

# Whole-genome Sequencing of Two Streptomyces Strains Isolated From Sand Dunes of Sahara

Chahira Zerouki (✉ [chahira.zerouki@uef.fi](mailto:chahira.zerouki@uef.fi))

School of Forest Sciences, University of Eastern Finland, FI-80101 Joensuu, Finland

Farid Bensalah

Universite Oran 1 Ahmed Ben Bella Faculte des Sciences de la Nature et de la Vie

Suvi Kuittinen

University of Eastern Finland: Ita-Suomen yliopisto

Ari Pappinen

University of Eastern Finland: Ita-Suomen yliopisto

Ossi Turunen

University of Eastern Finland: Ita-Suomen yliopisto

---

## Research article

**Keywords:** Sahara, soil, Streptomyces, complete genome sequence, secondary metabolites, enzymes

**Posted Date:** October 26th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-96085/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

**Version of Record:** A version of this preprint was published at BMC Genomics on July 27th, 2021. See the published version at <https://doi.org/10.1186/s12864-021-07866-x>.

# Abstract

**Background.** Sahara is one of the largest deserts in the world. The hard conditions especially high temperature and aridity lead to unique adaptation of organisms that could be a potential source of new metabolites. In this respect, two Saharan soils from El Oued Souf and Beni Abbes in Algeria were collected. The bacterial isolates were selected by screening the antibacterial, antifungal, and enzymatic activities. The whole genomes of the two native Saharan strains were sequenced to study the desert *Streptomyces* microbiology and ecology from genomic perspective.

**Results.** The strains Babs14 (from Beni Abbes, Algeria) and Osf17 (from El Oued Souf, Algeria) were initially identified by 16S rRNA sequence as belonging to the *Streptomyces* genus. The whole genome sequencing of the two strains was performed by using Pacific Biosciences Sequel II technology (PacBio) and it showed that Babs14 and Osf17 have a linear chromosome of 8.00 Mb and 7.97 Mb, respectively. The number of identified proteins is 6910 in Babs14 and 6894 in Osf17. No plasmids were found in Babs14, whereas three plasmids were detected in Osf17 and no genome duplications were detected in the two strains. Although the strains have different phenotypes and are from different regions, they are slightly different from each other in protein sequence level but, differ more from their earlier published relatives. 29 biosynthetic gene clusters (BGCs) coding for secondary metabolites were predicted in each strain indicating antifungal, antibiotic, and other activities of Babs14 and Osf17. Moreover, many other BGCs were related to biosynthesis of factors needed in response to environmental stress in temperature, UV light and osmolarity.

## Conclusion.

The genome sequencing of Saharan *Streptomyces* strains revealed factors that are related to their adaptation to an extreme environment and stress conditions. The genome information gives tools to study the ecological adaptation in desert environment and to explore the bioactive compounds of these microorganisms. The two whole genome sequences are among the first ones for the *Streptomyces* genus of Algerian Sahara. The present research was undertaken as a first step to explore more profoundly the desert microbiome.

## 1. Background

The actinomycetes are filamentous bacteria that inhabit various ecological niches including soil and water. The actinomycetes are economically and biotechnologically important and produce a wide range of bioactive metabolites such as antibiotics, antiparasitic agents, antitumor agents, immunosuppressive agents, vitamins, nutritional materials, pesticides, herbicides, and different enzymes [1, 2, 3, 4].

The search for new substances of microbial origin has been largely based on their isolation and depends on the number and diversity of the isolated and screened strains [5]. One of the strategies of obtaining isolates with desired properties is the search for cultivable bacteria from extreme biomes which include deserts. Sahara is characterized by extreme conditions, including high aridity and temperature varying between about -5 and +45°C. Typically, there are high UV radiation and desiccation conditions. Moreover, the presence of inorganic oxidants, but also, the starvation, and the physical instability caused by strong winds make the environment very challenging for the microorganisms [6]. Many organisms of extreme environments may represent new taxa and could provide a valuable resource for finding new bioactive compounds and enzymes for the use in commercial applications [7,8,9]. New actinobacteria have been discovered from microbiologically poorly explored areas. Moreover, the previous investigations have demonstrated the abundance and diversity of actinomycetes in the Sahara soils [10], but, also in arid Atacama Desert soils in Chile [10,11]. The isolation and studies of actinobacteria from Algerian Sahara confirmed the possibility of discovering novel species. In 2013, Meklat *et al*, published three novel *Actinopolyspora* species [12,13,14], and two novels halophilic actinobacteria were published by Saker *et al* in 2014, Meklet *et al*, in 2015, and a novel *Actinomadura* species by Lahoum *et al*, in 2016. All these strains were isolated from Algerian Sahara soil [15, 16, 17]. Although plant-free desert ecosystems represent lowered variation capacity [18], the Saharan soils under hard climate conditions still represents an ecosystem with significant biodiversity [19]. The present

study describes the isolation and characterization of Actinomycetes strains from South-West Algeria (Beni Abbes) and South-East Algeria (El Oued Souf), and their screening for antimicrobial, antifungal, and enzymatic activities. In addition, whole genome sequencing and bioinformatic tools were used to analyze the protein sequences of these two strains.

## 2. Results

### 2.1. Strain isolation and selection

A total of twelve isolates were identified showing filamentous features, which may indicate *Actinobacteria*. Two strains from different samples were the subject of our study, Babs14 from Beni Abbes and Osf17 from El Oued Souf (Fig. 1). Babs14 formed a raised, dry, and beige-cream colony with irregular and white margins and irregular surface, and Osf17 formed a dry, white-gray, convex, and circular colony with a brown pigment diffusion into the surrounding medium. Both strains showed clear hyphae structure and their morphology and hypha represent those of typical actinomycetes. The isolates are gram-positive, and catalase and phosphate solubilization positive. The optimum growth temperature for Babs14 and Osf17 is ca. 30°C. The phylogenetic analysis of strains Babs14 and Osf17 based on 16S rRNA sequences was conducted by MEGA 7.0 with the related species to show their phylogenetic relationships (Fig. 2). Babs14 and Osf17 were clearly clustered as belonging to the *Streptomyces* genus. The general features of Babs14 and Osf17 are listed in the Table 1.

### 2.2. Screening for potential activities

The Babs14 and Osf17 showed inhibition against all used microorganisms. It was possible to quantify the regions of growth and measure the clear (inhibited) areas, which are indicators of antimicrobial activity and are proportional to the concentration and diffusion of the antimicrobial compounds. The antimicrobial activities are expressed as the inhibition zone diameter (mm). Both strains showed a clear inhibition activity against *A. niger* ATCC 6275, *E. coli* ATCC 25922 and *S. aureus* ATCC 43300. However, Osf17 demonstrated a stronger inhibition with an inhibition diameter higher than Babs14. The enzymatic screening showed a wide enzymatic activity pattern in Babs14 and Osf17. The strains were able to degrade casein, lipid, starch, pectin, cellulose and birchwood xylan. The only difference between the two strains was the absence of detectable proteolytic activity in Osf17 (Table 2).

### 2.3. General Genome Features of Babs14 and Osf17

The complete genomes of strains Babs14 and Osf17 were sequenced by PacBio platforms and the genome completeness was checked by Busco [20]. The genome was 92.6% complete with 92.6% of the genome representing single copy genes and 0% was duplicated in Babs14. The genome of Osf17 was 94.6% complete and 0% of the genome was duplicated. The assembly of Babs14 by Flye 2.4.1 was generated from 2185080676 total bases with read depth of 266 over the genome, and the read length N50/N90 was 11022/6592. For the strain Osf17 the assembly was generated from 2269979737 total bases with read depth of 275 over the genome and the read length N50/N90 was 11555/6692. The total genome size of Babs14 is 7999102 bp in two contigs of 114649 bp and 7884453 bp with a GC content of 72,6%. The strain Osf17 has a genome size of 7967258 bp with two contigs of 7883883 pb and 83375 pb with a GC content of 72.6 %.

### 2.4. Genome annotation

The genome annotation of Babs14 with Prokka version 1.11 identified 6910 protein-coding genes (CDSs), and a single tmRNA, 87 tRNA and 18 rRNA genes, and two CRISPR regions, whereas no plasmids were found. The strain Osf17 has 6894 protein-coding genes, and a single tmRNA, 86 tRNA and 18 rRNA genes, and two CRISPRs and three native plasmids were found (Table 3). The BlastN run of the complete genome of the strains Babs14 and Osf17 revealed that the closest strains are *Streptomyces* sp. SGAir0924 (gb|CP027297.1) and *Streptomyces* sp. SS52 (gb|CP039123.1), respectively, with 99% similarity. The average nucleotide identity (ANI) searches were performed using [Kostas lab ANI calculator \(http://enve-omics.ce.gatech.edu/ani/index\)](http://enve-omics.ce.gatech.edu/ani/index) [21] and the results were confirmed with JSpeciesWS server (<http://jspecies.ribohost.com/jspeciesws/>) [22] against the reference genome sequences, which resulted in an ANI match

over 95% cutoff indicating that the strain Babs14 belongs to the same species as *Streptomyces* sp. SGAir0924 (CP027297.1) [23] that was isolated from outdoor air in Singapore. Correspondingly, Osf17 belongs to the same species as *Streptomyces* sp. SS52. (CP039123), an endophytic strain isolated in Vietnam [24].

The genomic map comparison (Fig. 3) using CGViewer webserver [25] showed that the differences between Babs14 and the three other strains are mostly evenly distributed and there exists only few areas with bigger differences. SEED Viewer maps show clearly that Babs14 and Osf17 are much more similar with each other than with *Streptomyces* sp. SGAir 0924 (CP027297.1) and *Streptomyces* sp. SS52 strains (CP039123) (Additional File 1, Figs. S1-S6). The RAST map shows that the identity between Babs14 and Osf17 is all over the genome close to 100%, but the RAST maps also show that there are below 80% identity regions between Sahara strains and their closest relatives (Figs. S1-S3). This indicates that the two Saharan strains may have ecological reasons for their genetic similarity despite of long distance between their isolation site and phenotypic differences. The dot plot comparison between Babs14 and Osf17 showed that the genome sequences have been divided in a different way at the genome ends between the contigs in the genome cloning (Fig. S4). The comparison of Babs14 and Osf17 genomes to their closest relatives also shows that the Babs14 and Osf17 genome sequences are fully or almost fully complete (Fig. S5 and Fig. S6).

