

Identity of carotenoid biosynthetic genes of *Streptomyces* and their activation in *S. globisporus* 1912-4Crt

Bohdan P. Matselyukh*

Department of Genetics of Microorganisms, DK Zabolotny Institute of Microbiology and Virology, National Academy of Sciences of Ukraine, Kyiv, Ukraine

*Corresponding author: Bohdan P. Matselyukh, Department of Genetics of Microorganisms, DK Zabolotny Institute of Microbiology and Virology, National Academy of Sciences of Ukraine, Kyiv, Ukraine, Tel: (380)-44-5226762; E-mail: bohdan.matselyukh@gmail.com

Received date: May 03, 2019; Accepted date: May 15, 2019; Published date: May 22, 2019

Copyright: © 2019 Matselyukh BP. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

High identity of the carotenoid biosynthetic gene (*crt*) sequences (92–95%) was shown in *S. globisporus* 1912, *S. globisporus* TFH56 and *S. griseus* NBRC 13350. The homology of the *crt* genes of the other 53 *Streptomyces* strains was lower, ranging from 75% to 88%. Two direct and non-punctual repeats (NPRs) of 21 bp and spontaneous deletion of 119 bp, including the sequence of 96 bp between two non-punctual repeats, one NPR from 3'-side and 2 neighboring bp, have been identified in the sequence of the lycopene cyclase gene *crtY* of the strains 1912-2 and 1912-4Crt of *S. globisporus*, correspondingly. Both NPRs begin from the same specific site (5'-GGGGCG-3') that may play a role in the site-specific recombination, appearance of deletion and activation of the carotenoid biosynthesis in *S. globisporus* 1912-4Crt. The *crtY* gene of the different species of *Streptomyces* genus and bacteria contain the not identical 5'-GGGGCG-3' site in the both NPRs. The mutant strain 1912-4Crt presents biotechnological interest for genetic selection of the active producer of beta-carotene and lycopene.

Keywords: *Streptomyces globisporus* 1912; *Crt* gene cluster; *Crt* gene identity; Non-punctual repeats; Spontaneous deletion

Introduction

Carotenoids belong to the natural pigments, synthesizing by plants, some bacteria, Streptomycetes and fungi, and playing an important role in the life of animals and humans as bio stimulants, vitamin A substitutes, antioxidants, coloring and tumor inhibiting compounds [1]. Representatives of the genus *Streptomyces* are known to be the main industrial producers of the different antibiotics widely used in medicine, veterinary and agriculture [2]. Genetic study of the different strains of *Streptomyces* showed the presence of the carotenoid biosynthetic gene clusters in their genomes in the functionally inactive state. Activation of the transcription of these cryptic *crt* genes in *S. coelicolor* A3(2) and *S. griseus* IFO 13350 requires the induction of a stress-responsible sigma factor through illumination of the culture with blue or increasing the copy number of the *crtS* gene [3,4]. In *S. globisporus* 1912 and *S. albus* J1074 the carotenoid producing mutants can appear in a spontaneous manner [5,6]. At present the genomes of many *Streptomyces* strains of the different species were sequenced and submitted to GenBank. Complete sequence of the carotenoid biosynthetic gene cluster of the strain *S. globisporus* 1912 and whole genome shotgun sequencing project of the strain *S. globisporus* 1912-4Crt were submitted to Genbank (Click: GenBank Overview; Nucleotide: *Streptomyces globisporus* strain 1912; *Streptomyces globisporus* strain 1912-4Crt (Contigs 307 and 141 FASTA).

The aim of this study was the comparative analysis of the identity of the carotenoid biosynthetic genes of the sequenced genomes of the different *Streptomyces* species and search of a possible cause of a spontaneous activation of the *crt* gene cluster in the red-orange mutant *S. globisporus* 1912-4Crt. The Biogenesis and Maturation of Exosomes

Exosomes are formed during earlier steps of endosomal pathway and released upon fusion of multiple vesicle bodies (MVBs) with the

plasma membrane. As part of cellular system for membrane traffic, formation of exosomes shares a lot of key components with endocytic process. For instance, it was shown that the formation of exosomes requires Endosomal Sorting Complex Required for Transport (ESCRT), although ESCRT-independent mechanisms exist in certain circumstances. ESCRT consists of four complexes, ESCRT-0, -I, -II and -III, plus several accessory components, such as VPS4, TSG101. ESCRT-0, together with flat clathrin coats, forms a protein network on endosomal membranes, capturing ubiquitinated cargo proteins and initiating their sorting into the MVB pathway. ESCRT-I and -II complexes was shown to be responsible for membrane deformation into buds containing sequestered cargo, while ESCRT-III provides the core membrane-remodeling activity driving MVB formation.

