Mechanisms and consequences of horizontal gene transfer in cell wall-deficient cells of Kitasatospora viridifaciens

Kapteijn, R.

Chapter 2.

Hyperosmotic stress conditions lead to major phenotypic and genetic diversity in *Kitasatospora viridifaciens*

Renée Kapteijn, Valerie Elise Gerritsen, Denzel Ignacia, Ulrika Victoria Ferner, Daniel E. Rozen, Gilles P. van Wezel and Dennis Claessen

Abstract

Bacteria need to respond adequately to overcome stresses associated with fluctuating environments. Some streptomycetaceae display a striking morphological response upon exposure to hyperosmotic stress by forming transient cell wall-deficient cells. Hyperosmotic stress also induces significant morphological heterogeneity among colonies of *Kitasatospora viridifaciens*. Here we show that colonies with an aberrant phenotype after exposure to osmotic stress frequently contain large genomic rearrangements, exemplified by deletions within the chromosomal arms, amplifications and (partial) loss of a 1.7 Mbp megaplasmid. The presence of genes associated with transposable elements at genomic breakpoints suggests a role for mobile genetic elements in genomic instability. These findings provide new insights into the bacterial response to hyperosmotic stress and how such conditions contribute to genetic diversity and evolution.

Introduction

Bacteria are exposed to changing environmental osmolarity. This is often caused by fluctuations in water availability due to high temperature, rainfall or desiccation, or by the dispersal of bacteria to a new environment [176]. To adapt to changes in extracellular osmolarity, bacteria adjust the composition of their cytoplasm to maintain the correct turgor pressure. An increase in external osmotic pressure is typically counteracted by increasing the cytoplasmic concentration of ions (potassium and chloride ions) and/or compatible solutes (organic osmolytes such as amino acids, ectoine and trehalose)[177, 178]. These compounds increase the internal osmotic potential which prevents water efflux and cell shrinkage.

Streptomyces and *Kitasatospora* (hereafter: streptomycetes) frequently need to respond to changes in environmental conditions. These filamentous actinobacteria are abundant in soil environments where they play an important role in the recycling of organic matter [61]. Their life cycle involves complex morphological differentiation starting with the germination of a spore, which develops into a network of branching hyphae, called the substrate mycelium [9]. In response to environmental signals such as nutrient limitation, part of the substrate mycelium undergoes autolytic degradation, providing nutrients for the development of aerial hyphae which rise into the air, producing unigenomic spores that can be dispersed [6].

As sessile, soil dwelling bacteria, streptomycetes frequently encounter osmotic challenges such as low (desiccation, freezing) or high water availability (e.g. rainfall) [179]. Hyperosmotic stress results in a temporary arrest in growth, followed by a change in cell polarity giving rise to a hyperbranching mycelium in *Streptomyces coelicolor*, *Streptomyces venezuelae* and *Kitasatospora viridifaciens* [15, 180, 181]. In addition, around 5-10% of natural *Streptomyces* and *Kitasatospora* isolates produce viable cell wall-deficient cells called S-cells (for stress-induced cells)[15]. During this process, hyphae temporarily stop growing and DNA-containing vesicles are extruded from the tip. These S-cells are neither enveloped by their cell wall nor proliferate in this wall-deficient state. Instead, they increase in size and, under specific conditions, rebuild their cell wall to revert to their canonical mycelial mode-of-growth.

In addition to S-cell formation, growth under hyperosmotic conditions also leads to morphologically diverse colonies in *K. viridifaciens*, some of which cannot fully sporulate and appear to have lost the large linear megaplasmid KVP1, which may be an indication of genetic instability[15]. The generation of phenotypic and/or genetic diversity can be a strategy to increase survival of bacteria when exposed to environmental stress[182-184]. The cause of the heterogeneity observed in *K. viridifaciens* after growth under hyperosmotic stress conditions is unknown, as well as whether this diversity could function as adaptation to survive osmotic stress.

Here we show that growth of *K. viridifaciens* under hyperosmotic stress conditions results in drastic phenotypic and genetic changes. Hyperosmotic stress induces the formation of a high frequency of heterogenous colonies with irreversible phenotypes, which are, surprisingly, no more resistant to hyperosmotic stress than the wild-type. An extensive analysis of heterogenous colonies using whole-genome sequencing indicates the occurrence of large chromosomal deletions, amplifications and putative rearrangements. Several of these genetic changes correlate with the presence of genes associated with transposable elements (TEs), suggesting a role for mobile genetic elements (MGEs) in genomic instability and in the bacterial response to osmotic stress.

Results

Colonies from S-cell filtrate originate from a mixture of walled and wall-less cells

Growth of *K. viridifaciens* under hyperosmotic conditions by liquid culture in LPB medium containing 22% (w/v) sucrose leads to the formation of phenotypically diverse colonies [15]. The same conditions induce the formation of temporary wall-less S-cells, which can revert to a mycelial-mode-of-growth to enable proliferation of the bacterium in a new environment. We therefore wondered whether S-cells give rise to the phenotypically diverse colonies. S-cells can be isolated by successive filtration of 48 h *K. viridifaciens* liquid cultures in LPB medium trough Ecocloth™ and a 5 µm membrane filter. However, besides S-cells, these filtrates can contain small mycelial fragments and ungerminated spores.

To determine which percentage of colonies truly originate from S-cells, we treated S-cell filtrate either with water (which causes S-cells to burst) or with 22% sucrose. Cells were then plated on different solid media with increasing sucrose concentrations that were specific for mycelial growth (MYM medium – 0% sucrose), reversion of wall-less protoplasts (R5 medium – 10.3% sucrose) and for proliferation of wall-less cells (LPMA medium – 20% sucrose). As expected, water treatment drastically reduced the average CFU ml^{-1} obtained from S-cell filtrate by around 9- and 53-fold on R5 medium (n = 3, paired t-test gives $t(2)$ = 10.03, P = 0.010) and LPMA medium (n = 3, paired t-test gives $t(2)$ = 5.934, P = 0.027), respectively (Fig. 1A). Unexpectedly, water treatment did not significantly reduce the CFU ml^{-1} of S-cell filtrate on MYM medium (n = 3, paired t-test gives $t(2)$ = 1.964, P = 0.189), which is the medium used to detect heterogenous colonies. Previous research indicated that S-cells with abundant cell wall material remain intact after water treatment [15]. MYM medium may therefore positively select for S-cells that contain a certain amount of peptidoglycan. Therefore, the heterogenous colonies likely originate from a mix of S-cells containing abundant cell wall material, small mycelial fragments and spores exposed to hyperosmotic stress conditions.

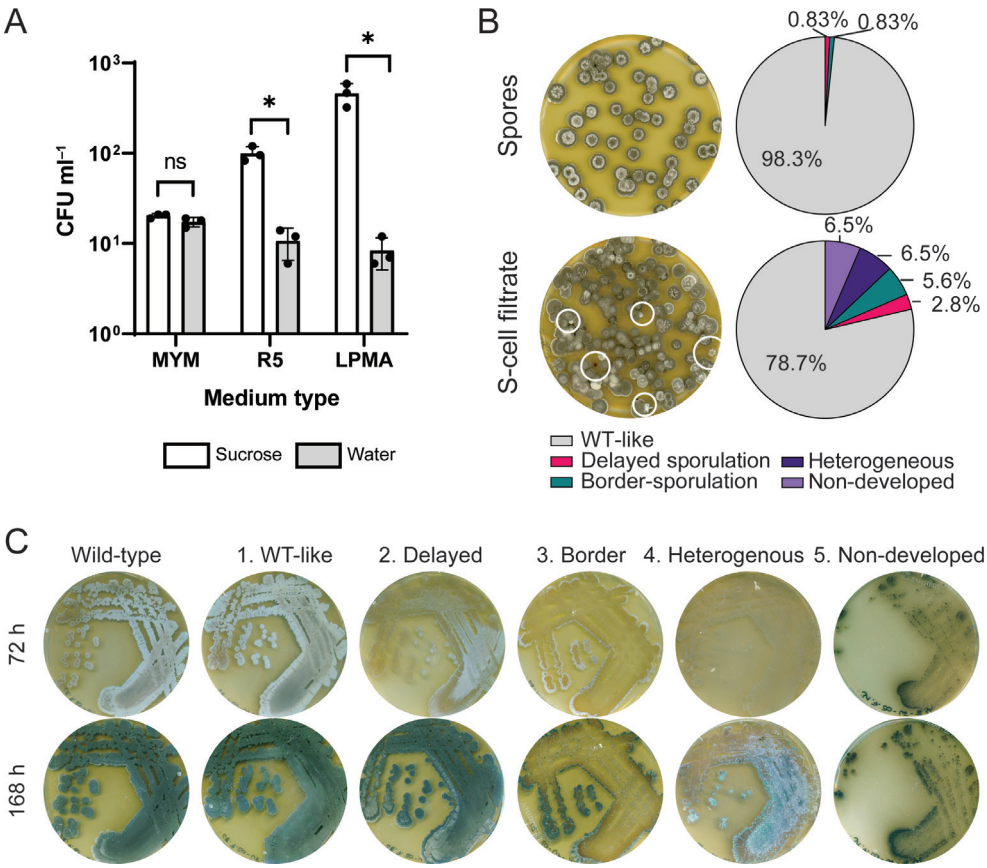


Figure 1. Phenotypic heterogeneity in *K. viridifaciens* after exposure to osmotic stress
(A) Viable units obtained after exposure of S-cell filtrate to 22% (w/v) sucrose or water as measured by colony forming units (CFU) per ml obtained after 168 h incubation on MYM, R5 or LPMA medium at 30 °C. Asterisk (*) indicates $P \leq 0.05$ between treatments ($n = 3$, paired t-test gives $t(2) = 1.964$, $P = 0.189$ for MYM medium; $t(2) = 10.03$, $P = 0.010$ for R5 medium and $t(2) = 5.934$, $P = 0.027$ for LPMA medium). ns = not significant. (B) Frequency of colonies (right) with an aberrant phenotype of *K. viridifaciens* wild-type spores (120 colonies) and S-cell filtrate (108 colonies) obtained after growth on MYM medium for 120 h at 30 °C. Left: Representative MYM agar plates to illustrate the colony diversity for each cell type. Colonies with an aberrant phenotype are indicated by white circles. (C) Phenotype of wild-type *K. viridifaciens* and representative isolates from each isolate category (1 = wild-type-like, 2 = delayed sporulation, 3 = border sporulation, 4 = heterogenous sporulation, 5 = non-developed) obtained after growth of S-cell filtrate on MYM medium for 72 h and 168 h. See also Supplementary Table 2.

Conditions other than MYM medium promote the reversion of S-cells that are sensitive to water treatment. The total average CFU ml^{-1} of S-cell filtrate on R5 medium and LPMA medium was around 5- and 23-fold higher, respectively, as compared to MYM medium (Fig. 1A, 'sucrose treatment'). Conversely, for spores the CFU ml^{-1} on R5 and LPMA medium was similar or lower, respectively, compared to MYM medium (Supplementary Fig. 1A). Higher sucrose concentrations may therefore promote S-cell survival and reversion to mycelial

growth on solid medium. Drying of agar is used in some *Streptomyces* protoplast regeneration protocols to increase the total CFU on solid medium [185, 186]. Prolonged drying of solid medium plates before or after application of S-cell-filtrate indeed increased the formation of colonies on LPMA medium, with up to 30% of S-cells forming a colony (Supplementary Fig. 1B, C and Supplementary Table 1). These results confirm that S-cells are viable but require specific conditions for reversion to a walled state on solid medium.