### 2.5. Protein comparison

OrthoVenn2 (<http://www.bioinfogenome.net/OrthoVenn/>) [26] was used to compare between the proteins of Babs14 and Osf17 and the proteins of their closest relatives using the files generated by Prokka analysis (Fig. 4). 6807 and 6802 orthologous protein clusters were found in Babs14 and Osf17, respectively. Altogether 293 protein clusters (284 common in Babs14 and Osf17, and 4 or 5 that are specific to each strain in Venn diagram) were unique to the strains Babs14 and Osf17. In addition, 39 singleton proteins were found in Babs14 and 48 singleton proteins in Osf17 (Table 4). The 284 proteins shared by Babs14 and Osf17 but not found in the other strains are responsible of different biological functions as DNA modification, RNA metabolic process, response to stimulus, carbohydrate metabolic process and many other functions (Additional file 2, Table S1). The OrthoVenn2 analysis indicated that although most of the proteins are very similar, there are clear genetic differences between all the studied strains. Cluster 6455 was unique to Babs14 and Osf17 and is identified as heat response protein (GO:0009408). In Gene Ontology (GO) Term, the heat response protein is assigned to any changes of the activity of a cell or an organism including enzyme production, secretion, gene expression etc. caused by a heat stimulus [27]. This property is apparently important in the desert environment.

Both strains, Babs14 and Osf17 showed a biological ability for the degradation of different substrates such as starch, gelatin, xylan, cellulose and pectin. However, the strain Osf17 showed negative protease (caseinase) activity during the screening. A genomic investigation was done to understand the missing activity. The comparison of the protein sequences showed that both strains have 23 proteases in common, 25 proteases were found in Babs14 and 24 proteases in Osf17, and the carboxy-terminal processing protease CtpA precursor was missing in Osf17. Moreover, three modulators of FtsH and protease HflK were found in Babs14, and only two sequences of the modulators FtsH, and protease HflK were found in Osf17. Furthermore, one protease HtpX was unique to Osf17. The differences between the proteases and the missing sequences but possibly also differing regulation could be related to the absence of the detected ability of Osf17 to degrade casein. It is also possible that the testing time or conditions were not proper for inducing the protease enzyme activity in Osf17.

The enzymatic activities were searched at the sequence level using the Prokka and CAZy annotations. Various xylanases, cellulases, glucanases, pectin esterases, lipases and amylases were found in Babs14 and Osf17. Concerning xylanases, the CAZy annotation assigned two proteins to the glycoside hydrolase family GH10 xylanases in both strains and only one glycoside hydrolase family GH11 xylanase was found in Babs14 and Osf17. One of the GH10 xylanases is exoxylanase and one endoxylanase according to PROKKA, and both are 100% identical between the two strains. The exoxylanase and endoxylanase are 46% identical. GH11 xylanases are 100% identical to each other. Already this short examination shows that at the protein level the enzymes between the two species tend to be identical.

## 2.6. COG annotation

A total of 1865 and 1852 genes were assigned to the COG databases for Babs14 and Osf17, respectively. Numbers of genes annotated by COG were similar in the two strains (Fig. 5; see also RAST result in Fig. S7), the genes encoding transcription accounted for the largest proportion 14,20% and 14,13% for Babs14 and Osf17, respectively. The genes encoding carbohydrate transport accounted for 10,01% for Babs14 and 10,04% for Osf17. The amino acid transport and metabolism accounted in Babs14 for 9,63% and 9,68% in Osf17 (Fig. 5). The RAST annotation showed many subcategory distributions and most features confirmed the COG analysis. The largest features were attributed to the amino acids and derivatives, followed by carbohydrate metabolism proteins, protein metabolism and cofactors, vitamins, prosthetic groups, and pigments for Babs14 and Osf17 (Fig.S7). These annotations indicated the ability of these two strains to use the carbohydrates and the amino acids and protein resources available in their living environment.

## 2.7. Island pathways and phages

The genomic islands (GEIs) could be related to a variety of functions and have a big role in the genome plasticity, evolution, and the biological adaptability. A part of the horizontal gene transfer is facilitated by genomic islands. The GEIs identified are also implicated in the antibiotic resistance, pathogenicity, symbiosis, xenobiotic degradation, and primary and secondary metabolism [28, 29]. In the strain Babs14, 12 GEIs were found with an average length of 17051 bp (Table S2). Osf17 contains 13 GEIs and the average length is up to 17741 bp (Table S3). Some CDS sequences were predicted to be in the GEIs of Babs14 and Osf17, although many CDSs were identified as hypothetical proteins. The majority of defined CDSs in GEIs of Babs14 and Osf17 are related to DNA transcription or response to regulators, and enzymes. In the GEIs of Babs14 and Osf17, two transposases were predicted; the transposases are associated to the genome plasticity and its adaptation to environment enabling the strain to horizontal transfer of genes [30].

Phages were investigated using PHASTER, and one incomplete phage region with a score of 40% was assigned to the *Rhodococcus* REQ1 phage (NC\_016655) in the second contig of Babs14. In the first contig of Osf17, three incomplete phage regions with a score of 30% were assigned to the *Bacillus* G phage (NC\_023719), *Gordonia nymphadora* phage (NC\_031061) and *Erwinia* phage vB\_EamM EarlPhillipIV

(NC\_031007) (Table S4).

## 2.8. Genetic basis for secondary metabolites

The strains Babs14 and Osf17 were isolated from the sand dunes of Sahara. Their activities against pathogenic microorganisms indicated the existence of antimicrobial and antifungal gene clusters. AntiSMASH predicted the presence of different types of secondary metabolite biosynthesis gene clusters (BGCs). The potential gene clusters encoding secondary metabolites of Babs14 and Osf17 are listed in the Table S5 and Table S6 (Additional file 2), respectively. 29 biosynthetic gene clusters were predicted in both strains and they are quite similar, among which 26 clusters showed similarity to gene clusters with a known function. Eight gene clusters (clusters 1, 7, 13, 17, 20, 24, 25 and 27 in both species) showed 100 % similarity to the reported clusters (Table 5).

The cluster 1 presented 100% similarity for the biosynthesis of **isorenieratene**, which is a carotenoid light harvesting pigment and is biosynthesized by brown-colored strains [31,32]. If this compound influences the brown pigment of Osf17, then in Babs14, the expression level could be different (see the colors in Fig. 1). The cluster 2 shows 95% similarity to the known cluster responsible for biosynthesis of candidicin, which is an aromatic heptane that shows antifungal activity [33]. The cluster 7 is responsible for biosynthesis of ectoine, which is a microbial protectant against osmotic stress and serves as a versatile nutrient [34]. The cluster 13 is responsible for the biosynthesis of albaflavenone, a novel antibiotic sesquiterpene [35].

The genus *Streptomyces* belongs to odor producing *Actinomycetes* [36]. The cluster 17 was assigned to the biosynthesis of geosmin, which is earthy-muddy-smelling metabolite. The cluster 20 is predicted to biosynthesize hopene, which is the precursor of a many hopanoids that play an important role in stabilizing the membranes [37]. The cluster 24 was assigned to the biosynthesis of **coelichelin**, which is a peptide siderophore, iron-binding compound [38]. The cluster 25 for biosynthesizes coelibactin, which is a putative non-ribosomally synthesized peptide with a zincophore activity implicated in antibiotic regulation [39]. The cluster 27 was predicted as SapB biosynthetic gene cluster which is a small, secreted morphogenetic peptide involved in the growth in complex medium and is considered as an antibiotic-like molecule [40]. The cluster 28 for biosynthesizes streptothricin and has in Babs14 and Osf17 91% and 87 % similarities, respectively, with the known gene clusters. Streptothricin is a broad-spectrum antibiotic, which has showed a strong antimicrobial activity against many species of fungi and bacteria [41].

Furthermore, the similarities of cluster 8, cluster 9, cluster 11, cluster 14, cluster 19 and cluster 29 in Babs14 and Osf17 with the known gene clusters were more than 60%, which might indicate that the ability for the biosynthesis of melanin, deferrrioxamin B / deferrrioxamin E, catenulipeptin, spore pigment, **CDA1b / CDA2a / CDA2b / CDA3a / CDA3b / CDA4a / CDA4b** (calcium dependent antibiotic) and fluostatins M-Q exists, if the missing biosynthesis proteins exist but were not detected.

### 2.9. Adaptation to environment and stress responses

In extreme and rapidly changing conditions, the microorganisms must sense the changes and respond with appropriate mechanisms. Stress responses were investigated in Babs14 and Osf17 using SEED viewer version 2.0 [42]. Twenty-eight stress response entries were found for both strains including osmotic, oxidative, and periplasmic stress responses, and detoxification and dimethylarginine metabolism (Additional file 2, Table S7). The extracytoplasmic function (ECF) factors sigma B ( $\sigma^B$ ) and sigma E ( $\sigma^E$ ) were found in Babs14 and Osf17. These factors are involved in gene regulation and expression in response to various extracellular changes. Sigma B plays a role in the resistance to a variety of stresses such as high and low pH, heat, high osmolarity, high ethanol concentrations and oxidizing agents [43,44]. Sigma E is involved in the control of the periplasmic heat shock regulon, which is activated by unfolded proteins in the periplasm [45].

The genome annotation confirmed the presence of heat shock proteins (Hsp) in each Sahara strain including: Hsp 15, Hsp18, heat-inducible transcription repressor HrcA and putative heat shock protein HspR. Previous studies on *Streptomyces albus* showed that HSP18 is involved in the thermotolerance at extreme temperatures [46]. Chaperone proteins ClpB, DnaK, DnaJ and HtpG were found in Babs4 and Osf17, which are a part of a stress-induced multi-chaperone system involved in the recovery of cells from heat-induced damages [47].

**During the wintertime, the temperature in the Algerian Sahara can drop to below freezing. In this sense, cold shock responses were investigated.** Cold-shock proteins (Csp) are induced upon temperature downshift and are involved in the adaptation of cells to cold. Cold shock protein ScoF, cold shock-like protein 7.0 and CspA, which is a major shock protein, were found in Babs14 and Osf17. Recent studies have shown that Csps might also have a wider role in stress tolerance of bacteria [48]

Programmed cell death and toxin-antitoxin systems Phd-Doc and YdcE-YdcD were found in Babs14 and Osf17 and many studies have suggested that they can activate the programmed cell death to survive in different environmental stresses caused by nutrient deprivation and antibiotics. Toxin-antitoxin systems have also been shown to be responsible for the stress management, bacterial persistence, plasmid maintenance, and biofilm formation [49].