ALG-2-interacting protein X (ALIX) is a protein that interacts with several ESCRT proteins and is thought to be involved in the germination and shedding process. *In vitro* binding assays shown that syndecans, syntenin and ALIX form tripartite complexes (syndecan-syntenin-ALIX), which plays a specific role in the biogenesis of a major class of exosomes, the loading of exosomes with specific cargo, or both. In addition, as regulators of syntenin exosomes, the small GTPase ADP ribosylation factor 6 (ARF6) and its effector phospholipase D2 (PLD2) were found to be able to affect exosomes by controlling the budding of intraluminal vesicles (ILVs) into multivesicular bodies (MVBs). Apart from this, autophagy-related gene 5 (*ATG5*) has been shown to mediate acidification of the MVB lumen and allows MVB-PM (plasma membrane) fusion, knocking-out of *ATG5* significantly reduces exosome release. Interestingly, the *ATG12-ATG3* complex has also been found to regulate exosome biogenesis through their interaction with ALIX, indicating potential reciprocal regulation between autophagosome formation and exosome biogenesis.

Materials and Methods

Isolation of strains

The initial wild-type strain *S. globisporus* 1912 was isolated from a soil sample of Armenia and is keeping in the Ukrainian Collection of Microorganisms at the D.K. Zabolotny Institute of Microbiology and Virology of NASU as the strain *S. globisporus* Ac-2098. The strain 1912-2, the producer of 200 mg L⁻¹ of the antitumor antibiotic landomycin E, was received by action of nitrosoguanidine on the spores of 1912 strain (Figure 1a) [7]. The strain 1912-4Crt, the spontaneous mutant of 1912, produced beta-carotene (7.0 mg L⁻¹) and lycopene (3.24 mg L⁻¹) in flask culture (Figure 1b) [7].

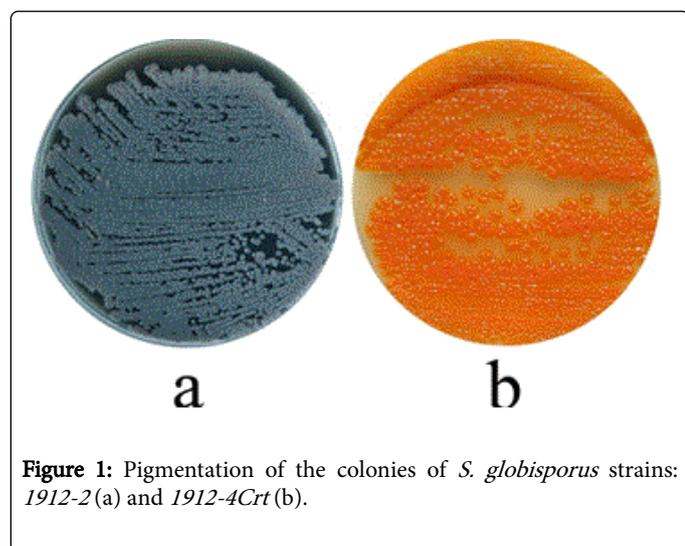


Figure 1: Pigmentation of the colonies of *S. globisporus* strains: 1912-2(a) and 1912-4Crt(b).

Isolation of chromosomal DNA

The young mycelium of the strains 1912-2 and 1912-4Crt after sedimentation by centrifugation was washed by TE buffer, pH 8.0 and resuspended in the same buffer containing 300 mg lysozyme in 1.0 mL. The procedure of mycelium lysis, DNA deproteinization and purification followed the standard protocol of the Kirby method [8]. The precipitated DNA was dissolved in TE buffer, pH 8.0 and stored at 4°C. The A260/A280 index of the DNA preparations was equal to

2.1-2.3 and electrophoresis showed one compact and high-situated strip of DNA in agarose gel. These data confirm the high molecular weight of the DNA preparations.