Exposure to hyperosmotic stress leads to phenotypic heterogeneity

To study the frequency of heterogenous colonies arising from S-cell filtrate compared to spores, more than 100 random single colonies were obtained on MYM medium from S-cell filtrate and compared to colonies originating from spores. The frequency of heterogeneous colonies obtained after plating S-cell filtrate was more than ten-fold higher than observed with spores (Fig. 1B; 21.3% and 1.7%, respectively). Further investigation of the colonies obtained from S-cell filtrate on MYM medium revealed that the majority possessed a wild-type-like phenotype (Fig. 1B) (78.7%, category 1 WT-like). The remaining colonies with an aberrant phenotype could be further divided into strains with delayed development as compared to the wild-type (2.78%, category 2 'delayed sporulation'), sporulation only at the edges of the mycelial biomass (5.56%, category 3 'border-sporulation'), heterogenous sporulation (nonconfluent sporulation, white and non-developed biomass visible) (6.48%, category 4 'heterogenous') and strains that did not appear to sporulate under the chosen conditions, namely those colonies that developed white aerial mycelium or non-developing ('bald' phenotype) colonies (6.48%, category 5 'non-developed') (Fig. 1C). Five isolates per category were stored for further analysis, with colonies having a bald phenotype representing category 5 (Supplementary Table 2). Isolates are represented by their category number followed by a unique identifier (e.g. 1.10 is isolate nr. 10 from cat. 1).

Microorganisms can use transient phenotypic (nongenetic) variability, to adapt to fluctuating environments [184, 187]. To test the phenotypic stability of the isolates, three isolates per category were grown for five rounds on MYM medium without sucrose. For each round their phenotype was observed after 72 h and 168 h growth (Fig. 2A). None of the strains from the mutant categories regained a wild-type mode-of-growth at any point in time (Fig. 2B, Supplementary Fig. 2A, B). This strongly indicates that genetic changes underly these mutant phenotypes.

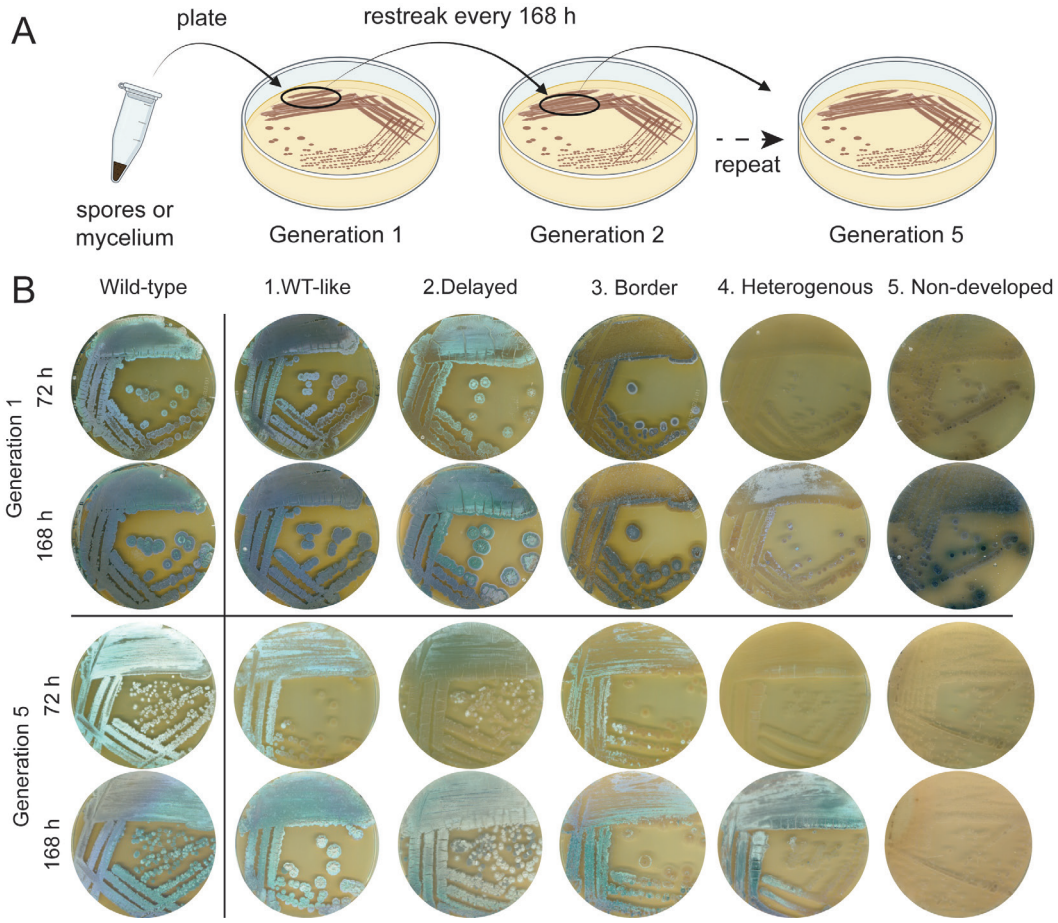


Figure 2. Osmotic stress-induced aberrant phenotypes do not change back to a wild-type phenotype
(A) Schematic overview of the procedure to test the stability of isolates with an aberrant phenotype. Spores (cat. 1 – 4 and wild-type) or mycelial biomass (cat. 5) of *K. viridifaciens* wild-type and three strains per phenotype category were grown on MYM medium for 168 h at 30 °C (Generation 1), after which a sample of biomass was restreaked on fresh MYM medium (Generation 2) and grown for 168 h. This was repeated three times up to Generation 5. Plates were imaged every 72 h and 168 h to determine the isolate phenotype. (B) Phenotypic stability of *K. viridifaciens* wild-type and isolates 1.38 (cat. 1), 2.12 (cat. 2), 3.24 (cat. 3), 4.4 (cat. 4) and 5.6 (cat. 5) at Generation 1 and 5, imaged after 72 and 168 h of growth at 30 °C. Note that none of the strains regain a wild-type phenotype. See also Supplementary Fig. 2.

Hyperosmotic stress induces genomic instability

Formation of spontaneous mutant colonies is common in *Streptomyces* and is linked to genetic instability [188, 189]. Genetic instability often leads to loss of genes involved in amino acid synthesis [190, 191]. Growth of five isolates per category on minimal medium (MM) showed that only colonies with a bald phenotype (cat. 5) were not able to grow without the supplementation of casamino acids, in contrast to the wild-type and sporulating strains (Fig. 3). Specifically, the presence of arginine rescued the lack of growth, suggesting that isolates of cat. 5 lost part of their chromosome that encodes for genes involved in arginine biosynthesis.

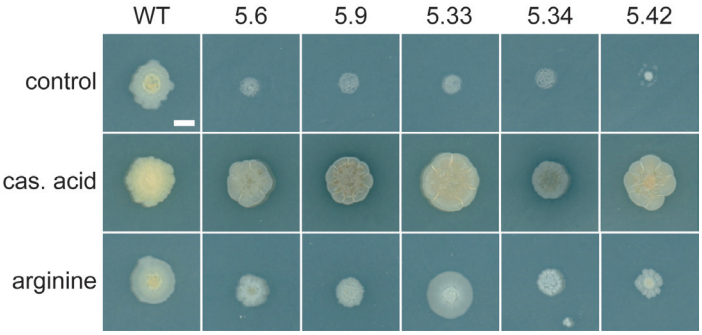


Figure 3. Isolates with a bald phenotype are auxotrophic for arginine
Growth of mycelium of *K. viridifaciens* wild-type and isolates from cat. 5 after a 96-h incubation at 30 °C on Minimal Medium (MM) without supplements (control) or with the addition of 0.5% (w/v) Casamino Acids (cas. acid) or 0.0037% (w/v) L-arginine. Addition of arginine rescues the lack of growth observed for all isolates with a bald phenotype. Note that the biomass observed in the control condition for cat. 5 is the mycelial biomass used as inoculum, and does not indicate growth. Scale bar = 0.5 cm.

To study the genetic changes underlying the heterogenous phenotypes, five isolates per category were subjected to whole-genome sequencing (WGS) using short-read Illumina sequencing technology (Supplementary Fig. 3). Mapping of the Illumina reads of an isolate to the *K. viridifaciens* DSM40239 reference genome, consisting of the linear chromosome (7.9 Mbp), a linear megaplasmid (pKVP1; 1.7 Mbp) and a separate linear contig (accession nr. CP090842; 115 Kbp), allows the detection of large chromosomal deletions by loss of read coverage, as well as detection of single nucleotide polymorphisms (SNPs).

As expected, mapping of the Illumina reads of the wild-type and wild-type-like strains (cat. 1) resulted in consistent coverage across all contigs (Supplementary Fig. 3A). In contrast, the genomes of 15 out of 25 isolates contained large deletions ranging from 20 kbp – 1.7 Mbp in size and included (part of) the megaplasmid, the ends of one or both chromosomal arms and contig 3 (Fig. 4 A, B and Supplementary Fig. 3B – E). Whereas a subset of the strains from cat. 2 and 3 were only affected in the megaplasmid, more drastic deletions were observed in strains from cat. 4 and 5, which were affected in both the megaplasmid and the chromosome. Some common patterns were revealed by study of the deleted regions across mutant strains.

Deletions at the end of the chromosomal arm(s) always correlated with (partial) loss of the megaplasmid (e.g. strain 4.4), but the megaplasmid itself could be lost without affecting other regions (e.g. strain 3.24).

In four cases (4.4, 4.8, 4.36 and 5.9) the right chromosomal arm was truncated by around 1 Mbp at the exact same position, namely at a locus encoding for a predicted ABC transporter ATP-binding protein (6,881,451 – 6,880,681), suggesting that this deleted region may be sensitive to genetic instability (Supplementary Fig. 3D, E). In addition, all strains with a bald phenotype (cat. 5) had lost part of the left chromosomal arm containing *argG* (BOQ63_08205; 164,956 – 166,368), which encodes an argininosuccinate synthase essential for arginine biosynthesis [192], and explains the arginine auxotrophy of these strains (Supplementary Fig. 3E, dotted orange line).

Next to deletions, some genomic regions had increased read coverage, suggesting genomic amplification events took place, such as on the left chromosomal arm of strain 5.9 (Fig. 4B, orange box, Fig. 4C). This region of around 0,2 Mbp had an average coverage depth approximately 20 times higher than the rest of the chromosome (Around 4000-fold versus 200-fold read coverage). These results show that growth under hyperosmotic stress conditions induces large-scale genetic instability in *K. viridifaciens*.

A SNP analysis indicated the presence of SNPs in all strains compared to the ancestral strain. However, no SNPs were shared between all strains within a phenotypic category. Eight SNPs in coding sequences were exclusively found in all strains of cat. 2 and 3 that still retained the partial or complete megaplasmid (these SNPs were not present in the strains of cat. 2 and 3 that lost KVP1) (Supplementary Table 3). Five of these SNPs resulted in a protein substitution, one of which was located in *bldG*, which encodes a regulator of morphological differentiation and osmotic stress response in *S. coelicolor* [193] , and may contribute to a mutant phenotype.

A possible role for transposable elements in genomic rearrangements

Transposable elements (TEs) are DNA sequences that can move, or transpose, across a genome as mobile genetic elements (MGEs) [120, 194]. TEs are known to affect gene expression or cause genomic instability, the latter resulting in chromosomal rearrangements and mutations [195-197] . Genes that are commonly part of TEs (hereafter named TE-associated genes) such as those encoding transposases, are abundantly present on the megaplasmid and chromosome of *K. viridifaciens* as predicted by genome annotation software (Electronic Supplementary Data, Table 1 and 2; megaplasmid: 113 genes encoding for transposases, 15 for integrases; 3 maturases; chromosome: 43 genes encoding for transposases, 12 for integrases, one resolvase and one maturase). The frequency of occurrence of TE-associated genes is higher on the megaplasmid than the chromosome, accounting for ~ 9% of all genes compared to ~ 0.8% on the chromosome (Fig. 5A), with an especially high density of TE-associated genes on the outer right arm of the megaplasmid.