## 3. Discussion

*Streptomyces* strains are of a great commercial importance due to their biologically active compounds. Previous reports demonstrated that actinobacteria residing in desert have the capability to survive under extreme conditions and possess

many gene clusters to produce bioactive compounds [50]. Connecting the natural abilities of production of active metabolites to the genes that encode them is enabling the utilization of the genome sequence data in the discovery of new molecules. The antagonistic potentials of actinomycetes isolated from Algerian desert against plant pathogenic fungi have been reported [51,52]. The previous investigation showed the potential of these microbes to produce antifungal compounds but also many hydrolytic enzymes [51,52]. To investigate further the potential bioactivities in Saharan microorganisms, the strains Babs14 and Osf17 were isolated from the sand dunes in *Beni Abbes* and El Oued Souf, which are arid areas in the Eastern great erg and the Western great erg in Algeria, respectively. The strains were selected for their activity profiles including the antibacterial, antifungal, and enzymatic activities. The molecular identification allowed us to classify the strains Babs14 and Osf17 as *Streptomyces sp.* with 99 % similarity to *Streptomyces sp.* SGAir 0924 and *Streptomyces sp.* SS52, respectively.

Babs14 and Osf17 were found to genetically resemble each other in enzymes, BGCs and stress responses. However, the two Saharan strains differ from each other in few aspects as plasmids, genomic islands, phages, and in some proteins. Based on OrthoVenn2 analysis, 284 proteins were shared by the Sahara strains that are missing in their closest relatives. These proteins were assigned to different metabolic and cellular functions, in which some proteins could be related to the adaptation of these bacteria to the available resources in their biotope. The high genomic similarity of these phenotypically differing strains supports the earlier finding that plant-free deserts show low levels of phylogenetic and taxonomic diversity and low diversity in protein-coding gene categories [18]. The high similarity might reflect their ecological adaptation.

Sand dune habitats are changing by many factors e.g. dune encroachment, sand accumulation, and wind erosion [53]. Previous studies of dust samples collected from Saharan region in Cap Blanc showed that dusts contain organic compounds like charcoal-like fragments of burnt vegetation, leaf wax-derived lipids adsorbed on clays, pollen grains and amorphous material [54]. The existence of a wide range of hydrolytic enzyme activities and carbohydrate-active enzymes indicates that these strains survive well in the sand areas containing plant debris, e.g. distributed by wind. Dust and sandstorms also play a role in spreading Saharan microbes and spores to various places [55]. The whole genome sequencing of Saharan microbes gives possibilities to identify the spreading microbes and thus create a more complete picture about the influence of Saharan microbes in other geographic areas.

Genome mining and metabolite analysis indicates that the isolated strains have a great potential in secondary metabolites production. Many genes involved in antibiotic biosynthesis showed a high similarity with the known genes and suggest that Babs14 and Osf17 and similar strains could be sources for commercially applicable secondary metabolites. It has been reported that the *Streptomyces* genus is abundant in Atacama Desert and 50% of these strains produce antibacterial and other compounds such as carotenoids with antibacterial activity, even against Gram-negative bacteria [55]. AntiSMASH has been used widely to identify biosynthesis gene clusters e.g. in *Streptomyces* [57,58,59,60,61]. A genome-wide study showed that *Streptomyces* genomes possess 25–70 known BGCs, but the genomes could contain more biosynthetic gene clusters coding for secondary metabolites [62]. It is also possible that antiSMASH did not identify all potential BGCs. Novel metabolites might be identified by further genome mining with the support of mass spectrometry analysis.

Microorganisms from arid soil are exposed to periodic nutrient starvation and various abiotic and biotic stresses. These conditions are unfavorable for the bacterial growth and to survive, bacteria must respond. A strategy of the desert microbiota is to increase the abundance of genes involved in osmoregulation and dormancy, which contribute to their survival in hostile environment [55,63]. The stress responses investigated in this study showed the presence of BGCs in Babs14 and Osf17 that could be involved in the adaptation to stresses caused by Saharan environment, e.g. pigments, ectoine, carotenoid, and also, factors responding to heat, cold, and oxidative and osmotic stresses. The spore coats contain pigments that absorb UV radiation and play a significant role in the resistance to UV-A and UV-B radiation [64,65].

The low optimal growth temperature (~30°C) compared to highest daytime temperatures (~45°C) probably reflects the living of these microbes under the soil surface. Therefore, the sand is probably partly protecting the microbes isolated from the layers below the surface.

## 4. Conclusion

Studying the two Saharan strains revealed reasons for adaptation to the desert environments. The capacity for sporulation and competitive production of enzymes and secondary metabolites synthesis, including the stress response mechanisms make them capable for surviving in the extreme ecological niches [18]. Sahara is one of the largest deserts and it spreads upon several countries of Africa, but it has one of the smallest registration rates of biodiversity in biological databases [66]. Actinobacteria seem to dominate the microbial desert soil communities [67]. There are not many earlier reported whole genome sequences from the Saharan microbes [7]. Our study opens ways to investigate more profoundly the microbial diversity, adaptation, and genomic evolution in the desert conditions of Sahara. The genomic data comparison and analysis could increase our understanding of the adaptation capability [68] but also, pathways and enzymes of the *Streptomyces* living in extreme environments that could have a biotechnical and pharmaceutical significance.

## 5. Methods

### 5.1. Sampling and microbial isolation

Non-rhizosphere soil samples from 5–20 cm of depth, were collected from Beni Abbas and El Oued Souf. The samples were packed in sterile containers and stored at 4°C [51]. The samples were air-dried and heated aseptically for 1 hour at 45°C in a hot air oven, then cooled to room temperature to remove the undesired Gram-negative bacteria. Selective media were used to promote *Actinomycetes* growth [69, 70].

Dilution techniques and different media such as yeast extract–malt extract agar medium (ISP2) and starch-casein agar (SCA) were used for the isolation. The plates were incubated at 28 °C to 40°C for 14 days at pH 7. Colonies were picked out and streaked until purity. All the isolates were examined by light microscopy to detect the actinomycetes and stored at 4°C for further studies [51, 69, 70].

### 5.2. Morphological and cultural characterization of the isolates

Cultural characteristics of the isolates were observed on yeast extract malt extract agar (ISP2), oatmeal agar (ISP3) and inorganic salt starch agar (ISP4). The plates were incubated at 30°C for 7–21 days. The sporulation was observed by light microscopy [51,71]. The Gram stain and phosphate solubilization tests [72] were performed as described. The catalase activity was determined by adding 3% H<sub>2</sub>O<sub>2</sub>.

### 5.3. Screening of the antimicrobial and antifungal activities

The antimicrobial activity was tested *in vitro* against *Escherichia coli* ATCC 25922 and *Staphylococcus aureus* ATCC 43300 and antifungal activity against *Aspergillus niger* ATCC 6275. The antibacterial tests were done by the agar antimicrobial spot assay. The pure isolates were spot inoculated (6-8mm) on ISP2 agar and incubated for 5 days at 30°C. A second layer was made by pouring on the plate 10 ml of semi-agar (0.7%) inoculated with 1.5 10<sup>8</sup> CFU ml<sup>-1</sup> (counted at 600 nm value 0,5 according to McFarland standard) of the pathogenic microorganisms. The plates were incubated at 37°C for 24 h. A clear zone around the spot is considered as positive [73,74]. Cross streak method was used for the screening of antifungal activity. The isolates were streaked by a central streak on the agar plates and incubated for five days at 30°C. The pathogenic fungus was streaked perpendicular to the first streak and incubated for 48–72 h. The antibacterial and antifungal interactions were analyzed by observing the inhibition zone size expressed in millimeters [73].

### 5.4. Enzymatic Screening

Enzymatic activities, including amylolytic, proteolytic (caseinase), lipolytic, pectinolytic, cellulolytic and xylolytic activities, were screened using zone clearance assays. Gelatin hydrolysis was screened using a nutrient gelatin medium; the tubes were inoculated and incubated at 30°C up to 7 days. After incubation, the tubes were kept on ice for 15 to 30 minutes.

Hydrolyzed gelatin results in the liquefaction of the medium even exposed to cold temperature [75]. The enzymatic potential screening was done by streaking the strains in spots on agar media containing (g/L): 0.05 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 0.05 g NaCl, 0.01 g CaCl<sub>2</sub>, 0.2 g yeast extract, 0.5 g peptone, 2 g agar [76]. 1% (w/v) amount of the different substrates was incorporated as a principal carbon source. The plates were incubated 3-7 days at 30°C, and then amylase and pectinase plates were flooded with an iodine solution [77,78]. Xylanase and cellulase plates were stained by using Congo red solution and washed after 10 min with 1 M NaCl [79,80]. A clear zone around the colony indicates the presence of the screened activity. The production of lipolytic enzyme was detected using Sorbitan monolaurate (Tween 20). The indication of lipolytic enzymes in the colony was a visible precipitate due to the formation of the crystals of calcium salt of lauric acid liberated by the enzyme [81]. The proteolytic activity was tested qualitatively on casein agar using skim milk; the ability of the bacterial strains to use protein (casein) is shown through clear zones surrounding the colonies [82,83].

#### *5.5. PCR amplification and phylogenetic analysis / Molecular identification*

The bacterial agar plates were shipped to MacroGen Europe. Polymerase chain reaction (PCR) was performed by MacroGen Europe to amplify and sequence the 16S rRNA sequences using primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3'). The sequences were aligned using the homology search by comparing the sequence with those present in the public database (NCBI) using the standard Basic Local Alignment Search Tool (BLAST). Phylogenetic analysis was conducted using the Molecular Evolutionary Genetics Analysis (MEGA) software version 7 [84] and multiple alignments of data were performed by Clustal W [85]. The phylogenetic tree was reconstructed with the neighbor-joining algorithm. Topology of the resultant tree was evaluated by bootstrap analysis of the neighbor-joining dataset, based on 1000 replications [86].