DNA sequencing

The sequencing of the genome DNA of the strains 1912-2 and 1912-4Crt was carried out in BaseClear B.V., Leiden, Netherlands using the following procedures. The FASTQ sequence reads were generated according to Illumina Casava pipeline version 1.8.3. Initial quality assessment was based on data passing the Illumina Chastity filtering. Subsequently, reads containing adapters and/or PhiX control signal were removed using an in-house filtering protocol. The second quality assessment was based on the remaining reads using the FASTQ quality control tool version 0.10.0. The quality of the FASTQ sequences was enhanced by trimming off low-quality bases using the “Trim sequences” option of the CLC Genomics Workbench version 6.0.4. The quality-filtered sequence reads were puzzled into a number of contig-sequences. The analysis has been performed using the “De novo assembly” option of the CLC Genomics Workbench version 6.0.4. The optimal k-mer size was automatically determined using KmerGenie [9]. The contigs were linked and placed into scaffolds or supercontigs. The orientation, order and distance between the contigs were estimated using the insert size between the paired-end and/or matepair reads. The analysis has been performed using the SSPACE Premium scaffolder version 2.3 [10]. The gapped regions within the scaffolds are (partially) closed in an automated manner using GapFiller version 1.10 [11].

This work was supported by the National Academy of Sciences of Ukraine (Grant N 0111U000606).

Results

The comparative analysis of the *crt* genes identity of the representatives of the different *Streptomyces* species was carried out by means of the BLAST program [www.ncbi.nlm.nih.gov/blast] (Table 1). High identity (92-95%) of the *crt* genes of *S. globisporus* 1912, *S. globisporus* TFH56 and *S. griseus* NBRC 13350 was shown. The *crt* genes of these three strains have also the same size. The homology of *crt* genes of the other 53 strains of *Streptomyces* was lower (71-88%). The *crtE* gene among all 7 genes of the *crt* cluster has the lowest identity (72-85%).

NN	Strain (GenBank accession N)	<i>Crt</i> genes, direction of transcription (arrows), bp N and identity with 1912 <i>Crt</i> genes (%)						
		<i>crtE</i>	<i>crrtI</i>	<i>crtB</i>	<i>crtV</i>	<i>crtU</i>	<i>crtT</i>	<i>crtY</i>
		?	?	?	?	?	?	?
1	<i>S. globisporus</i> 1912 (KM349312)	1206	1518	1029	1005	1554	717	1260
2	<i>S. globisporus</i> TFH56 (CP029361.1)	1209 (85)	1518 (93)	1029 (92)	1002 (94)	1554 (94)	717 (95)	1260 (93)
3	<i>S. griseus</i> NBRC 13350 (AP009493.1)	1209 (85)	1524 (92)	1029 (92)	1005 (95)	1554 (93)	717 (94)	1260 (93)
4	<i>S. violaceoruber</i> S21 (CP020570)	1158 (80)	1495 (92)	990 (88)	1007 (89)	1565 (86)	712 (88)	1226 (83)
5	<i>S. fulvissimus</i> DSM 40593 (CP005080.1)	1064	1475	923	990	1532	696	1226

		(74)	(86)	(83)	(83)	(85)	(81)	(78)
6	<i>S. sp. CFMR-7</i> (CP011522.1)	1172 (73)	1518 (86)	924 (83)	1007 (82)	1517 (84)	717 (79)	1148 (77)
7	<i>S. sp. S063</i> (CP021707.1)	1148 (72)	1518 (86)	921 (81)	986 (82)	11499 (82)	690 (80)	1163 (75)
8	<i>S. globisporus C-1027</i> (CP013738.1)	1054 (75)	1476 (86)	878 (76)	987 (82)	1481 (83)	698 (79)	1229 (75)
9	NN of rest tested strains	48 (72)	48 (75)	48 (75)	48 (75)	48 (75)	48 (74)	48 (71)

Table 1: Identity of *Crt* genes of the different *Streptomyces* strains.