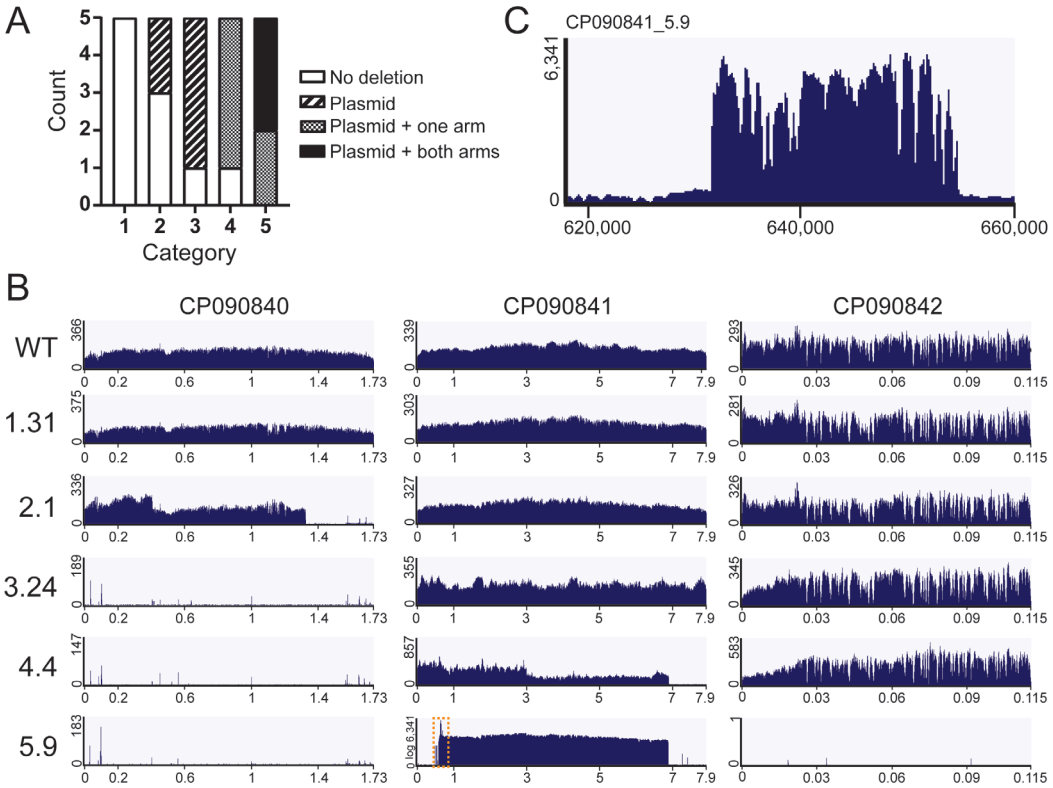


Figure 4. Growth under hyperosmotic conditions leads to genome instability in *K. viridifaciens*
(A) Number of isolates per category (1 = wild-type-like, 2 = delayed sporulation, 3 = border sporulation, 4 = heterogenous sporulation, 5 = non-developed) having intact genomes ('no deletion'), having (partially) lost the megaplasmid KVP1 ('Plasmid'), or that carry deletions on one or both of the outer chromosomal arms in addition to (partial) loss of KVP1 ('Plasmid + one arm' or 'Plasmid + both arms', respectively) based on Illumina sequencing coverage of contig CP090840 and CP090841. Five isolates were analyzed per category. See also Fig. 4B and Supplementary Fig. 3. **(B)** Mean coverage plots of Illumina sequencing reads from one isolate per category (1.31, 2.1, 3.24, 4.4 and 5.9) mapped to the reference genome of *K. viridifaciens* DSM40239 contigs CP090840 (linear megaplasmid, KVP1), CP090841 (linear chromosome) and CP090842. The x-axis indicates the genomic position in Mbp. The y-axis indicates the mean read coverage. The lack of consistent read coverage along the genome is a strong indication that part of the sequence has been lost. Note that for strain 5.9 contig CP090841 a log scale is used ('log'), and the genomic region indicated by the orange box is shown in C without log scale. See also Supplementary Fig. 3. **(C)** Mean coverage plot of Illumina sequencing reads mapped to CP090841, showing chromosome nucleotide position ~ 620,000 to 660,000 bp. Note that a ~0.2 Mbp region shows a large increased coverage compared to the surrounding region (~ 4000 versus 200 coverage).

Interestingly, TE-associated genes were found at genomic locations where the read coverage suddenly changed in strains 5.34 and 5.6. The mapped reads of strain 5.34 indicate that the strain has a truncated chromosome and a drastically varying read coverage on the megaplasmid, with two distinct regions of higher coverage (Fig. 5B, region A and B). TE-associated genes were present at both outer ends of these regions on the megaplasmid, as well as on both ends of the truncated chromosome (Fig. 5B, indicated with orange doted lines).

In addition, a roughly four-fold higher coverage was observed for a chromosomal region of strain 5.34 that contained genes encoding a maturase (a predicted reverse transcriptase/ maturase family protein found in mobile group II intron retrotransposable elements [199]), IS5 family transposase and a IS5/IS1182 family transposase (Supplementary Fig. 4A). Strain 5.6 had a similar increase in coverage on the left end of the megaplasmid for genes encoding two site-specific integrases (Supplementary Fig. 4B), suggesting genomic amplifications have occurred. Strain 5.6 also lost both outer ends of the chromosome but retained ~30% of the megaplasmid sequence. Genes encoding site-specific integrases (BOQ63_00480 & BOQ63_00490), a transposase (BOQ63_02810) and *istA* – *istB* genes (encoding an IS21 family transposase and IS21-like element helper ATPase, respectively) were present at the left and right borders of the megaplasmid region, while *istA* and *istB* were also found at the breakpoint of the right chromosomal arm (Fig. 5C). As it is unlikely that such a small region of the megaplasmid can still replicate autonomously, it may have recombined with the chromosome using the *istA-istB* locus.

Previous research indicated that a protoplast revertant of *K. viridifaciens* with a bald phenotype, strain B3.1, contained large genomic deletions [198]. Interestingly, the read coverage of strain B3.1 is similar to strain 5.34, with roughly the same region of the megaplasmid retained and the right chromosomal arm truncated at exactly the same position (Fig. 5D). In both strains, TE-associated genes were present at the location where the read coverage dropped, further suggesting that these genes may be involved in recombinations and/or deletions.

To study recombination events between the megaplasmid and the chromosome, whole-genome alignment (WGA) of different *K. viridifaciens* isolates was performed. *Streptomyces viridifaciens* ATTC11989 (accession no. CP023698 originating from KCTC) is listed as the same strain as *K. viridifaciens* DSM40239 (accession no. CP090840-42 from DSMZ), but its genome only consists of a linear chromosome, and lacks a megaplasmid. WGA indicated that a 341 kbp region homologous to the megaplasmid KVP1 is present in the right chromosomal arm of strain ATTC11989 (Supplementary Fig. 5A). It is interesting to note the presence of a gene encoding a IS3 family transposase (CP971_32970) on the chromosome of ATTC11989, next to the megaplasmid-homologous region (Supplementary Fig. 5B). This suggests that recombination events between the megaplasmid and the chromosome can indeed occur in *K. viridifaciens*, possibly mediated via TEs, and can presumably be triggered by hyperosmotic stress.

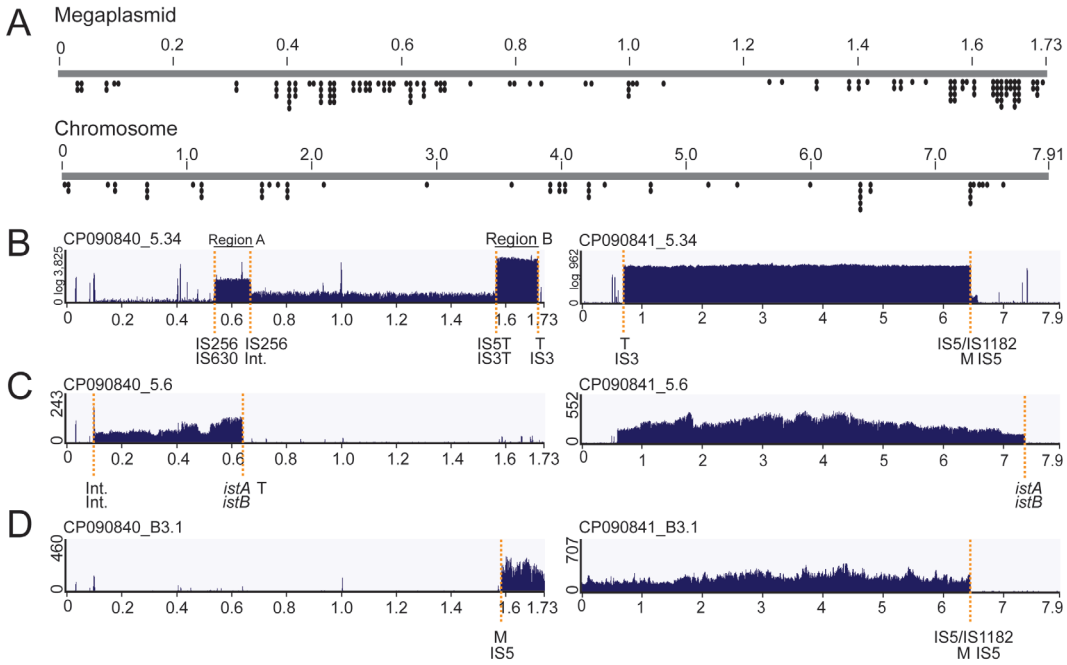


Figure 5. A role for transposable elements in genome instability of *K. viridifaciens*

(A) Location of predicted TE-associated genes on the megaplasmid (CP090840) and chromosome (CP090841) of *K. viridifaciens* DSM40239. Numbers indicate genomic position in Mbp. No TE-associated genes were present on CP090842. Each dot represents a gene encoding a transposase, integrase, resolvase, maturase or *istB* near that genomic position. Stacked dots indicate the presence of multiple genes near that location. The occurrence of TE-associated genes is higher on the megaplasmid (135) than on the chromosome (62) despite its smaller size (1.7 Mbp versus 7.9 Mbp). Note the high abundance of TE-associated genes around nucleotide position 1.6 – 1.7 Mbp on the megaplasmid. See also Electronic Supplementary Data, Table 1 and 2. **(B – D)** Mean coverage plots of Illumina reads of isolate 5.34 (B), 5.6 **(C)** and protoplast revertant B3.1 **(D)** from Ramijan *et al.*, (2020)[198] mapped to the reference genome of *K. viridifacien*s, showing CP090840 and CP090841. The x-axis indicates the genomic position in Mbp. The y-axis indicates the mean read coverage. Note the y-axis log scale in **(B)**. Orange dotted lines indicate the presence of TE-associated genes within 5 kb from the position where read coverage changes. Abbreviations indicate presence of predicted genes encoding the following proteins: T = transposase, M = maturase, Int. = integrase (site-specific or integrase-core domain containing protein), IS3 = IS3 family transposase, IS5 = IS5 family transposase, IS5/IS1182 = IS5/IS1182 family transposase, IS256 = IS256 family transposase or indicate the presence of *istA*/*istB* genes. See also Supplementary Table 4 and Supplementary Fig. 4.

Isolates with a mutant phenotype are not resistant against osmotic stress

Genetic changes such as mutations, gene deletions and genome rearrangements, including those mediated by TEs [200], can lead to altered phenotypes and adaptation of bacterial species to environmental stress [201]. We hypothesized that the generation of colony heterogeneity in *K. viridifaciens* could function as an adaptation to withstand osmotic stress. To test whether the isolates with an altered phenotype had adapted to better cope with osmotic stress, growth of five candidates per category (except the isolates with a bald phenotype from cat. 5 due to their apparent inability to produce spores) was compared to that of the parental strain in liquid cultures using a 'BioLector' (m2p-labs) microbioreactor. Strains were grown in LPB medium with and without sucrose at 800 rpm orbital shaking for 72 h and 25 h, respectively, during which growth was monitored using the optical backscatter signal every 15 min.