#### *5.6. Whole genome sequencing*

The microbial strains were cultivated in Yeast-Malt-Glucose (YMG) broth for 48 h at 30°C to produce cell mass. The genomic DNA of the strains Babs14 and Osf17 was initially extracted using Invitrogen PureLink® Genomic DNA kit. The DNA quantity and quality were tested by the NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific). The DNA was purified further by using the Quick-DNA Miniprep Plus kit by Zymo research at the sequencing center of the University of Oregon Genomics & Cell Characterization Core Facility (GC3F) and the whole genome sequencing was done there by Pacific Biosciences Sequel II technology (PacBio). The DNA was made into SMRTbell libraries using the Express Template Prep Kit 2.0 from PacBio according to the manufacturer's protocol. Samples were pooled into a single multiplexed library and size selected using Sage Sciences' BluePippin using the 0.75% DF Marker S1 High-Pass 6 kb-10 kb v3 run protocol and S1 marker. A size selection cutoff of 8000 (BPstart value) was used. The size-selected SMRTbell library was annealed and bound according to the SMRT Link Set Up and sequenced on a Sequel II. Raw PacBio reads were converted to fasta format with Samtools Fasta (<http://www.htslib.org/doc/samtools.html>) and then assembled with Flye 2.6 (<https://github.com/fenderglass/Flye>) with parameters-plasmids-iterations 2-asm-coverage 120.

#### *5.8 Genome annotation and analysis*

The assembled bacterial genomes were annotated with Prokka 1.11 (<https://github.com/tseemann/prokka>) [87] and RAST Rapid Annotation using Subsystem Technology (<https://rast.nmpdr.org/>) [88]. Ribosome RNA (rRNA) genes and transfer RNA (tRNA) genes were predicted using tRNAscan-SE galaxy version 0.4 [89]. The genomic islands (GEIs) were predicted by IslandViewer 4 (<http://www.pathogenomics.sfu.ca/islandviewer/upload/>), which is an integrated interface for computational identification and visualization of genomic islands [90]. PHASTER (<https://phaster.ca/>) was used for the prophage prediction [91] and the CRISPRCasFinder 1.1.2 (<https://crisprcas.i2bc.paris-saclay.fr/>) for the CRISPR identification. The gene functions were analyzed by BLASTP using Cluster of Orthologous Groups (COG) of proteins on WebMGA server [92], Kyoto Encyclopedia of Genes and Genomes (KEGG) [93], and CAZy enzymes using dbCAN2 (<http://bcb.unl.edu/dbCAN2/index.php>) [94]. The potential secondary metabolite biosynthetic gene clusters were

investigated using antiSMASH v.4.1.0 [95]. KEGG and COG annotations were done for proteins predicted with Prokka. dbCAN2 and antiSMASH were run with the genomic DNA sequence.

## 6. List Of Abbreviations

ANI: average nucleotide identity

BGCs: biosynthetic gene clusters

COG: Cluster of Orthologous Groups

Csp: Cold shock proteins

ECF: the extracytoplasmic function

GEIs: the genomic islands

Hsp: heat shock protein

ORF: open reading frame

PCD: Programmed Cell Death

TAs: toxin-antitoxin systems

## Declarations

### *Ethics approval and consent to participate*

Not applicable

### *Consent for publication*

Not applicable

### *Availability of data and materials*

All data generated or analyzed during this study are included in this published article [and its supplementary information files]. The genomic sequences described here, have been submitted to NCBI GenBank under BioProject ID PRJNA665615, and BioSample accessions SAMN16261965 and SAMN16261966.

### *Competing interests*

The authors declare that they have no competing interests.

### *Funding information*

This project was supported by the Algerian Ministry of Higher Education and Research by a grant for visit of CZ to the University of Eastern Finland. The University of Eastern Finland provided financial support for CZ to do this research at UEF, including design of the study and collection, analysis, interpretation of data, writing the manuscript, and for the laboratory costs of the whole genome sequencing.

### *Authors' contributions*

CZ did the experiments and most sequence analyses, OT supervised the genome research, FB supervised the strain isolation and characterization, CZ and OT interpreted the genome analyses and wrote the manuscript. SK and AP contributed to the experimental work. All authors read and approved the final manuscript.

### ***Acknowledgements***

The authors thank Dr S. Larouci, Dr A. Guemouche and Dr K. Bouacem for their valuable advices and Maini Mononen and Risto Ikonen for the technical assistance.

### ***Sample collection statement***

The collection of soil samples was done by CZ as a member of the Laboratory of Microbial Genetics, Oran 1 University, accredited by the Algerian Ministry of High Education and Research to do research with environmental samples. No separate permission was needed for the sample collection from Saharan dunes.

## **References**

- [1] Bérdy J. Thoughts and facts about antibiotics: Where we are now and where we are heading. *J Antibiot.* 2012; doi : [10.1038/ja.2012.27](https://doi.org/10.1038/ja.2012.27).
- [2] Bredholt H, Fjærvik E, Johnsen G, Zotchev SB. Actinomycetes from sediments in the Trondheim fjord, Norway: Diversity and biological activity. *Mar Drugs.* 2008; doi: [10.3390/md6010012](https://doi.org/10.3390/md6010012).
- [3] Abdelmohsen UR, Grkovic T, Balasubramanian S, Kamel MS, Quinn RJ, Hentschel U. Elicitation of secondary metabolism in actinomycetes. *Biotech Adv.* 2015; doi: [10.1016/j.biotechadv.2015.06.003](https://doi.org/10.1016/j.biotechadv.2015.06.003).
- [4] Nouioui I, Carro L, García-López M, Meier-Kolthoff JP, Woyke T, Kyrpides NC, et al. Genome-based taxonomic classification of the phylum actinobacteria. *Front Microbiol.* 2018; doi: [0.3389/fmicb.2018.02007](https://doi.org/0.3389/fmicb.2018.02007).
- [5] Mazza P, Monciardini P, Cavaletti L, Sosio M, Donadio S. Diversity of *Actinoplanes* and related genera isolated from an Italian soil. *Microb Ecol.* 2003; doi [10.1007/s00248-002-2038-4](https://doi.org/10.1007/s00248-002-2038-4).
- [6] Makhalanyane TP, Valverde A, Gunnigle E, Frossard A, Ramond JB, Cowan DA. Microbial ecology of hot desert edaphic systems. *FEMS Microb Rev.* 2015; doi [10.1093/femsre/fuu011](https://doi.org/10.1093/femsre/fuu011).
- [7] Sayed AM, Hassan MHA, Alhadrami HA, Hassan HM, Goodfellow M, Rateb ME. Extreme environments: microbiology leading to specialized metabolites. *J Appl Microb.* 2020; doi [10.1111/jam.14386](https://doi.org/10.1111/jam.14386).
- [8] Aguilar A. Extremophile research in the European Union: From fundamental aspects to industrial expectations. *FEMS Microb Rev.* 1996; doi [10.1016/0168-6445\(96\)80462-1](https://doi.org/10.1016/0168-6445(96)80462-1).
- [9] Bezuidt OKI, Makhalanyane TP, Gomri MA, Kharroub K, Cowan DA. Draft genome sequence of thermophilic *Geobacillus* sp. strain Sah69, isolated from Saharan soil, Southeast Algeria. *Genome Announc.* 2015; doi:[10.1128/genomeA.01447-15](https://doi.org/10.1128/genomeA.01447-15).
- [10] Goodfellow M, Busarakam K, Idris H, Labeda DP, Nouioui I, Brown R, et al. *Streptomyces asenjonii* sp. nov., isolated from hyper-arid Atacama Desert soils and emended description of *Streptomyces viridosporus* Pridham et al. 1958. *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol.* 2017; doi:[0.1007/s10482-017-0886-7](https://doi.org/0.1007/s10482-017-0886-7).
- [11] Abdelkader MSA, Philippon T, Asenjo JA, Bull AT, Goodfellow M, Ebel R, et al. Asenjonamides A-C, antibacterial metabolites isolated from *Streptomyces asenjonii* strain KNN 42.f from an extreme-hyper arid Atacama Desert soil. *J Antibiot.* 2018; doi:[10.1038/s41429-017-0012-0](https://doi.org/10.1038/s41429-017-0012-0).

- [12] Meklat A, Bouras N, Zitouni A, Mathieu F, Lebrihi A, Schumann P, et al. *Actinopolyspora mzabensis* sp. nov., a halophilic actinomycete isolated from an Algerian Saharan soil. *Int J Syst Evol Microbiol*. 2013;
- [13] Meklat A, Bouras N, Zitouni A, Mathieu F, Lebrihi A, Schumann P, Spröer C, Klenk HP, Sabaou N. *Actinopolyspora righensis* sp. nov., a novel halophilic actinomycete isolated from Saharan soil in Algeria. *Antonie Van Leeuwenhoek*. 2013; doi: 10.1007/s10482-013-9948-7.
- [14] Meklat A, Bouras N, Zitouni A, Mathieu F, Lebrihi A, Schumann P, et al. *Actinopolyspora saharensis* sp. nov., a novel halophilic actinomycete isolated from a Saharan soil of Algeria. *Antonie van Leeuwenhoek*. 2013; doi: 10.1007/s10482-012-9859-z.
- [15] Saker R, Bouras N, Zitouni A, Ghouil M, Rohde M, Schumann P, et al. *Mzabimyces algeriensis* gen. nov., sp. nov., a halophilic filamentous actinobacterium isolated from a Saharan soil, and proposal of *Mzabimycetaceae* fam. nov. *Antonie van Leeuwenhoek*. 2014; doi: 10.1007/s10482-014-0271-8.
- [16] Meklat A, Bouras N, Mokrane S, Zitouni A, Schumann P, Spröer C, et al. *Bounagaea algeriensis* gen. nov., sp. nov., an extremely halophilic actinobacterium isolated from a Saharan soil of Algeria. *Antonie van Leeuwenhoek*. 2015; doi: 10.1007/s10482-015-0500-9.
- [17] Lahoum A, Bouras N, Verheecke C, Mathieu F, Schumann P, Spröer C, et al. *Actinomadura adraensis* sp. Nov., an actinobacterium isolated from saharan soil. *Int J Syst Evol Microbiol*. 2016; doi: 10.1099/ijsem.0.001114.
- [18] Fierer N, Leff JW, Adams BJ, Nielsen UN, Bates ST, Lauber CL, et al. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc Natl Acad Sci U S A*. 2012; doi: 10.1073/pnas.1215210110.
- [19] Reghiooua S, Boughachiche F, Zerizer H, Oulmi L, Kitouni M, Boudemagh A, et al. Activité antibactérienne d'actinomycètes rares isolés d'échantillons de sol aride du Sud-est Algérien. *Antibiotiques*. 2006; doi: 10.1016/S1294-5501(06)70814-7.
- [20] Seppey M, Manni M, Zdobnov EM. BUSCO: Assessing genome assembly and annotation completeness. *Meth Mol Biol*. 2019; doi:10.1007/978-1-4939-9173-0\_14.
- [21] Rodriguez-R L, Konstantinidis K. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints*. 2016; doi : 10.7287/peerj.preprints.1900v1.
- [22] Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics*. 2016; doi: 10.1093/bioinformatics/btv681.
- [23] Gupta AB, Uchida A, Purbojati RW, Wong A, Kushwaha KK, Putra A, et al. Complete Genome Sequence of *Streptomyces* sp. Strain SGAir0924, an Actinobacterium Isolated from Outdoor Air in Singapore. *Microbiol Resour Announc*. 2019; doi: 10.1128/mra.00899-19.
- [24] Nguyen H Van, Truong PM, Duong HT, Dinh HM, Nguyen CH. Genome sequence data of *Streptomyces* sp. SS52, an endophytic strain for daidzein biosynthesis. *Data Br*. 2019; doi: 10.1016/j.dib.2019.104746.
- [25] Grant JR, Stothard P. The CGView Server: a comparative genomics tool for circular genomes. *Nucl Acids Res*. 2008; doi:10.1093/nar/gkn179.
- [26] Xu L, Dong Z, Fang L, Luo Y, Wei Z, Guo H, et al. OrthoVenn2: A web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucl Acids Res*. 2019; doi : 10.1093/nar/gkz333.
- [27] Balakrishnan R, Harris MA, Huntley R, Van Auken K, Michael Cherry J. A guide to best practices for gene ontology (GO) manual annotation. *Database*. 2013; doi.org/10.1093/database/bat054.

