, assembly, and annotation and wrote the manuscript; E.S. performed and analyzed the phage plating time-course experiments and assisted with infection morphology experiments; L.H. assisted with genome sequencing, assembly, and annotation; and M.G. assisted with genome annotation. A.R.D. assisted with bioinformatics and manuscript preparation. K.L.M. helped design and supervised all experiments, analysed data, and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-36938-z>.

Correspondence and requests for materials should be addressed to K.L.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023