Without sucrose, the isolates behaved similarly to the parental wild-type strain (Fig. 6, No Sucrose). In the presence of sucrose, strains showed delayed growth and reached a lower final biomass than the parental strain (cat. 1 and 2) or behaved similarly (cat. 3 and 4) to the parental strain (Fig. 6, Sucrose). Furthermore, strains from all categories retained the response to form S-cells upon growth under hyperosmotic stress conditions (Supplementary Fig. 6). These results indicate that isolates with a mutant phenotype are still susceptible to hyperosmotic stress, but in some cases behave differently than the wild-type strain under these conditions.

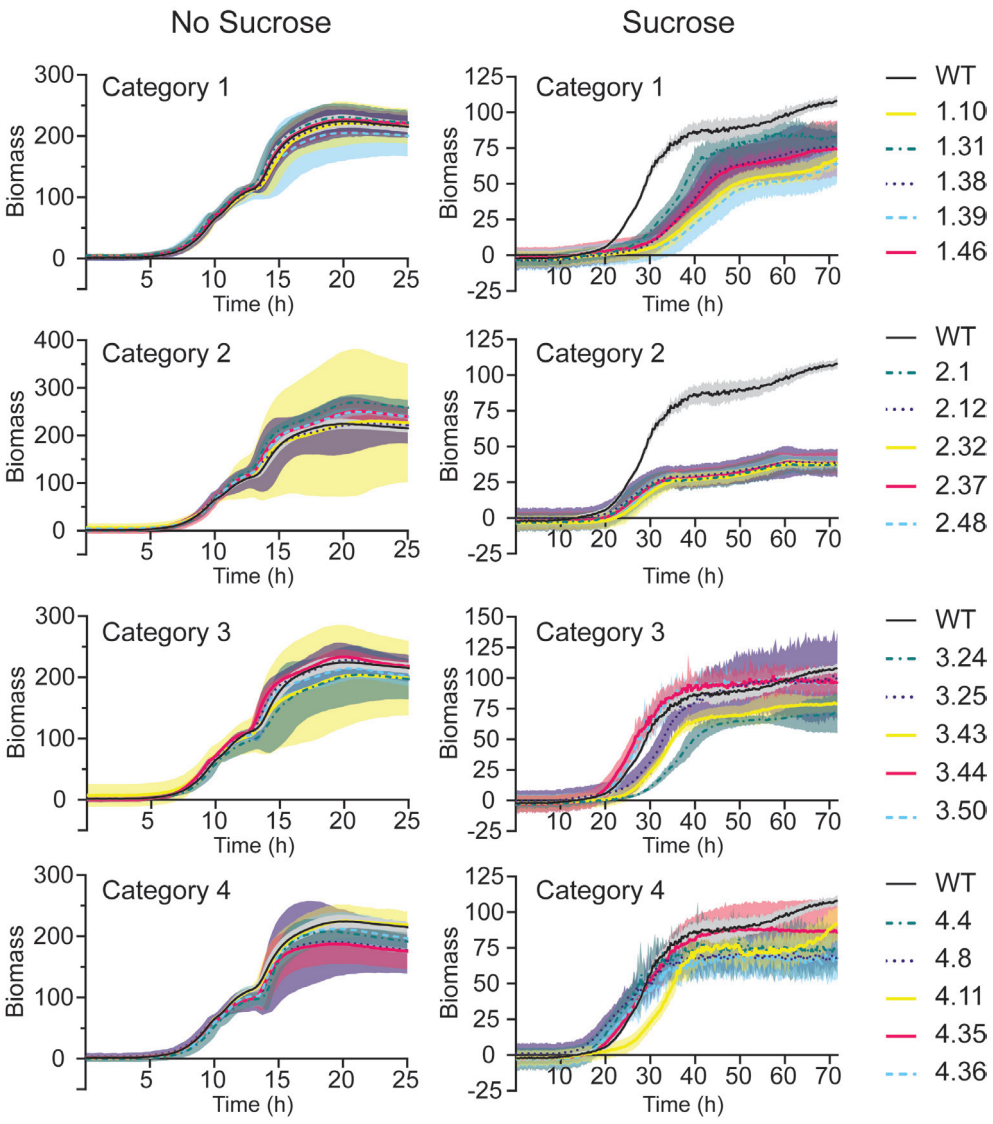


Figure 6. Hyperosmotic stress-induced mutants are not resistant to osmotic stress

K. viridifaciens wild-type (WT) and five isolates from the sporulating categories 1 to 4 were grown in LPB medium with (Sucrose) and without (No Sucrose) 22% (w/v) sucrose in a microbioreactor at 800 rpm and 30 °C. Biomass, measured using backscattered light (arbitrary units), was measured every 15 min for ~25 h (No Sucrose) and ~72 h (Sucrose). Lines indicate the mean of four biological replicates, shaded areas indicate the 95% confidence interval. Note that in general, the isolates grow similarly or slower than the wild-type in the presence or absence of sucrose.

Discussion

Bacteria are often exposed to fluctuating environments and need to adapt to cope with diverse stresses. Growth of *K. viridifaciens* under hyperosmotic stress conditions leads to the formation of phenotypically diverse colonies [15] . In this study, we isolated and characterized these diverse colonies to study the cause and consequences of this heterogeneity. We discovered that *K. viridifaciens* is prone to genetic instability under hyperosmotic stress conditions. The genome of *K. viridifaciens* contains a high number of genes associated with TEs, which often flank the amplified or deleted genomic regions. As TEs can be activated by stress conditions, this suggests a role for MGEs in genomic rearrangements in response to hyperosmotic conditions.

Genetic instability is widespread in *Streptomyces* and has been studied for decades. The formation of mutant phenotypes (as observed by changes in pigment synthesis, arginine and antibiotic biosynthesis, antibiotic resistance and sporulation) is coupled to large deletions of up to 2 Mb at the chromosome ends [202-205]. In addition, deletions are often accompanied with tandem amplifications (often up to 100-500 copies), rearrangements and/or circularization of the linear chromosome [190, 206-210]. Various stress treatments such as UV-mutagenesis, DNA intercalating agents or storage of spores at – 20 °C increase genetic instability in streptomycetes [188]. This is in line with our findings in which growth of *K. viridifaciens* under osmotic stress conditions led to colonies with a mutant phenotype, of which the majority contained genomic rearrangements, such as loss of the megaplasmid, large deletions of chromosomal arms as well as amplified regions.

The high number of TEs located in the terminal regions of *Streptomyces* chromosomes indicate that TEs may play a major role in genetic instability in these bacteria [211-213]. TEs can be activated by a wide range of abiotic and biotic environmental stressors in eukaryotes and prokaryotes, such as osmotic stress, UV-irradiation, oxidative stress, increased temperature, growth stage as well as conjugative interactions between prokaryotes (as reviewed by Negi *et al*., (2016) [214] and Vandecraen *et al*., (2017) [215]). TEs can move around the genome by non-replicative (cut-and-paste), replicative (copy-and-paste) or alternative transposition mechanisms (formation of hybrid elements) [197]. Repair of double-strand breaks (DSBs) formed by the excision of TEs from the DNA can lead to mutations, deletions, loss of chromosome arms and genome rearrangements [216]. In addition, chromosomal rearrangements can be caused by homologous recombination (HR) between repetitive sequences such as TEs [197]. TE activity or recombination in response to osmotic stress may have led to deletions of outer chromosomal arms and/or rearrangements between the megaplasmid KVP1 and the chromosome. Hyperosmotic stress can thus be a driver for major genetic changes in streptomycetes that are often exposed to such conditions in soil habitats.

The colonies obtained from S-cell filtrate grown on MYM medium originated from a mix of walled and wall-less cells. We propose that osmotic stress induces genomic rearrangements already in hyphal cells, resulting in diverse genotypes. These altered chromosomes end up inside S-cells produced from these hyphae. The S-cell formation process may further contribute to loss of genetic material, for example if the megaplasmid is not extruded in the S-cell. Subsequently, the mixture of genetically diverse cells contributes to the heterogenous colonies observed grown on MYM medium (Fig. 7). Similar genomic rearrangements were observed in colonies originating from *K. viridifaciens* protoplasts, which was linked to replicative transposition between the chromosome and megaplasmid [198]. These protoplasts were generated from walled cells grown under hyperosmotic conditions, further suggesting that osmotic stress leads to genomic rearrangements.

The increased CFU of S-cell filtrate as well as its sensitivity to water observed on R5 and LPMA medium indicate that S-cells can revert. The process of S-cell reversion has also been captured using timelapse imaging of S-cells in LPB medium covered with an R5 medium agar pad [15]. There are several factors that may contribute to the differences in CFU obtained across the solid medium types when plating S-cell filtrate. As the increased sucrose content of the medium positively correlated with CFU, this suggests that S-cells require an osmotically protective medium to revert. Additionally, rapid growth of walled or wall-less cells on MYM medium may inhibit colony formation by remaining S-cells in a process called auto-inhibition. Auto-inhibition is the inhibition of protoplast regeneration by non-protoplasted cells or by earlier regenerated protoplasts, as observed for some *Streptomyces* species [185, 217, 218]. Auto-inhibition in *Bacillus subtilis* protoplasts is presumably caused by secretion of an autolysin-like protein that targets cell wall synthesis [219]. Reduction of auto-inhibition during protoplast regeneration is achieved by use of dryer solid medium, which may retard the diffusion of auto-inhibitory substances, as well as minimize protoplast lysis due to the agar containing less water [185]. These factors may explain the enhanced CFU of S-cell filtrate on LPMA medium under dryer conditions, which likely increases S-cell reversion.

Isolates with a mutant phenotype grow similarly or slower than the parental strain during growth under hyperosmotic stress conditions, and still form S-cells. Further experiments are required to determine whether this altered growth can improve survival under hyperosmotic stress, or whether the mutant strains have any other benefits. It has been shown previously that formation of mutant cells by genomic instability is linked to a division of labor strategy in *S. coelicolor*, in which antibiotic production is performed in part of the colony by less-fit mutants, while the colony-wide fitness remains intact [191]. In addition, diversity can make species and populations more robust to environmental stress and plays a key role in the evolution of biological systems [220].

In summary, this study shows how growth under hyperosmotic stress induces drastic phenotypic changes and genetic instability in *K. viridifaciens,* indicating how environmental stress conditions can contribute to bacterial evolution. It would be interesting to study whether hyperosmotic stress induces genetic instability in other filamentous actinobacteria, and whether this response correlates with their ability to form cell wall-deficient cells, which may act as an indicator of osmotic stress. The generation of genomic diversity via induction of MGEs may be used as a tool to alter the production of bioactive molecules in streptomycetes for antibiotic research.

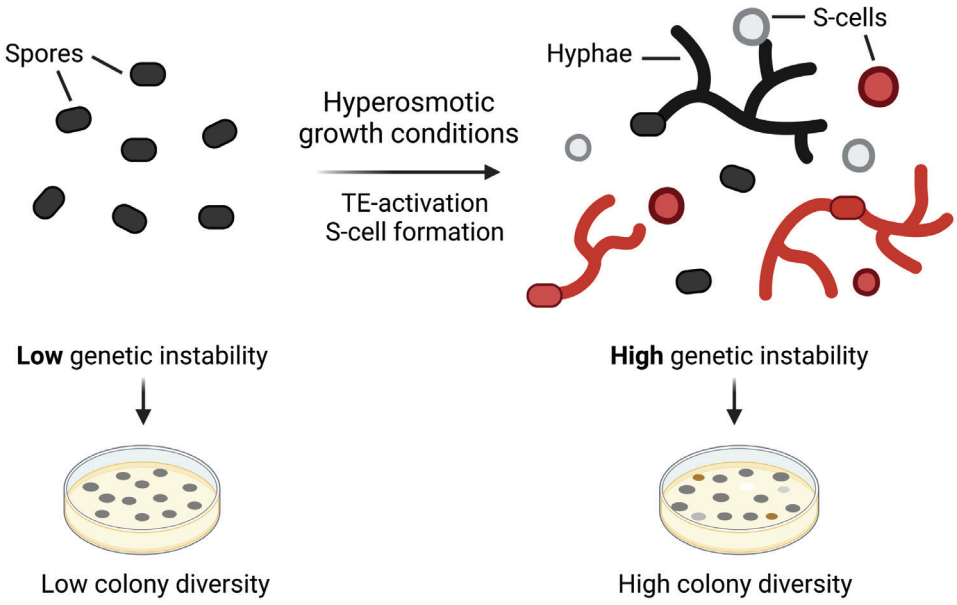


Figure 7. Model for genetic instability in *K. viridifaciens* observed during growth under hyperosmotic conditions.