- [28] Juhas M, Van Der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. Genomic islands: Tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiology Reviews*. 2009; doi: 10.1111/j.1574-6976.2008.00136.x
- [29] Dobrindt U, Hochhut B, Hentschel U, Hacker J. Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol*. 2004; doi : [10.1038/nrmicro884](https://doi.org/10.1038/nrmicro884).
- [30] Cuecas A, Kanoksilapatham W, Gonzalez JM. Evidence of horizontal gene transfer by transposase gene analyses in *Fervidobacterium* species. *PLoS One*. 2017; doi: [10.1371/journal.pone.0173961](https://doi.org/10.1371/journal.pone.0173961).
- [31] Sinninghe Damsté JS, Schouten S, Van Duin ACT. Isorenieratene derivatives in sediments: Possible controls on their distribution. *Geochim Cosmochim Acta*. 2001; doi: 10.1016/S0016-7037(01)00549-X.
- [32] Maresca JA, Romberger SP, Bryant DA. Isorenieratene biosynthesis in green sulfur bacteria requires the cooperative actions of two carotenoid cyclases. *J Bacteriol*. 2008; doi: 10.1128/JB.00758-08.
- [33] Szczeblewski P, Laskowski T, Kubacki B, Dziergowska M, Liczmańska M, Grynda J, et al. Analytical studies on ascosin, candicidin and levorin multicomponent antifungal antibiotic complexes. the stereostructure of ascosin A2. *Sci Rep*. 2017; doi: [10.1038/srep40158](https://doi.org/10.1038/srep40158).
- [34] Schulz A, Hermann L, Freibert SA, Bönig T, Hoffmann T, Riclea R, et al. Transcriptional regulation of ectoine catabolism in response to multiple metabolic and environmental cues. *Environ Microbiol*. 2017; doi:[10.1111/1462-2920.13924](https://doi.org/10.1111/1462-2920.13924).
- [35] Yamada Y, Kuzuyama T, Komatsu M, Shin-ya K, Omura S, Cane DE, et al. Terpene synthases are widely distributed in bacteria. *Proc Natl Acad Sci U S A*. 2015; doi:[10.1073/pnas.1422108112](https://doi.org/10.1073/pnas.1422108112).
- [36] Jüttner F, Watson SB. Biochemical and ecological control of geosmin and 2-methylisoborneol in source waters. *Appl Environ Microbiol*. 2007; doi:[10.1128/AEM.02250-06](https://doi.org/10.1128/AEM.02250-06).
- [37] Reinert DJ, Balliano G, Schulz GE. Conversion of Squalene to the Pentacarbocyclic Hopene. *Chem Biol*. 2004; doi: 10.1016/j.chembiol.2003.12.013.
- [38] Challis GL, Ravel J. Coelichelin, a new peptide siderophore encoded by the *Streptomyces coelicolor* genome: Structure prediction from the sequence of its non-ribosomal peptide synthetase. *FEMS Microbiol Lett*. 2000; doi: [10.1016/S0378-1097\(00\)00184-1](https://doi.org/10.1016/S0378-1097(00)00184-1).
- [39] Zhao B, Moody SC, Hider RC, Lei L, Kelly SL, Waterman MR, et al. Structural analysis of cytochrome P450 105N1 involved in the biosynthesis of the zincophore, coelibactin. *Int J Mol Sci*. 2012; doi:[10.3390/ijms13078500](https://doi.org/10.3390/ijms13078500).
- [40] Kodani S, Hudson ME, Durrant MC, Buttner MJ, Nodwell JR, Willey JM. The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene ramS in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A*. 2004; doi: 10.1073/pnas.0404220101.
- [41] Ji Z, Wang M, Wei S, Zhang J, Wu W. Isolation, structure elucidation and antibacterial activities of streptothricin acids. *J Antibiot*. 2009; doi:[10.1038/ja.2009.16](https://doi.org/10.1038/ja.2009.16).
- [42] Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucl Acids Res*. 2014; doi:[10.1093/nar/gkt1226](https://doi.org/10.1093/nar/gkt1226)
- [43] Hecker M, Völker U. General stress response of *Bacillus subtilis* and other bacteria. *Advances in Microbial Physiology*. 2001; doi: [10.1016/s0065-2911\(01\)44011-2](https://doi.org/10.1016/s0065-2911(01)44011-2).

- [44] Holtmann G, Brigulla M, Steil L, Schütz A, Barnekow K, Völker U, et al. RsbV-independent induction of the SigB-dependent general stress regulon of *Bacillus subtilis* during growth at high temperature. *J Bacteriol.* 2004; doi:10.1128/JB.186.18.6150-6158.2004.
- [45] De Las Peñas A, Connolly L, Gross CA. The  $\sigma(E)$ -mediated response to extracytoplasmic stress in *Escherichia coli* is transduced by RseA and RseB, two negative regulators of  $\sigma(E)$ . *Mol Microbiol.* 1997 Apr 1;24(2):373–85; doi:10.1046/j.1365-2958.1997.3611718.x.
- [46] Servant P, Mazodier P. Characterization of *Streptomyces albus* 18-kilodalton heat shock- responsive protein. *J Bacteriol.* 1995; doi: 10.1128/jb.177.11.2998-3003.1995.
- [47] Kedzierska S, Akoev V, Barnett ME, Zolkiewski M. Structure and Function of the Middle Domain of ClpB from *Escherichia coli*. *Biochemistry.* 2003; doi:10.1021/bi035573d.
- [48] Keto-Timonen R, Hietala N, Palonen E, Hakakorpi A, Lindström M, Korkeala H. Cold Shock Proteins: A Minireview with Special Emphasis on Csp-family of Enteropathogenic *Yersinia*. *Front Microbiol.* 2016; doi: 10.3389/fmicb.2016.01151.
- [49] Hu MX, Zhang X, Li EL, Feng YJ. Recent advancements in toxin and antitoxin systems involved in bacterial programmed cell death. *Int J Microbiol.* 2010; doi:10.1155/2010/781430.
- [50] Mohammadipanah F, Wink J. Actinobacteria from arid and desert habitats: Diversity and biological activity. *Front Microbiol.* 2016; doi: 10.3389/fmicb.2015.01541.
- [51] Zitouni A, Boudjella H, Lamari L, Badji B, Mathieu F, Lebrihi A, et al. Nocardiosis and *Saccharothrix* genera in Saharan soils in Algeria: Isolation, biological activities and partial characterization of antibiotics. *Res Microbiol.* 2005; doi:10.1016/j.resmic.2005.05.006.
- [52] Merrouche R, Bouras N, Coppel Y, Mathieu F, Monje MC, Sabaou N, et al. Dithiopyrrolone antibiotic formation induced by adding valeric acid to the culture broth of *Saccharothrix algeriensis*. *J Nat Prod.* 2010; doi :10.1021/np900808u.
- [53] Liao C, Liu B, Xu Y, Li Y, Li H. Effect of topography and protecting barriers on revegetation of sandy land, Southern Tibetan Plateau. *Sci Rep.* 2019; doi:10.1038/s41598-019-43034-8.
- [54] Eglinton TI, Eglinton G, Dupont L, Sholkovitz ER, Montluçon D, Reddy CM. Composition, age, and provenance of organic matter in NW African dust over the Atlantic Ocean. *Geochem Geophys Geosyst.* 2002; doi10.1029/2001gc000269.
- [55] Behzad H, Mineta K, Gojobori T. Global Ramifications of Dust and Sandstorm Microbiota. *Genome Biol Evol.* 2018; doi: 10.1093/gbe/evy134.
- [56] Namitha KK, Negi PS. Chemistry and biotechnology of carotenoids. *Crit Rev Food Sci Nutr.* 2010; doi: 10.1080/10408398.2010.499811.
- [57] Najah S, Chong TM, Gerbaud C, Chan KG, Mellouli L, Pernodet JL. Complete genome sequence of *Streptomyces* sp. TN58, a producer of acyl alpha-L-rhamnopyranosides. *Genome Announc.* 2017; doi: 10.1128/genomeA.00828-17.
- [58] Herdini C, Mubarika S, Hariwiyanto B, Wijayanti N, Hosoyama A, Yamazoe A, et al. Secondary bioactive metabolite gene clusters identification of anticandida-producing *Streptomyces* Sp. GMR22 isolated from Wanagama forest as revealed by Genome mining approach. *Indones J Pharm.* 2017; doi: 10.14499/indonesianjpharm28iss1pp26.
- [59] Amin DH, Abolmaaty A, Borsetto C, Tolba S, Abdallah NA, Wellington EMH. In silico genomic mining reveals unexplored bioactive potential of rare actinobacteria isolated from Egyptian soil. *Bull Natl Res Cent.* 2019; doi: 10.1186/s42269-019-0121-y.