Phylogenetic tree of the eight *crtY* genes of *Streptomyces* was constructed by means of the Blastn (Distance tree) program of GenBank (Figure 2), using the data of Table 2). One can see good

coincidence of the identity results obtained for all seven *Crt* genes of the different strains and only one *crtY* gene.

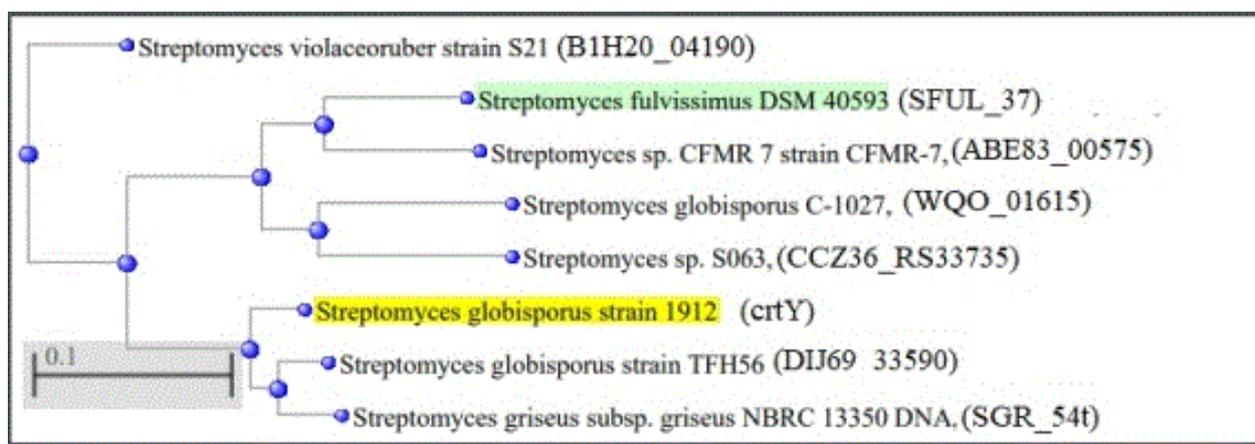


Figure 2: Phylogenetic tree of the *crtY* genes of the different *Streptomyces* strains.

Strain	GenBank accession N of <i>crtY</i> gene	Query cover%	E-value	Identity%
<i>S. globisporus</i> 1912	<i>crtY</i>	100	0	100
<i>S. globisporus TFH56</i>	DIJ69_33590	100	0	93.54
<i>S. griseus</i> NBRC 13350	SGR_54t	100	0	92.9
<i>S. violaceoruber</i> S21	B1H20_04190	98	0	82.64
<i>S. fulvissimus</i> DSM 40593	SFUL_37	98	0	77.03
<i>S. sp. CFMR-7</i>	ABE83_00575	98	0	75.1
<i>S. sp. S063</i>	CCZ36_RS33735	93	0	76.11
<i>S. globisporus C-1027</i>	WQO_01615	93	0	75.41

Table 2: Identity of *crtY* genes of *Streptomyces* with 1912 *crtY* gene.

The *crtY* gene encoding enzyme lycopene cyclase in *1912-2* strain contains two direct and non-punctual repeats (NPRs, underlined) of 21 bp, flanking the sequence of 96 bp. These two NPRs differ one from another by the four underlined bases. The specific sites of the first 6 bp (5'-GGGGCG-3') are the same in both NPRs of the strain *1912-2* and contain substitutions of some base pair in the *crtY* gene sequence of the other *Streptomyces* strains (Table 3).