Growth under hyperosmotic conditions leads to genetic instability and a high colony diversity in *K. viridifaciens*. After inoculation of *K. viridifaciens* spores in LPB medium containing 22% (w/v) sucrose, the spores germinate and form hyphae. We propose that the hyperosmotic stress conditions that trigger the formation of wall-less S-cells also activate transposable elements (TEs) encoded on the genome. Activation of TEs results in hyphal cells with a high genetic instability (depicted in red) as exemplified by genomic rearrangements, deletions and/or amplifications. S-cells extruded from these hyphal cells inherit this genetically heterogenous DNA. The loss of genetic material may be further enhanced during the S-cell extrusion process if only part of the genome ends up in these cells (e.g. without the megaplasmid). Upon growth of cells exposed to hyperosmotic stress on MYM medium, the mixture of wild-type and non-wild-type cells results in heterogenous colonies as compared to colonies arising from wild-type spores. Cells with wild-type genomes are indicated in black/grey, and cells with aberrant genomes are indicated in red. Note that the percentage of cells with genetic instability for each cell type is unknown. Created with BioRender.com.

Materials and methods

Bacterial strains and culture conditions

Bacterial strains were grown at 30 °C in liquid cultures at a density of 1×10^6 CFU ml^{-1} while shaking at 100 rpm during S-cell induction. For other applications, cultures were shaking at 200 rpm. To induce sporulation, *K. viridifaciens* DSM40239 [221] was grown confluently on maltose-yeast extract medium (MYM) for 72 – 96 h. MYM medium [222] contained 4 g maltose, 4 g yeast extract, 10 g malt extract, 2% (w/v) Iberian Agar dissolved in a one L 1:1 mix of tap water and dH_2O. After autoclaving, R5 trace elements [223] were added to a final concentration of 0.08 mg $ZnCl_2$, 0.4 $FeCl_3 \times 6H_2O$, 0.02 $CuCl_2 \times 2H_2O$, 0.02 $MnCl_2 \times 4H_2O$, 0.02 $Na_2B_4O_7 \times 10H_2O$ and 0.02 $(NH_4)_6MO_7O_{24} \times 4H_2O$. Spores were harvested by resuspension in MilliQ and filtration through cotton wool. Spores were stored in 20% (v/v) glycerol at – 80 °C until use. Mycelial biomass from liquid cultures was obtained by growth of spores or mycelium in tryptone soy broth medium [223] containing 30 g L^{-1} Tryptone Soy Broth powder and 10% (w/v) sucrose (TSBS) in a coiled flask at 200 rpm for 24 h, or until sufficient biomass was present for non-sporulating strains. S-cell formation was induced by growth in liquid L-phase broth (LPB) [15], containing 0.15% yeast extract, 0.15% oxoid malt extract, 1.5% oxoid tryptone soy broth powder, 0.25% bacto-peptone, 0.5% glucose (all w/v), 0.64M sucrose and 25 mM $MgCl_2$, at 100 rpm for 48 h or on L-phase medium agar (LPMA) unless stated otherwise. LPMA contained 20% (w/v) sucrose, 0.5% (w/v) glucose, 0.5% (w/v) yeast extract, 0.5% (w/v) peptone, 0.01% (w/v) $MgSO_4 \times 7H_2O$, 0.75% (w/v) Iberian agar [15]. After autoclaving, LPMA medium was supplemented with 5% (v/v) horse serum and $MgCl_2$ to a final concentration of 25 mM. LPB medium cultures were filtered by passing the culture through a sterile EcoCloth™ filter followed by a 5 µm Isopore™ membrane filter to obtain S-cell filtrate [15].

The effect of prolonged drying of solid medium on the efficiency of S-cell reversion was tested by applying agar medium in 8.5 cm round petri dishes and solidifying for 0.5, 2, 4 and 8 h in a Telstar Horizontal Laminar Air flow Cabinet. Alternatively, 1 ml samples were applied on LPMA agar plates and subsequently dried for 0.5, 1, 2 and 3 h in a biological safety cabinet (CleanAir, Baker). Around 1 ml dilutions of S-cell filtrate from two separate cultures were plated in duplo, followed by incubation for 168 h at 30 °C.

Viability of S-cells on solid medium

To determine the viability of S-cells on solid medium, S-cells were filtered from three separate cultures ($n = 3$), diluted ten times in MilliQ or 22% sucrose, gently mixed and incubated for 5 min at room temperature. Subsequently, cells were diluted ten times in LPB medium before plating 1 ml dilutions on MYM, R5 and LPMA medium in triplicate. R5 medium [223] contained 10.3% sucrose, 0.25 g K_2SO_4, 10.12 g $MgCl_2 \times 6H_2O$, 1% glucose, 0.1 g Difco Casaminoacids, R5 Trace element solution, 5 g Difco yeast extract, 5.73 g *N*-tris(hydroxymethyl)methyl-2-

amino ethane sulfonic acid (TES) and 22 g Difco agar in 1 L distilled water. After autoclaving, R5 medium was supplemented with 1 ml 0.5% (w/v) KH_2PO_4, 0.4 ml 5M $CaCl_2 \times 2H_2O$, 1.5 ml 20% (w/v) L-proline and 0.7 ml 1M NaOH per 100 ml agar [223]. Spores of *K. viridifaciens* were diluted in LPB medium and plated in the same manner ($n = 8$). Plates were incubated at 30 °C for 168 h before determining the CFU ml^{-1}.

Quantification of mutant phenotype frequency

S-cell filtrate or spores from *K. viridifaciens* were diluted in LPB medium or MilliQ, respectively, and plated on MYM medium. After 120 h incubation, single colonies from separate MYM agar plates were picked until at least 100 colonies were obtained (108 for S-cell filtrate and 120 for spores). These colonies were patched on MYM medium and incubated until sporulation, or sufficient growth, was observed (96 – 168 h incubation). Each patch was then grown individually on MYM agar plates and imaged every 72 h and 168 h using the Epson Perfection V600 Photo scanner. As a control, *K. viridifaciens* wild-type was grown in a similar manner.

Stability of mutant phenotypes

Spores (around 1×10^7 CFU) or 50 µl mycelial stock (for non-sporulating strains) from three isolates per category (cat. 1 – 5), as well as from the *K. viridifaciens* wild-type strain, were streaked on MYM medium. These cultures were called Generation 1 (G1). Plates were incubated and imaged using an Epson Perfection V600 Photo scanner after 72 h and 168 h. Spores or mycelium from the biomass patch of G1 were streaked on fresh MYM medium, resulting in Generation 2 (G2), and were incubated and imaged in the same way. This procedure was repeated up to Generation 5 (G5). Morphology of the mycelial biomass regarding sporulation, aerial hyphae formation and production of coloured secondary metabolites was manually compared to *K. viridifaciens* wild-type of G1.

Growth in microbioreactors

Growth experiments were performed using the BioLector (m2p-labs) micro-fermentation system using 48-wells plates (MTP FlowerPlate, Beckman Coulter Life Sciences) closed with a gas-permeable, evaporation-reducing seal. Strains capable of sporulation (cat. 1 to 4) were grown in parallel with *K. viridifaciens* wild-type in 1 ml LPB medium cultures with or without 22% sucrose for ~72 h or 25 h, respectively. Spore stocks were adjusted to an equal concentration in 20% (v/v) glycerol before diluting 100 times to a density of 1×10^7 CFU ml^{-1} in the final liquid medium. Cultures were grown at 30 °C with 800 rpm orbital shaking and 85% humidity. Growth was monitored by measuring the optical backscatter signal every 15 mins and plotted with biomass gain set to 5. Medium without spores was used as a blank, and was subtracted from all biomass measurements. The location of each strain in the wells plate

was randomized, and any remaining empty wells were used for additional *K. viridifaciens* wild-type cultures. Experiments were carried out in quadruplicate.

Auxotrophy assay

Spore stocks or mycelial stocks were washed three times with MilliQ. 5 µl spots containing 5×10^6 spores or mycelial biomass were incubated for 96 h on minimal medium (MM)[223] containing 1% glucose, 0.5% mannitol, 0.05% L-Asparagine, 0.05% KH_2PO_4, 0.02% $MgSO_4 \times 7H_2O$, 0.001% $Fe_2SO_4 \times 7H_2O$, and 2% Difco agar (all w/v). Where necessary, MM was supplemented with 0.5% Bacto™ Casamino Acids or 0.0037% L-arginine (Merck) after autoclaving. As a control, wild-type *K. viridifaciens* was grown using spores or mycelium as inoculum.

gDNA isolation

For gDNA isolation, strains were grown for 24 h or up to 72 h in TSBS medium for sporulating and non-sporulating strains, respectively, before harvesting the biomass for phenol:chloroform extraction [223]. In brief, the biomass pellet was washed with 10.3% (w/v) sucrose before incubation in P-buffer [186, 224] containing 20 mg ml^{-1} lysozyme (Lysozyme from chicken egg white, 70,000 U mg^{-1}) at 37 °C. After 1 – 2 h incubation, cells were pelleted and the biomass was resuspended in 10.3% sucrose containing 0.01 M ethylenediamine tetra acetic acid (EDTA) pH = 8, following addition of 10% sodium dodecyl sulfate (SDS) for cell lysis. The suspension was mixed 1:1 with 5 M NaCl before phenol:chloroform (pH = 8, 1:1 ratio) extraction was performed. Nucleic acids were precipitated using isopropanol and sodium acetate at – 20 °C. The gDNA was resuspended in Tris-EDTA buffer before removal of RNA using RNAse A (Thermo Fisher) and removal of proteins using Proteinase K (Qiagen). gDNA was extracted using phenol:chloroform, precipitated using absolute ethanol and sodium acetate at – 20 °C. Precipitated gDNA was resuspended in MilliQ. DNA quality was determined using gel electrophoresis and a NanoDrop 2000 spectrophotometer. The Quant-iT dsDNA Assay Kit (Sigma-Aldrich, Q33120) in combination with the Spark™ 10M multimode microplate reader (Tecan) with excitation 488 nm and emission 550 nm was used to accurately quantify the DNA concentration.

Kitasatospora viridifaciens DSM40239 reference genome

Comparison of Sanger sequencing results of PCR products from *K. viridifaciens* DSM40239 gDNA with the available reference genome (accession number MPLE00000000) as well as with *Streptomyces viridifaciens* ATTC11989 (accession number CP023698.1) indicated the presence of sequencing errors (mutations, short deletions) in the current genome version. Therefore, to obtain a high-quality reference genome sequence, *de novo* whole-genome sequencing was performed by BaseClear B.V. using Oxford Nanopore technology combined

with Illumina sequencing. Three contigs were assembled *de novo* and corrected using Flye version 2.8.3. In addition, paired-end Illumina sequencing reads were generated using Illumina NovaSeq600 or MiSeq system. FASTQ read sequence files were generated using bcl2fastq version 2.20 (Illumina). Initial quality assessment was based on data passing the Illumina Chastity filtering, and reads containing PhiX control signal were removed using a BaseClear inhouse filtering protocol. Reads containing (partial) adapters were clipped up to a minimum read length of 50 bp. The second quality assessment was performed using the FastQC quality control tool version 0.11.8, with an average FastQC quality higher than 35 (Q35). Polishing of the reference genome was performed based on ONT reads using Medaka version 1.2.0 followed by polishing based on Illumina reads using Pilon version 1.23. Genome annotation was performed via the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) [225] v6.0. The genome was submitted to NCBI with the name *Streptomyces viridifaciens* strain DSM40239 under accession numbers CP090840, CP090841 and CP090842 corresponding to the megaplasmid KVP1, the linear chromosome and a separate contig, respectively.