- [60] Gosse JT, Ghosh S, Sproule A, Overy D, Cheeptham N, Boddy CN. Whole genome sequencing and metabolomic study of cave *Streptomyces* isolates ICC1 and ICC4. *Front Microbiol.* 2019; doi: 10.3389/fmicb.2019.01020
- [61] Guerrero-Garzón JF, Zehl M, Schneider O, Rückert C, Busche T, Kalinowski J, et al. *Streptomyces* spp. From the Marine Sponge *Antho dichotoma*: Analyses of Secondary Metabolite Biosynthesis Gene Clusters and Some of Their Products. *Front Microbiol.* 2020; doi: 10.3389/fmicb.2020.00437.
- [62] Belknap KC, Park CJ, Barth BM, Andam CP. Genome mining of biosynthetic and chemotherapeutic gene clusters in *Streptomyces* bacteria. *Sci Rep.* 2020; doi 10.1038/s41598-020-58904-9
- [63] Roszak DB, Colwell RR. Survival strategies of bacteria in the natural environment. *Microbiol Rev.* 1987; doi 10.1128/mmr.51.3.365-379.1987.
- [64] Moeller R, Horneck G, Facius R, Stackebrandt E. Role of pigmentation in protecting *Bacillus* sp. endospores against environmental UV radiation. *FEMS Microbiol Ecol.* 2005; doi: 10.1016/j.femsec.2004.08.008.
- [65] Hullo MF, Moszer I, Danchin A, Martin-Verstraete I. CotA of *Bacillus subtilis* is a copper-dependent laccase. *J Bacteriol.* 2001; doi: 10.1128/JB.183.18.5426–5430.2001.
- [66] Selama O, Amos GCA, Djenane Z, Borsetto C, Laidi RF, Porter D, et al. Screening for genes coding for putative antitumor compounds, antimicrobial and enzymatic activities from haloalkalitolerant and haloalkaliphilic bacteria strains of Algerian Sahara soils. *Biomed Res Int.* 2014; doi:10.1155/2014/317524.
- [67] A. Belov A, S. Cheptsov V, A. Vorobyova E. Soil bacterial communities of Sahara and Gibson deserts: Physiological and taxonomical characteristics. *AIMS Microbiol.* 2018; doi: [10.3934/microbiol.2018.4.685](https://doi.org/10.3934/microbiol.2018.4.685).
- [68] Liang P, Saqib HSA, Ni X, Shen Y. Long-read sequencing and de novo genome assembly of marine medaka (*Oryzias melastigma*). *BMC Genomics.* 2020; doi: 10.1186/s12864-020-07042-7.
- [69] Mohamed H, Miloud B, Zohra F, García-Arenzana JM, Veloso A, Rodríguez-Couto S. Isolation and Characterization of Actinobacteria from Algerian Sahara Soils with Antimicrobial Activities. *Int J Mol Cell Med.* 2017; doi: 10.22088/acadpub.BUMS.6.2.5.
- [70] Nanjwade BK, Chandrashekhara S, Shamarez AM, Goudanavar PS, Manvi F V. Isolation and morphological characterization of antibiotic producing actinomycetes. *Trop J Pharm Res.* 2010; doi: 10.4314/tjpr.v9i3.56282.
- [71] Shirling EB, Gottlieb D. Methods for characterization of *Streptomyces* species. *Int J Syst Bacteriol.* 1966; doi 10.1099/00207713-16-3-313.
- [72] Pikovskaya, RI., Mobilization of phosphorus in soil connection with the vital activity of some microbial species. *Microbiology.* 1948;17:362-370 .
- [73] Dharumadurai Dhanasekaran, Yi Jiang, *Actinobacteria: Basics and Biotechnological Applications*, IntechOpen; 2016; doi: 10.5772/60457.
- [74] Jacobsen CN, Nielsen VR, Hayford AE, Møller PL, Michaelsen KF, Pærregaard A, et al. Screening of probiotic activities of forty-seven strains of *Lactobacillus* spp. by in vitro techniques and evaluation of the colonization ability of five selected strains in humans. *Appl Environ Microbiol.* 1999; doi:10.1128/AEM.65.11.4949-4956.1999.
- [75] De la Cruz TE, Torres JM. Gelatin Hydrolysis Test Protocol. *Am Soc Microbiol.* 2012; <https://www.asmscience.org/content/education/protocol/protocol.3776>. Accessed 29.9.2020.

- [76] Selvin J, Shanmughapriya S, Gandhimathi R, Seghal Kiran G, Rajeetha Ravji T, Natarajaseenivasan K, et al. Optimization and production of novel antimicrobial agents from sponge associated marine actinomycetes *Nocardioopsis dassonvillei* MAD08. *Appl Microbiol Biotechnol*. 2009; doi: [10.1007/s00253-009-1878-y](https://doi.org/10.1007/s00253-009-1878-y).
- [77] Chandwad S, Gutte S. Screening of Actinomycetes for  $\alpha$ -amylase Inhibitors Production. *Curr Enzym Inhib*. 2019; doi : [10.2174/1573408015666190110130544](https://doi.org/10.2174/1573408015666190110130544)
- [78] Oumer OJ, Abate D. Screening and Molecular Identification of Pectinase Producing Microbes from Coffee Pulp. Freire D, editor. *Biomed Res Int*.2018; doi: [10.1155/2018/2961767](https://doi.org/10.1155/2018/2961767).
- [79] Adesina FC, Onilude AA. Isolation, identification and screening of xylanase and glucanase-producing microfungi from degrading wood in Nigeria. *African J Agric Res*. 2013; doi:[10.5897/AJAR2013.699](https://doi.org/10.5897/AJAR2013.699).
- [80] Alves-Prado HF, Pavezzi FC, Leite RSR, de Oliveira VM, Sette LD, DaSilva R. Screening and Production Study of Microbial Xylanase Producers from Brazilian Cerrado. *Appl Biochem Biotechnol* 2010;161(1):333–46; doi:[10.1007/s12010-009-8823-5](https://doi.org/10.1007/s12010-009-8823-5).
- [81] Hankin L, Anagnostakis SL. The Use of Solid Media for Detection of Enzyme Production by Fungi. *Mycologia*. 1975; doi: [10.2307/3758395](https://doi.org/10.2307/3758395).
- [82] Yuratmoko D, Rachmania Mubarik N, Meryandini A. Screening of Proteolytic Enzymes of *Streptomyces* sp. Local Strain and Their Characterization. *Microbiol Indones*. 2007; [doi.org/10.5454/mi.1.2.5](https://doi.org/10.5454/mi.1.2.5).
- [83] Durham DR, Stewart DB, Stellwag EJ. Novel alkaline- and heat-stable serine proteases from alkalophilic *Bacillus* sp. strain GX6638. *J Bacteriol*. 1987; doi: [10.1128/jb.169.6.2762-2768.1987](https://doi.org/10.1128/jb.169.6.2762-2768.1987).
- [84] Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*. 2016; doi:[10.1093/molbev/msw054](https://doi.org/10.1093/molbev/msw054).
- [85] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994; doi: [10.1093/nar/22.22.4673](https://doi.org/10.1093/nar/22.22.4673).
- [86] Felsenstein J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution*. 1985; doi: [10.1111/j.1558-5646.1985.tb00420.x](https://doi.org/10.1111/j.1558-5646.1985.tb00420.x).
- [87] Seemann T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*. 2014; doi:[10.1093/bioinformatics/btu153](https://doi.org/10.1093/bioinformatics/btu153).
- [88] Aziz RK, Bartels D, Best A, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*. 2008; doi [10.1186/1471-2164-9-75](https://doi.org/10.1186/1471-2164-9-75).
- [89] Lowe TM, Eddy SR. tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. *Nucleic Acids Res*. 1997; doi: [10.1093/nar/25.5.955](https://doi.org/10.1093/nar/25.5.955)
- [90] Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, et al. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucl Acids Res*. 2017; doi: [10.1093/nar/gkx343](https://doi.org/10.1093/nar/gkx343).
- [91] Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucl Acids Res*. 2016; doi: [10.1093/nar/gkw387](https://doi.org/10.1093/nar/gkw387).

- [92] Wu S, Zhu Z, Fu L, Niu B, Li W. WebMGA: A customizable web server for fast metagenomic sequence analysis. *BMC Genomics*. 2011; doi: 10.1186/1471-2164-12-444.
- [93] Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucl Acids Res*. 2004; doi 10.1093/nar/gkh063.
- [94] Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. DbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucl Acids Res*. 2018; doi: 10.1093/nar/gky418.
- [95] Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, et al. AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucl Acids Res*. 2019; doi: [10.1093/nar/gkz310](https://doi.org/10.1093/nar/gkz310).

## Tables

<b>Table 1.</b> General features of the strains <i>Babs14</i> and <i>Osf17</i> according to MIGS mandatory recommendations.		
Item	Description	
	Babs14	Osf17
<b>Classification</b>	Domain <i>Bacteria</i>	Domain <i>Bacteria</i>
	Phylum <i>Actinobacteria</i>	Phylum <i>Actinobacteria</i>
	Class <i>Actinobacteria</i>	Class <i>Actinobacteria</i>
	Order <i>Streptomyces</i>	Order <i>Streptomyces</i>
	Family <i>Streptomyces</i>	Family <i>Streptomyces</i>
	Genus <i>Streptomyces</i>	Genus <i>Streptomyces</i>
	Species <i>Streptomyces sp.</i>	Species <i>Streptomyces sp.</i>
<b>Type strain</b>	Wild-type strain	Wild-type strain
Gram stain	Positive	Positive
Cell shape	Filamentous	Filamentous
Motility	Non-motile	Non-motile
Pigmentation	Beige-cream	White-grey
Sporulation	Sporulating	Sporulating
Temperature range	20-40°C	20-40°C
Salinity range	0–5%	0–5%
pH range	6.0–10.0	6.0–10.0
Catalase	Positive	Positive
Submitted to NCBI	Accession not attributed yet	Accession not attributed yet
BioProject ID	PJNA665615	PRJNA665615
Collection date	April 2014	October 2017
Latitude and longitude	30°07'57.9"N 2°10'38.6"W	<b>33°29'32.0"N</b> 6°53'46.7"E
Geographic location name	Algeria: Beni Abbes	Algeria: El Oued Souf
Environment biome	Desert	Desert
Environment material	Soil of sand dunes	Soil of sand dunes
Depth	5–20 cm	5–20 cm
Biotic relationship	Free-living	Free-living

**Table 2.** Screening of antibacterial, antifungal, and enzymatic activities of Babs14 and Osf17.

Strains	Antibacterial and antifungal activities: Diameter of inhibition (mm)			Enzymatic activities of the strains: Diameter of clear zone (mm)					
	<i>E. coli</i> ATCC 25922	<i>S. aureus</i> ATCC 43300	<i>A. niger</i> ATCC 6275	Proteolytic	Lipolytic	Amylolytic	Pectinolytic	Cellulolytic	Xylanolytic
Babs14	12 mm	15 mm	13 mm	24 mm	10 mm	12 mm	14 mm	16 mm	10 mm
Osf17	26 mm	20 mm	13 mm	-	8 mm	14 mm	20 mm	12 mm	12 mm

**Table 3.** Genome features of *Streptomyces* sp. strains Babs14 and Osf17.

Feature	Chromosome characteristics	
	Babs14	Osf17
Genome topology	Linear	Linear
Chromosome size (bp)	7999102	7967258
GC content (%)	72,6	72,6
Protein-coding genes	6910	6894
Contigs	2	2
rRNA genes	18	18
tRNA genes	87	86
tmRNA genes	1	1
Plasmids	0	3
Secondary metabolite gene clusters BGCs	29	29
Genes assigned to COG	1865	1852
Genes assigned to KEGG	3121	3099
Features assigned by RAST		7241
Gene entries assigned to CAZy	372	370

**Table 4.** Orthologous protein clusters and singletons of the Sahara strains and the closest relatives from OrthoVenn2 analysis.