641
TCGGGCTGGGGCGTGCGGAAGTCCATCAGGTCCGGTACTG
680

681
CCGGGTCGAACACGGGCCGTTTCGGTTCTGACGAACCAGCC
720

721
GGAGAAGTGCTGGAGCAGGGTGGTGC GGCGGGCAGG
760

761
ACCCGGGGCGGGCGGGAGTCGAACACGTAGCGCCGCGCA
800

648 5'-GGGGCGTGCGGAAGTCCATCA-3' 668

765 5'-GGGGCGGGCGGGAGTCGAACA-3' 785

Strain	Non-punctual repeat sequences
<i>S. globisporus</i> 1912	5'-GGGGCGGGCGGGAGTCGAACA... ...GGGGCGTGCGGAAGTCCATCA-3'
<i>S. globisporus</i> TFH56	5'-GGGGCGGACGGGAGTCGAACA... ...GGGATGTGCGGAAGTCCATCA-3'
<i>S. griseus</i> NBRC 13350	5'-GGGGCGGACGGGAGTCGAACA... ...GGGACGTGCGGAAGTCCATCA-3'
<i>S. violaceoruber</i> S21	5'-GGGGCGGCCGGGAGTCGAAGA... ...GCGGGGTGCGGAAGTCCATCA-3'
<i>S. fulvissimus</i> DSM 40593	5'-GGGGCGGGCGGGAGTCGAAGA... ...GCGGTGTGCGGAAGTCCATCA-3'
<i>S. sp.</i> CFMR-7	5'-GGGGCGGCCGGGAGTCGAAGA... ...GCGGGGTGCGGAAGTCCATCA-3'
<i>S. sp.</i> 063	5'-GGAGAGGACGGGAGTCGAAGA... ...GCGGCGTACGAAGTCCATCA-3'

Table 3: Non-punctual repeats in *crt* genes of *Streptomyces*.

Even very closely related strains *TFH56* and *NBRC 13350* have only one identical site of 6 bp from the 5'-side. The carotenoid producing strain *1912-4Crt* contains deletion of 119 bp in the *crtY* gene including the sequence of 96 bp between two non-punctual repeats, one NPR from 3'-side and 2 neighboring bp. The deleted sequence of 119 bp has homology with the *transposase* gene (722-747 bp) of *S. albus* J1074 [6].

The homology of the different bacterial *crtY* genes begins from the overlapped stop-start codon CATCA (GTAGT in the minus DNA strand) of NPR from 3'-side. Gene *crtY* of the bacteria of the genera *Micromonospora* (20 strains) and *Nocardiopsis* (5 strains) consists of from 637 to 659 bp and has about 70% identity with one of *S. globisporus 1912-2 strain*. It also contains two similar NPRs and finished together with second one. Bacteria and *streptomyces* have the classical monomeric lycopene β -cyclase as opposed to the order Actinomycetales which produces a heterodimeric enzyme encoded by two *crtYc* and *crtYd* genes [12].

Discussion

Identity of the carotenoid genes of the different *Streptomyces* strains was established ranging from 71% to 95%. The most identity (92% - 95%) of *crt* genes was shown for *S. globisporus 1912*, *S. globisporus TFH56* and related *S. griseus NBRC 13350*. Phylogenetic tree of the eight *crtY* genes of the different strains of *Streptomyces* agrees with the data presented in the Tables 1 and 2.

The cluster of *crt* genes is localized in the end of a linear chromosome of the streptomycetes in the neighborhood of the terminal inverted repeat (TIR) region where frequent DNA rearrangement is observed [13]. Direct repeats stimulated excision and generation of single point mutations, small insertions or replacements and deletions of different size in bacterial genes [14].

The frequency of the spontaneous appearance of the direct *crt*+ mutations in *1912-2* strain, activating beta-carotene production, was

relatively high [5]. Loss of the above-mentioned property by *1912-4Crt* strain in its turn was also increased [15]. Two direct and non-punctual repeats of 21 bp were found in the *crtY* gene of *1912-2* strain beginning from the same specific sites 5'-GGGGCG-3'. We suppose that these sites play role in a site-specific recombination and genetic instability. This assumption was confirmed by the spontaneous deletion of 21 bp observed in the red-orange strain *1912-4Crt*.

The representatives of the different *Streptomyces* species contain only one above mentioned specific site in the NPR of *crtY* gene meanwhile the second one has some substitutions of the base pairs preventing a site-specific recombination and deletion appearance (Table 3) [16-18].

Conclusion

One can suppose that NPRs were introduced into *crtY* genes of different bacteria in the distant past and their function is not clear today. One hypothesis suggests the increase of complexity of the carotenoid gene clusters of *Streptomyces* throughout evolution.