Whole-genome sequencing and variant analysis

Illumina-sequencing of gDNA was performed by BaseClear B.V. as described previously. Paired Illumina reads were trimmed in Geneious Prime v. 2022.1.1 software using the BBduk v. 38.84 plugin with a minimum quality of 20 set for 'Both Ends' to trim low quality reads, and discarding short reads under 20 bp. Remaining paired reads were mapped against the reference genome of *K. viridifaciens* DSM40239 with accession numbers CP090840, CP090841 and CP090842 using the 'Geneious Mapper' tool set to Medium-Low Sensitivity/Fast and mean insert size of 400 bp. Variant analysis (SNP calling) was performed using the 'Find Variations/SNPs' option in Geneious, with a minimum coverage set to 10 and a minimum variant frequency of 0.7. Variants called in the wild-type were removed from variants called in the mutant strains using the 'VCF substraction' tool from VCFtools [226] publicly available at Galaxy Pasteur [227] (Galaxy version 1.0.0). In addition, variants with a strand bias of 90% or higher were removed (strand-bias was calculated as the percentage of variant reads on one strand, giving a value between 50 and 100%).

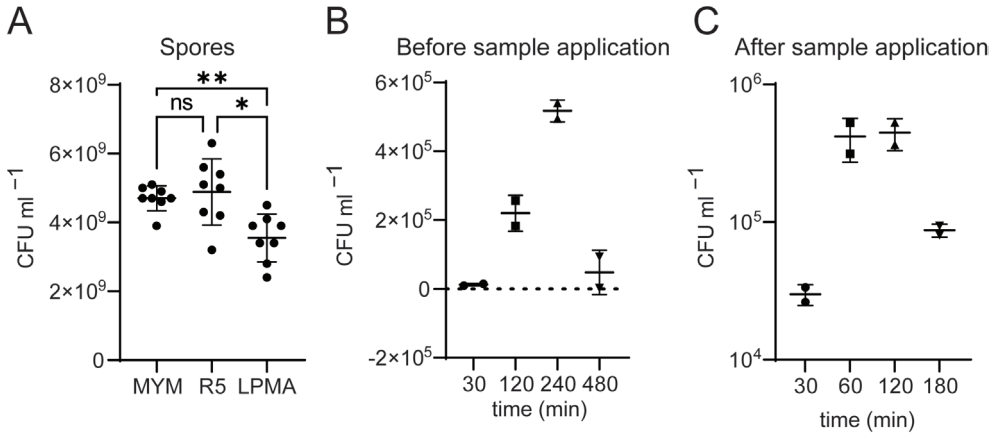
Whole-genome alignment

Whole-genome alignment was performed with MAUVE [228] (version 20150226) using default settings. The accession number of *Streptomyces viridifaciens* ATTC11989 was CP023698.

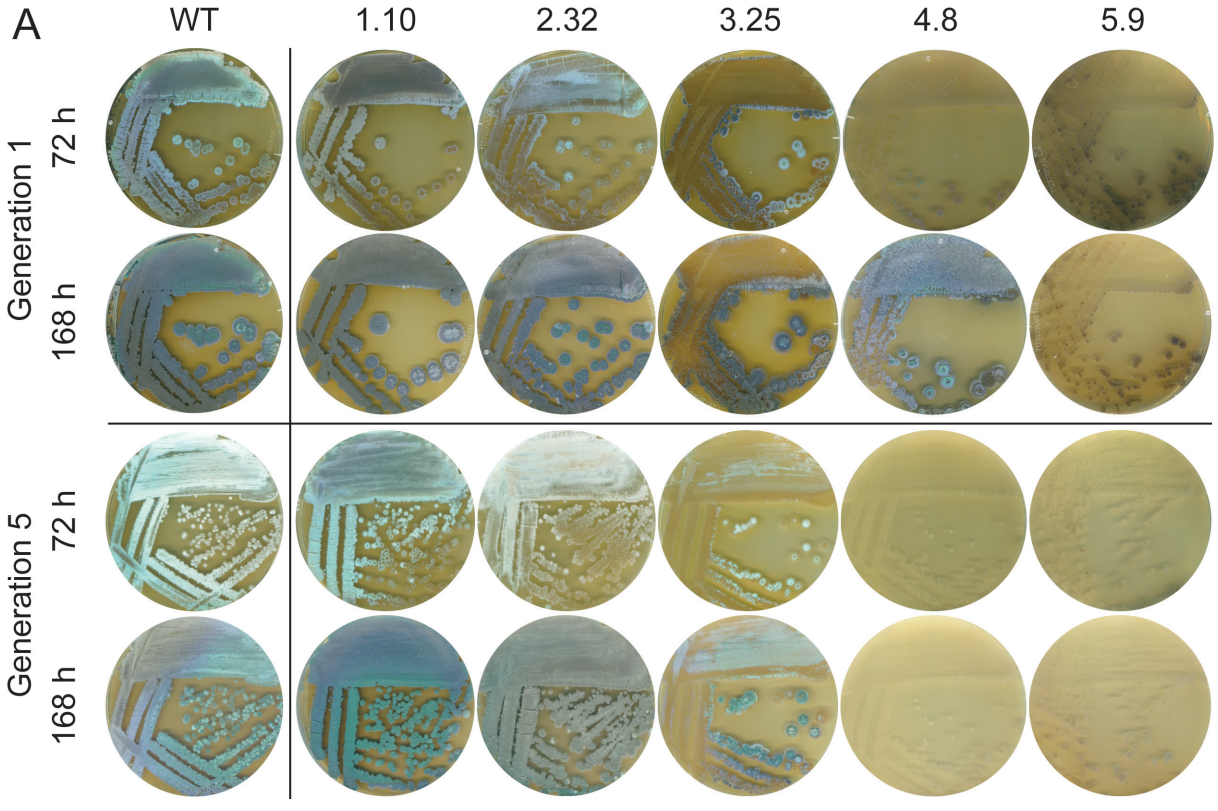
Microscopy

Formation of S-cells in liquid-grown cultures was assessed by using a Zeiss LSM 900 confocal microscope with Airyscan 2 module and Zeiss Zen 3.1 software (blue edition, Carl Zeiss Microscopy GmbH). Strains were grown for 48 h in LPB medium to induce S-cell formation, with the exception of cat. 5, which was grown up to 240 h due to delayed growth. Presence of nucleic acids was confirmed by incubation with SYTO 9 for 30 min prior to visualization using 488 nm excitation and 450-565 nm emission. Cell membranes were visualized using SynapseRed C2M at a final concentration of 40 μg ml^{-1} using excitation and emission wavelengths of 488 nm and 640 – 700 nm, respectively. Cells were placed on a microscopy slide topped with a glass coverslip before imaging, or placed in an μ-Slide 8 Well Ibidi® slide topped with an agar pad (LPMA medium) to clearly image mycelial cells. The concentration of S-cells in S-cell filtrate was estimated using a 0.1 mm deep Neubauer-improved Counting Chamber (Marienfeld) (Supplementary Table 1) and an Axio Lab A1 Microscope (Zeiss) equipped with an Axiocam 105 color camera (Zeiss) and ZEN 2.5 software (blue edition, Carl Zeiss Microscopy GmbH).

Supplementary Data

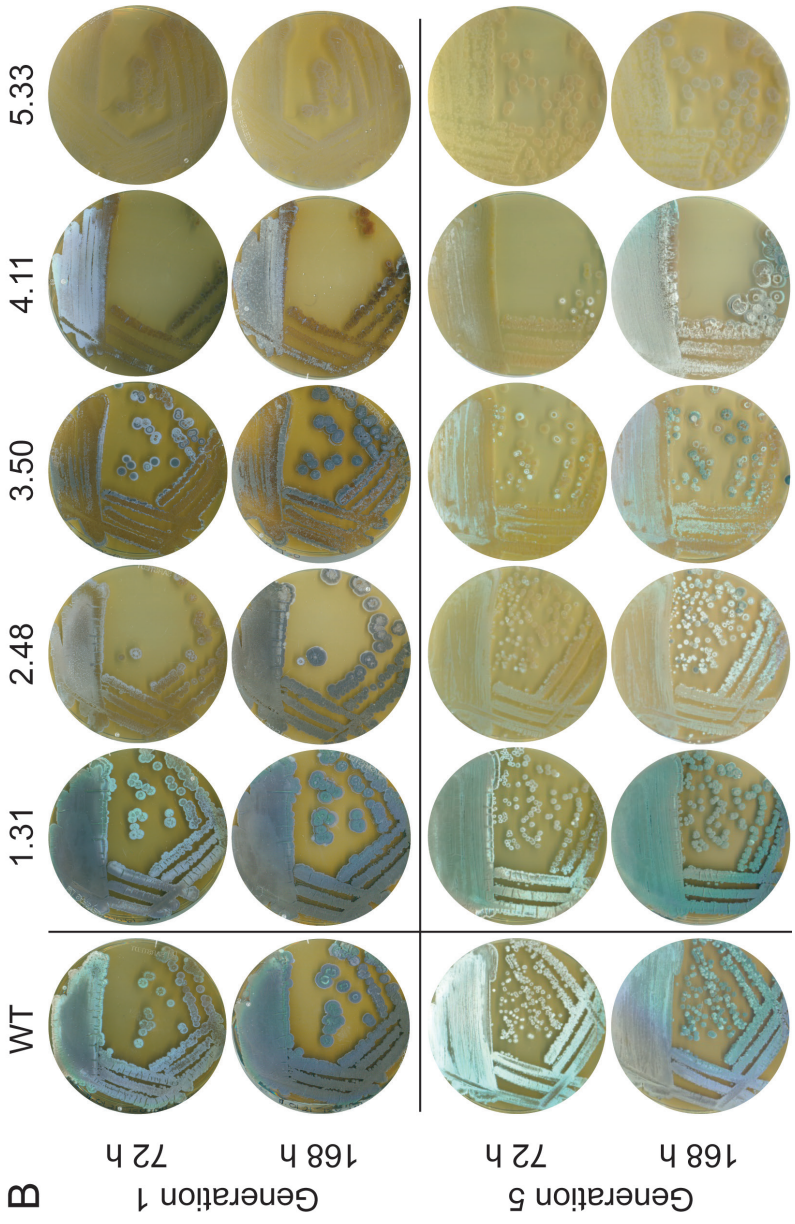


Supplementary Figure 1. Viability of spores and S-cells under different solid medium conditions
(A) Mean colony forming units (CFU) per ml of *K. viridifaciens* spores grown on MYM medium, R5 medium or LPMA medium. CFUs were analyzed after 168 h incubation at 30 °C. Asterisks indicate $P < 0.05$ (*) or $P < 0.01$ (**) ($n = 8$ replicates, Welch ANOVA, F (2, 12.10) = 8.828, Dunnett's T3 post hoc test, $P = 0.937$ (MYM – R5), $P = 0.0048$ (MYM – LPMA) and $P = 0.0207$ (R5 – LPMA)). Significantly lower CFU ml^{-1} is obtained by growing spores on LPMA medium as compared to MYM and R5 medium. Individual data points shown, error bars indicate the standard deviation. **(B – C)** Mean CFU ml^{-1} of *K. viridifaciens* S-cell filtrate after growth on LPMA medium that was dried before **(B)** or after **(C)** sample application. Agar was dried for 30 (standard), 120, 240 or 480 min before sample application, or dried for 30, 60, 120 or 180 min after sample application (before sample application the agar was dried for the standard 30 min). CFUs were analyzed after 168 h incubation at 30 °C. Shown is the mean of $n = 2$ biological replicates, with individual datapoints representing the average of two technical replicates. Error bars indicate the standard deviation.