Strain	Proteins	Clusters	Singletons
Babs14	6910	6807	39
Osf17	6894	6802	48
<i>Streptomyces sp.</i> SGAir_0924	6603	6269	288
<i>Streptomyces sp.</i> SS52	7152	6701	363

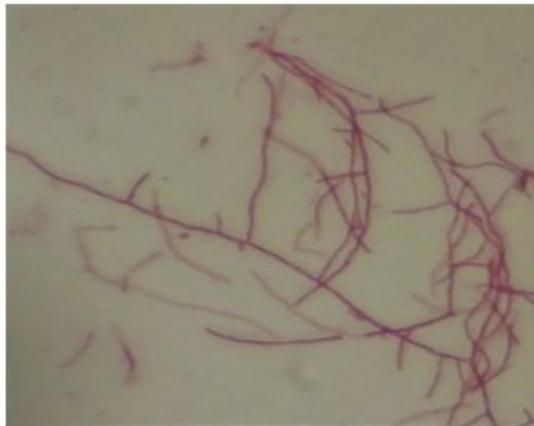
**Table 5.** Potential Biosynthetic Gene clusters assigned in Babs14 and Osf17 showing 100% similarity with the known clusters predicted by AntiSMASH.

Region	Type	Most similar known cluster
Cluster 1	terpene	isorenieratene
Cluster 7	ectoine	ectoine
Cluster 13	terpene	albaflavenone
Cluster 17	terpene	geosmin
Cluster 20	terpene	hopene
Cluster 24	NRPS	coelichelin
Cluster 25	NRPS	coelibactin
Cluster 27	lanthipeptide	SapB

**Table 6.** The COG functional classes. Abbreviations for Figure 5.

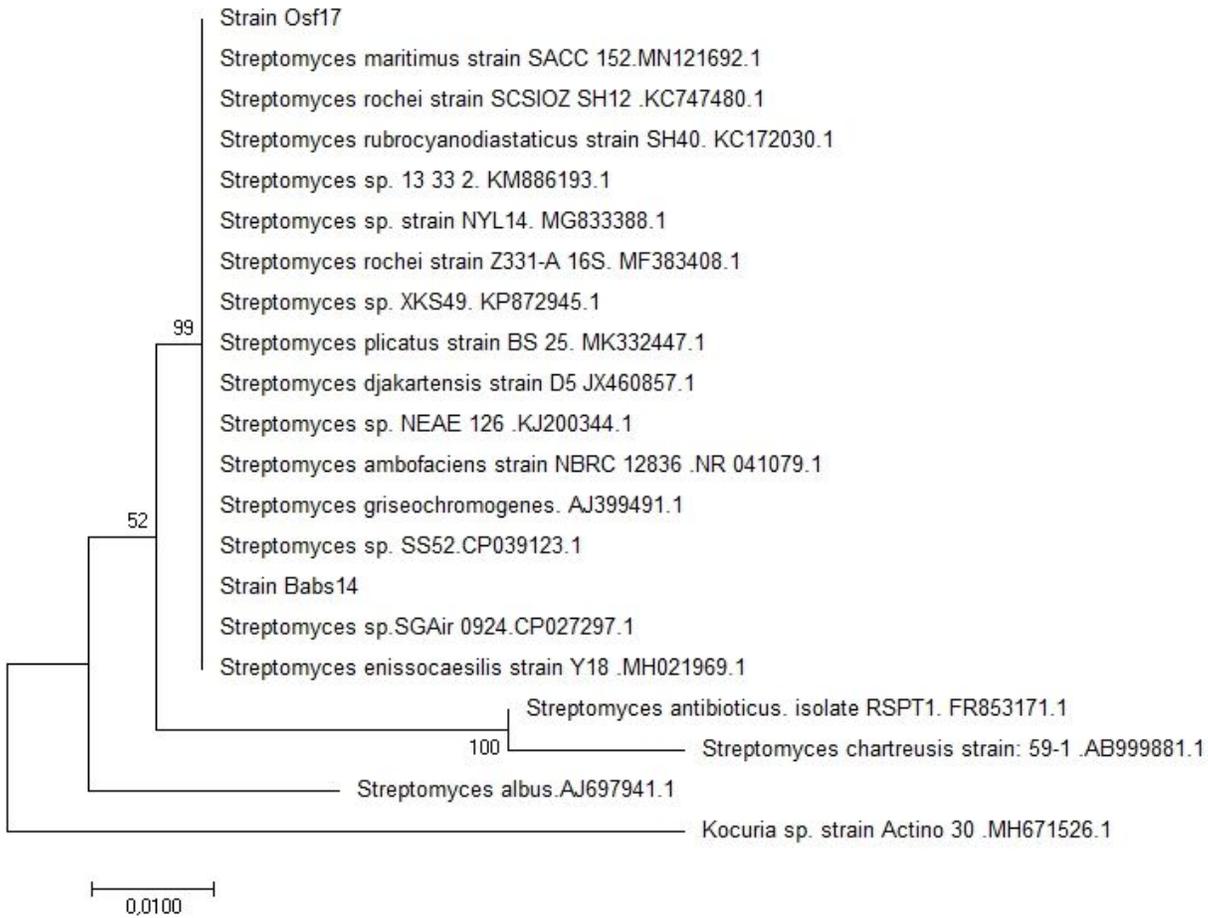
<b>Functional class</b>	<b>Class description</b>
A	RNA processing and modification
B	Chromatin structure and dynamics
C	Energy production and conversion
D	Cell cycle control /cell division / chromosome partitioning
E	Amino acid transport and metabolism
F	Nucleotide transport and metabolism
G	Carbohydrate transport and metabolism
H	Coenzyme transport and metabolism
I	Lipid transport and metabolism
J	Translation / ribosomal structure and biogenesis
K	Transcription
L	Replication / recombination, and repair
M	Cell wall/membrane/envelope biogenesis
N	Cell motility
O	Posttranslational modification / protein turnover/ chaperones
P	Inorganic ion transport and metabolism
Q	Secondary metabolites biosynthesis, transport, and catabolism
R	General function prediction only
S	Function unknown
T	Signal transduction mechanisms
U	Intracellular trafficking / secretion, and vesicular transport
V	Defense mechanisms
W	Extracellular structures
Y	Nuclear structure
Z	Cytoskeleton

## Figures

**A****B****C****D**

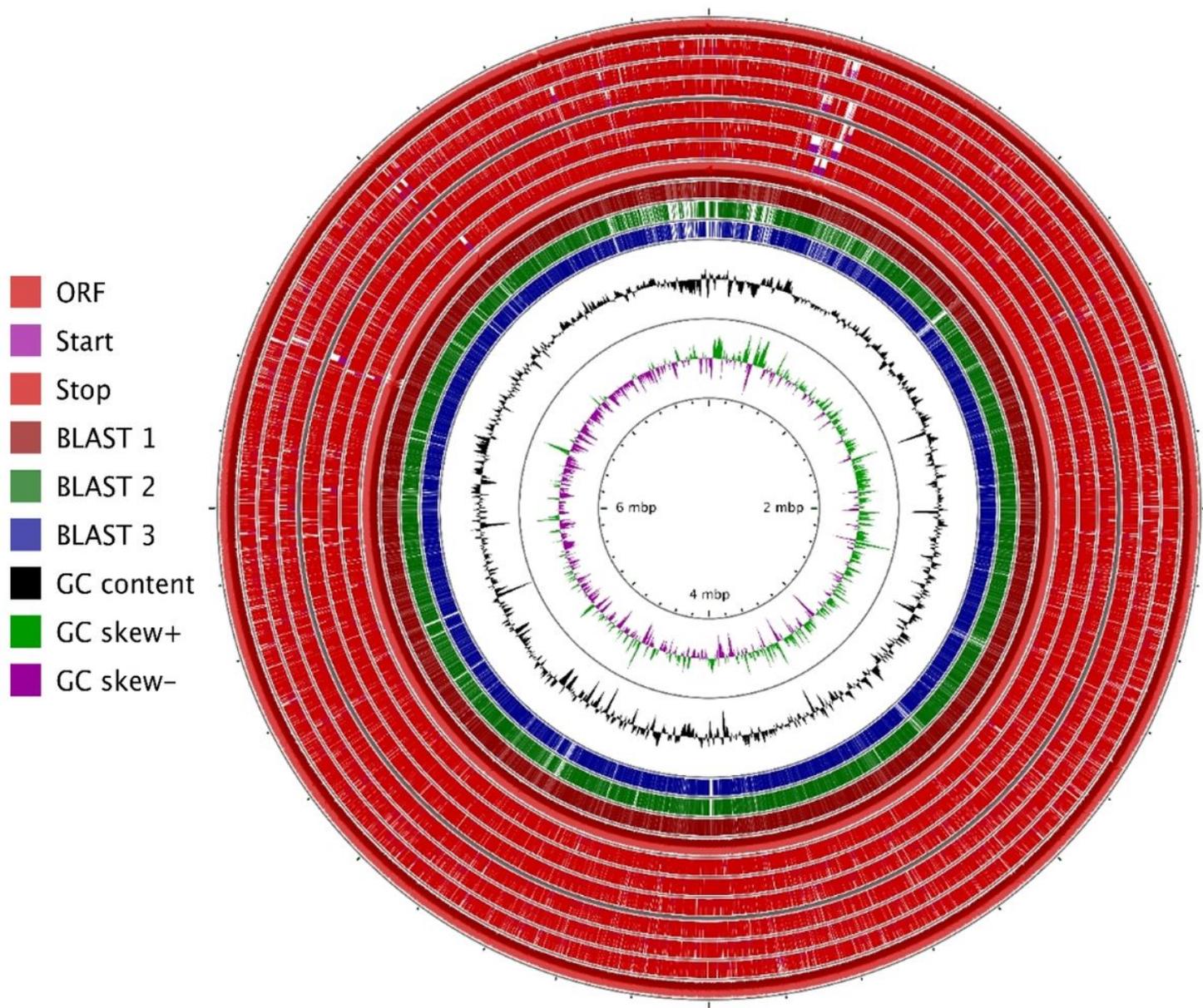
**Figure 1**

Morphological characteristics of the strains Babs14 (A, B) and Osf17 (C, D). Macroscopic morphology on ISP2 agar (pH 7) after incubation of 7 days (A, C) and microscopic filamentous morphology of the mycelium under light microscopy (magnification 10x100) stained with Fuchsin (C, D).