Scientists use metabolic engineering in order to increase the biotechnological production of carotenoids in non-carotenogenic microorganisms. The strain *1912-4Crt* presents biotechnological interest as the perspective one for genetic selection of a high producer of beta-carotene and lycopene. On the first stage of selection the mutant strain *1912-4Lcp Hp7*, a derivative of *1912-4Crt*, was received producing 63.0 mg L⁻¹ of lycopene in flask culture.

References

1. Frenklova G, Breshkova D (2009) Carotenoids from *Rhodotorula* and *Phaffia*: Yeasts of biotechnological importance. *J Industr Microbiol Biotechnol* 36: 163-180.
2. Van Wezel G, McKenzie N, Nodwell J (2009) Applying the genetics of secondary metabolism in model Actinomycetes to the discovery of new antibiotics. *Meth Enzymol* 458: 117-141.
3. Takano H, Obitsu S, Beppu T, Ueda K (2005) Lightinduced carotenogenesis in *Streptomyces coelicolor* A3(2): identification of an extracytoplasmic function sigma factor that directs photodependent transcription of the carotenoid biosynthesis gene cluster. *J Bacteriol* 187: 1825-1832.
4. Lee P, Schmidt-Dannert C (2002) Metabolic engineering towards biotechnological production of carotenoids in microorganisms. *Appl Microbiol Biotechnol* 60: 1-11.
5. Matselyukh B, Lutchenko V, Polishchuk L (2003) Synthesis of carotenoids by mutant strains of *Streptomyces globisporus* 1912. *Mikrobiol Zh* 65: 24-30.
6. Myronovskiy M, Tokovenko B, Brötz E, Rückert C, Kalinowski J, et al. (2014) Genome rearrangement of *Streptomyces albus* J1074 lead to the carotenoid gene cluster activation. *Appl. Microbiol. Biotechnol.* 98: 79-86.
7. Matselyukh B, Lavrinchuk V (1999) The isolation and characteristics of mutant *Streptomyces globisporus* 1912 defective for landomycin E biosynthesis. *Mikrobiol Zh* 61: 22-27.
8. Kieser T, Bibb M, Buttner M, Chater K, Hopwood D (2000) Practical streptomycetes genetics. The John Innes Foundation, Norwich.
9. Chikhi R, Medvedev P (2014) Informed and automated k-mer size selection for genome assembly. *Bioinformatics* 30: 31-37.
10. Boetzer M, Henkel C, Janssen H, Butler D, Pirovano W (2011) Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27: 578-579.
11. Boetzer M, Pirovano W (2012) Toward almost closed genomes with GapFiller. *Genome Biol* 13: R56.

12. Krubasik P, Sandman G (2000) Molecular evolution of lycopene cyclases involved in the formation of carotenoids with ionone end groups. *Biochem Soc Transact* 28: 806-810.
13. Fisher G, Wenner T, Decaris B, Leblond P (1998) Chromosomal arm replacement generates a high level of intraspecific polymorphism in the terminal inverted repeats of the linear chromosomal DNA of *Streptomyces ambifaciens*. *Proc Natl Acad Sci* 95: 14296-14301.
14. Naswall J (2017) Direct and inverted repeat stimulated excision (DIRex): Simple, single-step, and scar-free mutagenesis of bacterial genes. *PLoS One* 12: e0184126.
15. Polishchuk LV, Golembiovskaya SL, Matselyukh BP, Lukyanchuk VV (2013) Genetic variability of synthesis feature of carotenoids in *Streptomyces globisporus* 1912. *Mikrobiol Zh* 75: 40-46.
16. Lee HS, Ohnishi Y, Horinouchi S (2001) A sigma B-like factor responsible for arotenoid biosynthesis in *Streptomyces griseus*. *J Mol Microbiol Biotechnol* 33: 95-101.
17. Yang J, Guo L (2014) Biosynthesis of β -carotene in engineered *E. coli* using the MEP and MVA pathways. *Microb Cell Factor* 13:160.
18. Golembiovskaya S, Dvornyk T, Yanchenko L, Matselyukh B (2014) The productivity of mutant strain *Streptomyces globisporus* 1912-4Lcp Hp7 in the different conditions. *Microbiol Biotechnol* 2: 6-14.