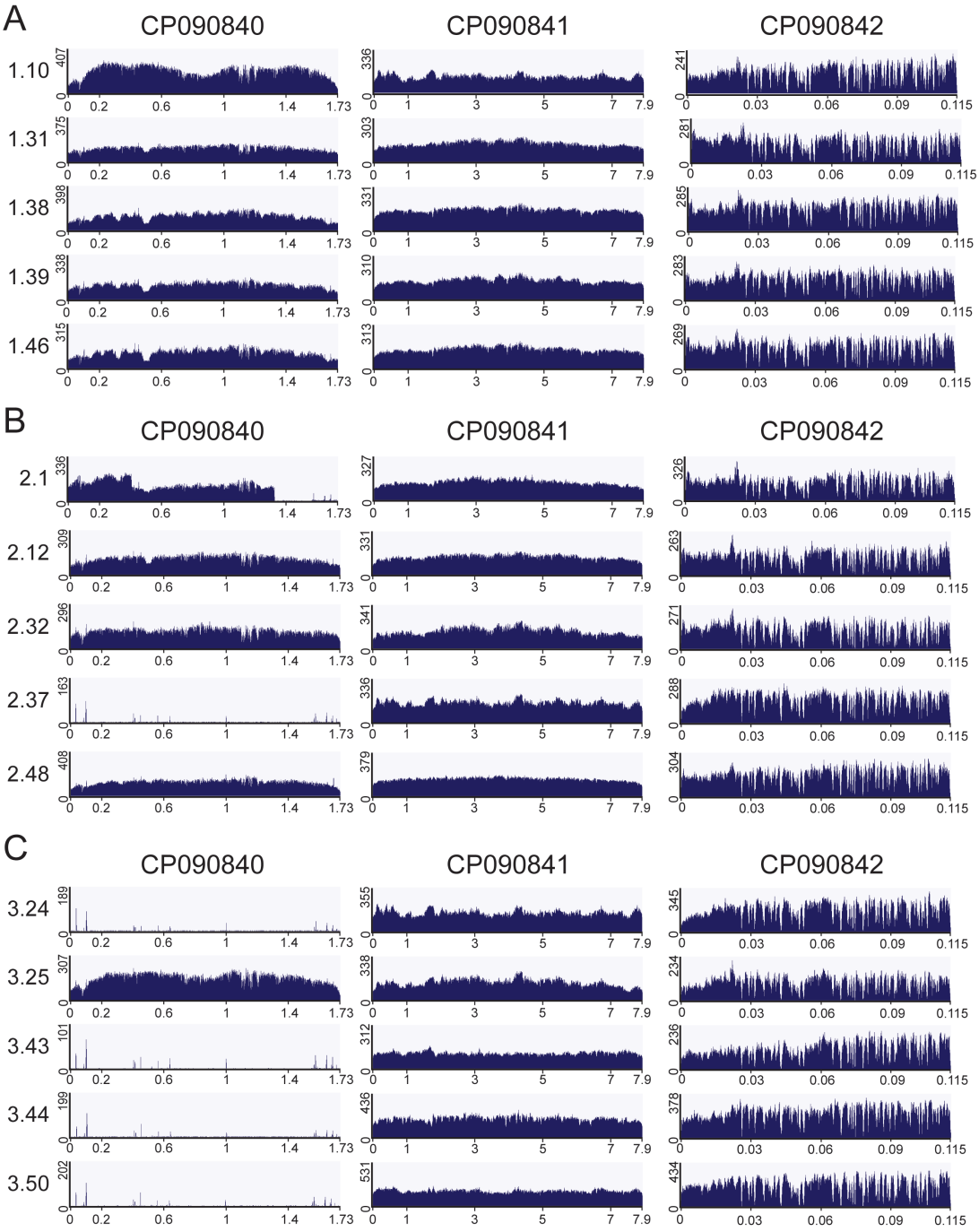


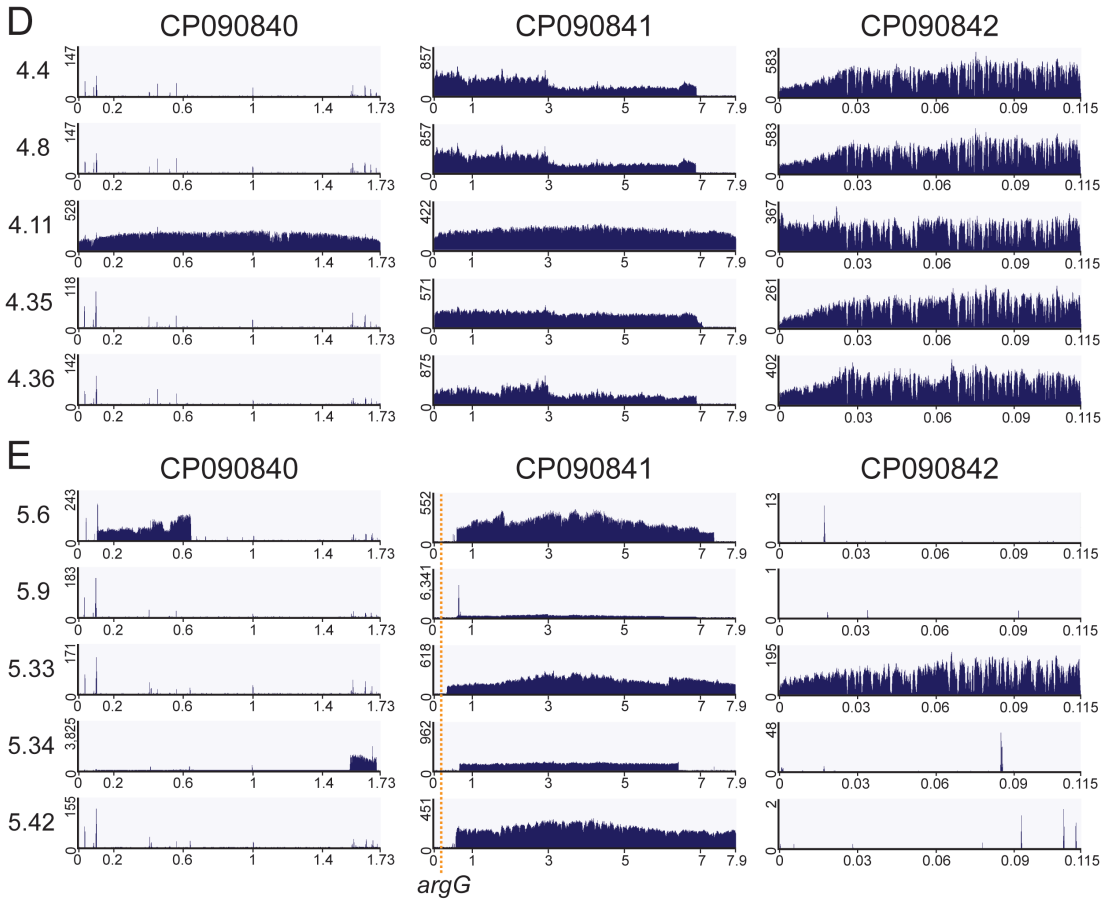
Supplementary Figure 2. Mutant phenotype of isolates does not revert to wild-type

Phenotypic stability of *K. viridifaciens* wild-type and two isolates per category at Generation 1 and 5, imaged after growth on MYM medium for 72 and 168 h at 30 °C. The tested isolates 1.10 (cat. 1), 2.32 (cat. 2), 3.25 (cat. 3), 4.8 (cat. 4) and 5.9 (cat. 5) are given in **(A)** and isolates 1.31 (cat. 1), 2.48 (cat. 2), 3.50 (cat. 3), 4.11 (cat. 4) and 5.33 (cat. 5) are given in **(B)**. Note that none of the strains with a mutant phenotype regain the ability to sporulate as wild-type at G1. See also Fig. 2.



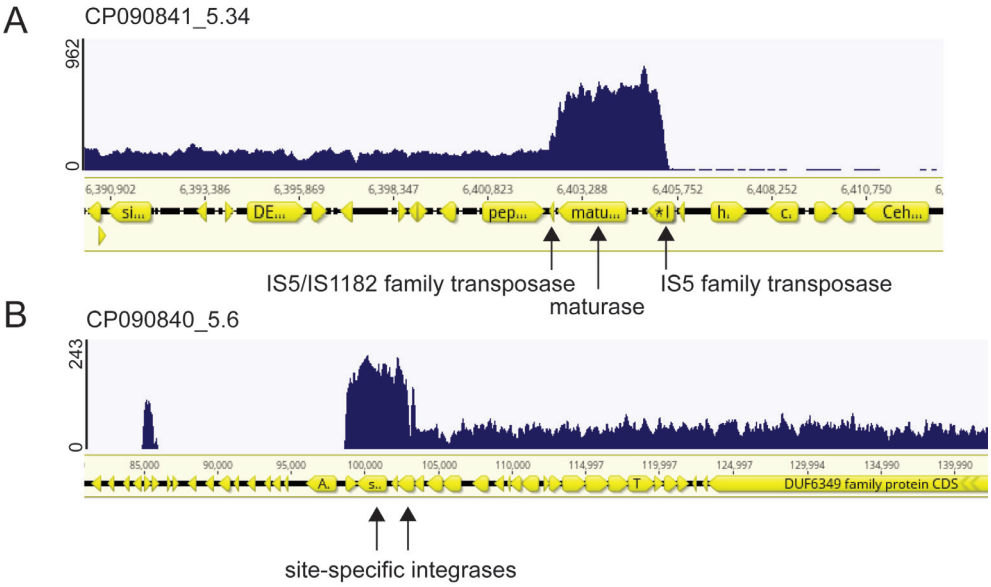
Supplementary Figure 2. (Continued)


A
CP090840
CP090841
CP090842
1.10
1.31
1.38
1.39
1.46
B
CP090840
CP090841
CP090842
2.1
2.12
2.32
2.37
2.48
C
CP090840
CP090841
CP090842
3.24
3.25
3.43
3.44
3.50

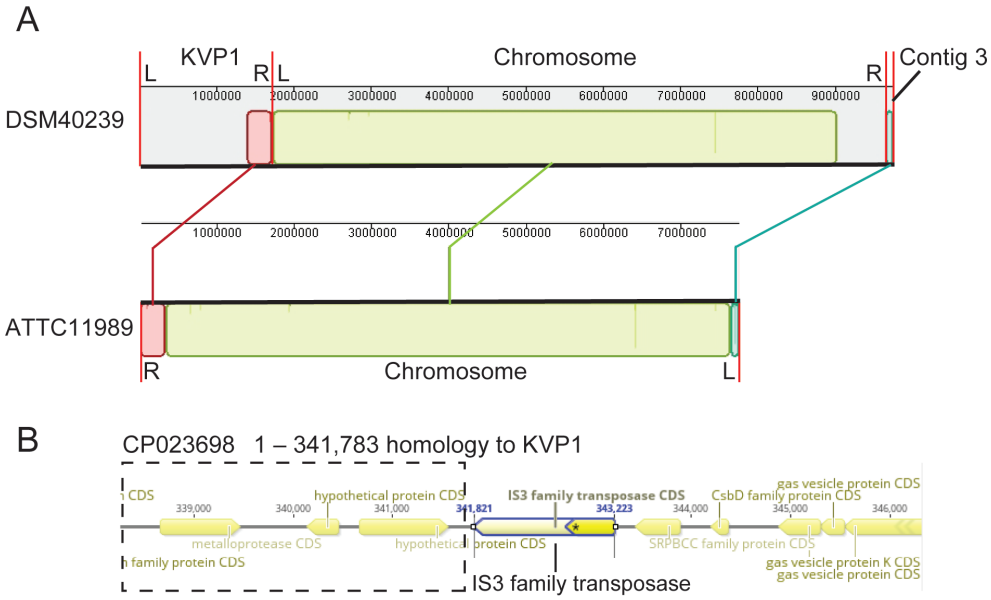


Supplementary Figure 3. Mean read coverage plots of Illumina sequences from isolates mapped to the *K. viridifaciens* DSM40239 reference genome