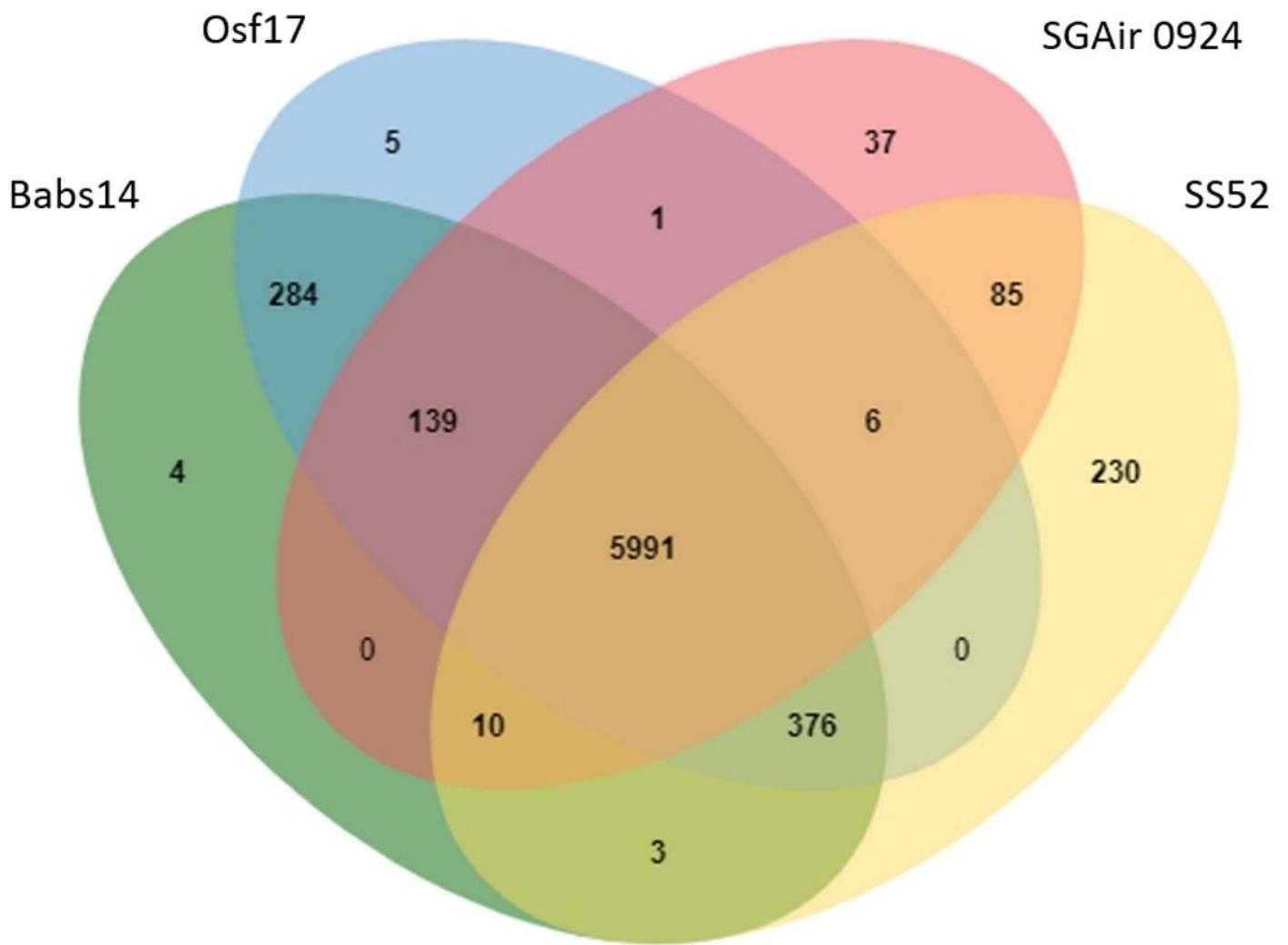
**Figure 2**

Molecular Phylogenetic analysis of the 16S rRNA sequences of the strains Babs14 and Osf17. The tree was constructed by MEGA7 using Maximum Likelihood method based on the Kimura 2-parameter model. Gamma distribution was used to model evolutionary rate differences among sites (+G, parameter = 0,7241). The rate variation model allowed for some sites to be evolutionarily invariable [+I], 40,17% of the sites. Bootstrap values are expressed as percentages of 1000 replication and the scale bar indicates 0.01 substitutions per site.



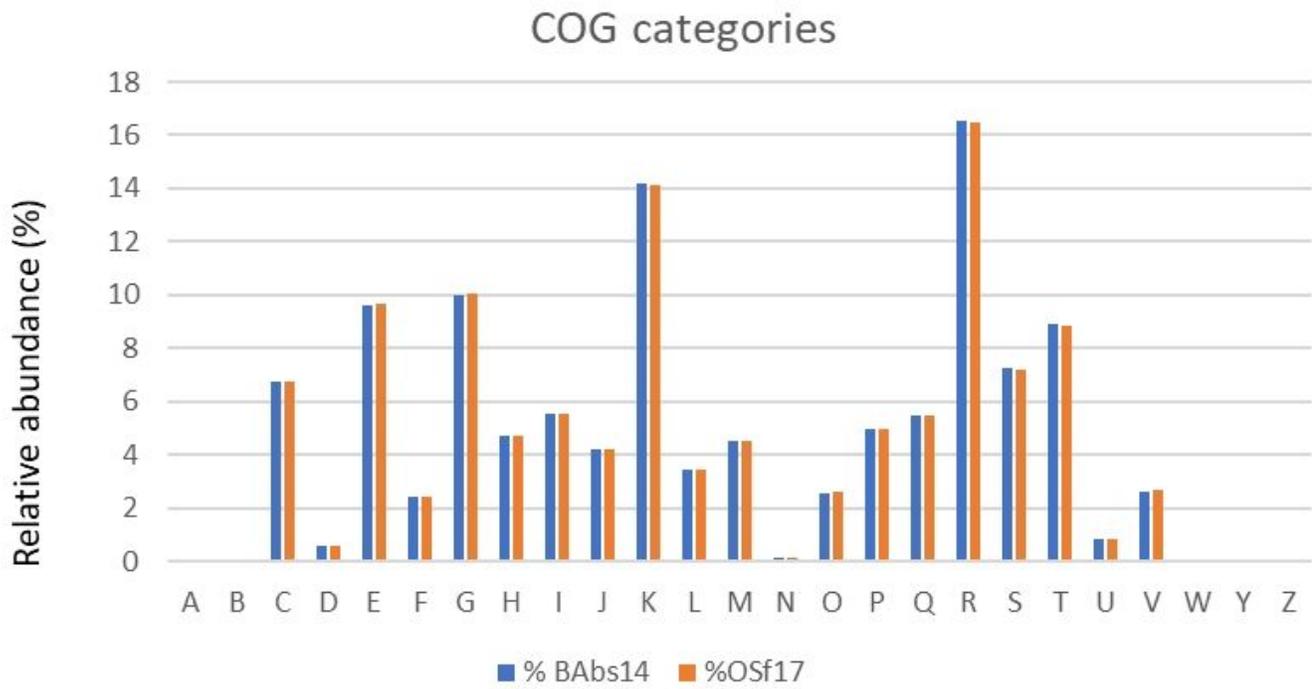
**Figure 3**

Genomic maps. Full view of the genome of Babs14 as a reference sequence blasted with *Osf17*, *Streptomyces* sp. SGAir 0924 and *Streptomyces* sp. SS52 using CGViewer Server V1.0 [39] is shown. From outside to inside: Ring 1 is the open reading frame of the forward strand (ORF). Rings 2,3 and 4 indicate the start and stop codons of the forward strand of Babs14. Rings 5,6 and 7 indicate the start and stop codons of the reverse strand of Babs14. Ring 8 is the ORF of the reverse strand of the primary sequence of Babs14. Ring 9 corresponds to the Blast 1 of the hits of Babs14 with *Osf17*. Ring 10 indicates the Blast 2 and shows the hits with *Streptomyces* sp. SGAir 0924 (CP027297.1). Ring 11 is the Blast 3 and corresponds to the hits with *Streptomyces* sp. SS52 (NZ\_CP039123). Ring 12 and 13 show GC content and GC skew, respectively, of the reference sequence Babs14. The map shows the Blast comparison result (BLASTN) with the primary DNA sequence of Babs14. The BLAST results are drawn at partial opacity. The darker regions indicate the multiple hits to the corresponding region of the reference sequence Babs14



**Figure 4**

Venn diagram of the number of shared and unique proteins between Babs14, Osf17 and the closest strains *Streptomyces* sp. SGAir0924 (CP027297.1) and *Streptomyces* sp. SS52(NZ\_CP039123) as created by OrthoVenn2.



**Figure 5**

COG database annotation of Babs14 and Osf17. The relative abundances of proteins (%) in the two genomes is shown. The letter codes are described in Table 6.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Zeroukietalsupplementalmaterial28.9.20.docx](#)