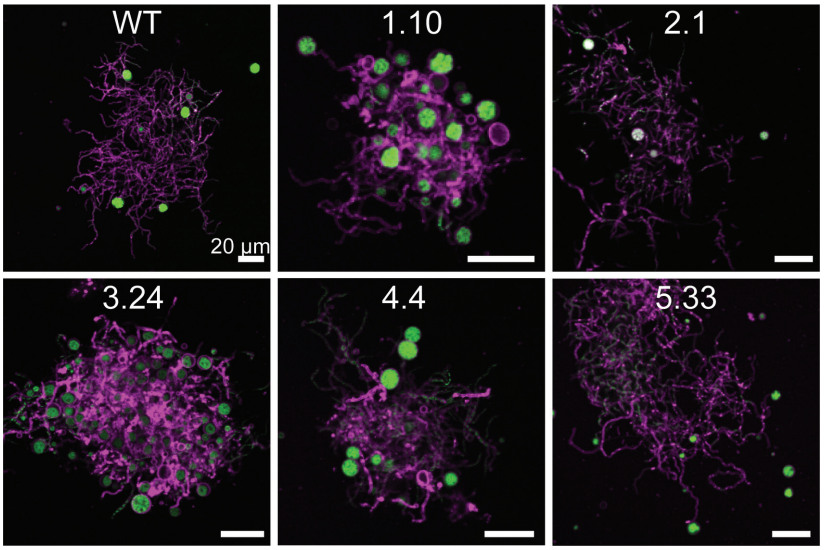
Mean coverage plots of Illumina sequencing reads from five isolates per cat. 1 **(A)** (wild-type-like), cat. 2 **(B)** (delayed sporulation), cat. 3 **(C)** (border-sporulation), cat. 4 **(D)** (heterogenous sporulation) and cat. 5 **(E)** (non-developed) mapped to the reference genome of *K. viridifaciens* DSM40239 contigs CP090840 (linear megaplasmid, KVP1), CP090841 (linear chromosome) and CP090842 (contig 3). The x-axis indicates the genomic position in Mbp and the y-axis indicates the mean read coverage. The lack of consistent read coverage along the genome is a strong indication that part of the sequence has been lost. Note that while isolates from cat. 1 have a wild-type-like coverage, for two isolates (2.1 and 2.37) the megaplasmid is (partially) lost. The majority of isolates from cat. 3, 4 and 5 have completely lost KVP1. Isolates from cat. 4 and 5 have deletions on one or both chromosomal arms. Orange dotted lines in **(E)** (CP090841) indicate the chromosomal location of *argG* (BOQ63_08205). See also Fig. 4 for wild-type coverage and Supplementary Fig. 4.



Supplementary Figure 4. Putative genomic amplifications contain TE-associated genes
(A – B) Mean coverage plots of Illumina sequencing reads from strain 5.34 **(A)** and 5.6 **(B)** (cat. 5) showing selected genomic regions of contig CP090841 and CP090840, respectively. The x-axis indicates the genomic position in bp and the y-axis indicates the mean read coverage. Note the nearly four-fold higher coverage for the region containing a gene encoding an IS5/ISS1182 family transposase (BOQ63_36085), maturase (BOQ63_36090) and IS5 family transposase (BOQ63_36095) in strain 5.34 as well as for the megaplasmid region in strain 5.6 containing genes encoding two site-specific integrases (BOQ63_00480 and BOQ63_00490). See also Fig. 5.



Supplementary Figure 5. Recombination between megaplasmid and chromosome in *K. viridifaciens*
(A) Whole-genome alignment of *K. viridifaciens* DSM40239 (contigs CP090840, CP090841 and CP090842) and *Streptomyces viridifaciens* ATTC11989 (contig CP023698). Coloured blocks indicate regions with homology. Note that the orientation of the ATTC11989 chromosome is reversed compared to DSM40239 as illustrated positioning of the coloured blocks below the black line and as indicated by the annotation of 'L' and 'R' relative to DSM40239. A 341 kbp region of the chromosome of ATTC11989 has homology to KVP1, suggesting a recombination event has occurred. **(B)** Chromosomal region of *S. viridifaciens* ATTC11989 indicating the presence of a gene encoding an IS3 family transposase (CP971_32970) flanking the genomic region (spanning nucleotide position 1 – 341,783 partially indicated by dotted box) that has homology to the megaplasmid KVP1.



Supplementary Figure 6. All isolates retain the ability to form S-cells
Growth of *K. viridifaciens* wild-type (WT) and strains from each isolate category (1 to 5) in LPB medium for 48 h (WT and cat. 1 to 4) or 240 h (cat. 5) 30 °C. Membranes were visualized with SynapseRed C2M (SynapseRed; red) and nucleic acids with SYTO 9 (green). Scale bar indicates 20 µm. One representative per isolate category is shown. All isolates formed S-cells upon growth under hyperosmotic stress conditions.

Supplementary Table 1. Effect of solid medium drying on S-cell reversion
Reversion of S-cells on LPMA medium for different drying times of the agar before sample application. Reversion is calculated as the percentage of S-cells that results in a colony forming unit (CFU). Concentration of S-cells ml^{-1} used to calculate the percentage of reversion was quantified based on haemocytometer counts. $n = 2$ biological replicates. See also Supplementary Fig. 1B.

Drying time (min)	CFU ml^{-1} Replica 1	CFU ml^{-1} Replica 2	Average CFU ml^{-1}	Reversion (%)
30	1.51×10^4	9.80×10^3	1.25×10^4	0.7
120	2.57×10^5	1.83×10^5	2.20×10^5	12.3
240	5.40×10^5	4.95×10^5	5.18×10^5	29.0
480	9.35×10^4	2.50×10^3	4.80×10^4	2.7
	Replica 1	**Replica 2**	**Average**	
S-cells ml^{-1}	1.83×10^6	1.75×10^6	1.79×10^6	

Supplementary Table 2. Phenotypes per isolate category

Phenotype is based on streaks of biomass on MYM medium, determined after 72 and 168 h growth. Strains were assigned to a category based on its phenotype after 72 and 168 h growth. Note that strains that were capable of sporulation, but did not have a phenotype belonging to category 1, 2 or 3 were combined to category 4. Each isolate identifier consists of its category number followed by its isolate number.

Isolate category	Phenotype after 72 h growth	Phenotype after 168 h growth	Isolates
1. Wild-type like	Confluent aerial hyphae formation along the biomass, mostly producing spores	Confluent sporulation	1.10 1.31 1.38 1.39 1.46
2. Delayed sporulation	Lack of visible aerial hyphae and/or sporulation OR Largely reduced aerial hyphae formation and sporulation compared to wild-type (e.g lighter spore lawn)	(near) confluent sporulation	2.1 2.12 2.32 2.37 2.48
3. Border sporulation	Large bald region(s) in center of biomass, with aerial hyphae formation along edges of biomass with/without spores	Large bald regions in sporulating biomass	3.24 3.25 3.43 3.44 3.50
4. Heterogenous	No aerial hyphae, or delayed like cat. 2 (4.11)	Heterogenous colouration of spore lawn with white specks, sometimes not confluent, some strains sporulate only occasionally	4.4 4.8 4.11 4.35 4.36
5. Non-developed	No visible aerial hyphae formation and sporulation	No visible aerial hyphae formation and sporulation OR Limited region with aerial hyphae that disappears upon restreaking	5.6 5.9 5.33 5.34 5.42

Supplementary Table 3. Overview of SNPs

SNPs found in all isolates 2.1, 2.12, 2.32, 2.48 and 3.25. Note that the SNPs in BOQ63_06755 and BOQ63_06225 could not be detected in mutant 2.1 as this region of the megaplasmid was lost. Nt. = nucleotide. CP090840 = megaplasmid KVP1, CP090841 = chromosome.

Contig	Locus_tag	CDS start	CDS end	Location SNP	Nt. change	Protein effect	Amino acid change	Product
CP090840	BOQ63_06755	1600667	1601458	1601017	G -> A	None	None	hypothetical protein
CP090840	BOQ63_06225	1489859	1496374	1494086	G -> A	A->T	Substitution	type I polyketide synthase
CP090840	BOQ63_04035	887674	888270	887718	C -> T	None	None	WHG domain-containing protein
CP090841	BOQ63_22455	3464669	3465952	3465495	C -> T	P->L	Substitution	adenylosuccinate synthase
CP090841	BOQ63_24775	3976035	3978449	3976237	T -> G	D->A	Substitution	bifunctional SulP family inorganic anion transporter /carbonic anhydrase
CP090841	BOQ63_24865	3996305	3996646	3996575	C -> T	R->W	Substitution	anti-sigma factor antagonist BldG
CP090841	BOQ63_13545	1371430	1373364	1372947	A -> T	S->T	Substitution	NAD-binding protein
CP090841	BOQ63_22150	3398051	3400327	3398596	G -> A	None	None	TOMM precursor leader peptide-binding protein

Supplementary Table 4. Overview of selected TE-associated genes

TE-associated genes located at genomic breakpoints or amplified regions in contig CP090840 and CP090841 of strains 5.34, 5.6, as well as protoplast revertant B3.1 [198], as annotated in Fig. 5 B–D and Supplementary Fig. 4. For the complete list of TE-associated genes annotated in *K. viridifaciens* DSM40239 see Electronic Supplementary Data, Table 1 and 2.

Strain	Contig	Location	Locus (+ gene)	Product	Min.	Max.
5.34	CP090840	Region A	BOQ63_02310	IS256 family transposase	538,367	539,662
		Region A	BOQ63_02315	IS630 family transposase	539,701	539,907
		Region A	BOQ63_03010	IS256 family transposase	670,880	672,175
		Region A	BOQ63_03015	integrase core domain-containing protein	672,247	672,589
		Region B	BOQ63_06540	IS5 family transposase	1,563,475	1,564,244
		Region B	BOQ63_06550	IS3 family transposase	1,564,717	1,565,607
		Region B	BOQ63_06555	transposase	1,565,604	1,565,894
		Region B	BOQ63_06565	transposase	1,566,526	1,566,960
		Region B	BOQ63_07300	IS3 family transposase	1,709,455	1,710,345
		Region B	BOQ63_07305	transposase	1,710,342	1,710,632
	CP090841	Left arm	BOQ63_10370	transposase	676,276	676,566
		Left arm	BOQ63_10375	IS3 family transposase	676,563	677,453
		Right arm	BOQ63_36085	IS5/IS1182 family transposase	6,402,423	6,402,539
		Right arm	BOQ63_36090	maturase	6,402,649	6,404,430
		Right arm	BOQ63_36095	IS5 family transposase	6,404,965	6,405,656
5.6	CP090840	Left cutoff	BOQ63_00480	site-specific integrase	99,472	101,490
		Left cutoff	BOQ63_00490	site-specific integrase	102,174	103,313
		Right cutoff	BOQ63_02810	transposase	635,350	636,825
		Right cutoff	BOQ63_02820 (*istB*)	IS21-like element helper ATPase IstB	637,750	638,475
		Right cutoff	BOQ63_02825 (*istA*)	IS21 family transposase	638,568	639,827
	CP090841	Right arm	BOQ63_40065 (*istA*)	IS21 family transposase	7,347,306	7,348,565
		Right arm	BOQ63_40070 (*istB*)	IS21-like element helper ATPase IstB	7,348,565	7,349,383
B3.1	CP090840	Left cutoff	BOQ63_06610	maturase	1,573,505	1,575,286
		Left cutoff	BOQ63_06615	IS5 family transposase	1,575,746	1,576,506
	CP090841	Right cutoff	BOQ63_36085	IS5/IS1182 family transposase	6,402,423	6,402,539
		Right cutoff	BOQ63_36090	maturase	6,402,649	6,404,430
		Right cutoff	BOQ63_36095	IS5 family transposase	6,404,965	6,405,656

Electronic Supplementary Data

Electronic Supplementary Tables 1 & 2 are available on the online repository OSF database and can be accessed via: https://doi.org/10.17605/OSF.IO/M6XUW

Electronic Supplementary Table 1.
List of transposable element-associated genes (TE-associated genes) annotated in *K. viridifaciens* DSM40239 contig CP090840 (megaplasmid KVP1).

Electronic Supplementary Table 2.
List of transposable element-associated genes (TE-associated genes) annotated in *K. viridifaciens* DSM40239 contig CP090841 (chromosome). Note that no TE genes were found on contig CP090842.