

<https://doi.org/10.15407/frg2025.03.268>

УДК 004.2:577.1:577.11:579.873.71

СТРЕПТОМІЦЕТИ, В ГЕНОМАХ ЯКИХ НАЯВНІ КЛАСТЕРИ ГЕНІВ БІОСИНТЕЗУ ПРОТИРАКОВОГО АНТИБІОТИКА ЛІДАМІЦИНУ

Л.В. ПОЛІЩУК

Інститут мікробіології і вірусології ім. Д.К. Заболотного Національної
академії наук України

03143 Київ, вул. Академіка Заболотного, 154

e-mail: LVPolishchuk@ukr.net

Лідаміцин — протипухлинний та антибактеріальний антибіотик, що синтезується мікроорганізмами виду *Streptomyces globisporus* (штами С-1027 і TFH56). Метою роботи було виявити стрептоміцети, геноми яких містять імовірні кластери біосинтетичних генів лідаміцину (LDM-кластер). Об'єктами дослідження були нуклеотидні послідовності стрептоміцетів, депоновані в Інтернет базах даних (Nucleotide Collection, Whole-Genome Shotgun Contigs) на сервері National Center for Biotechnology Information. Первинні структури ДНК аналізували за допомогою Basic Local Alignment Search Tool. Послідовність LDM-кластера *S. globisporus* С-1027 використовували як запит для комп'ютерного аналізу нуклеотидних послідовностей. Інтернет бази даних містять інформацію про первинні структури сотень тисяч хромосом і плазмід стрептоміцетів, які повністю (Nucleotide Collection) або частково (Whole-Genome Shotgun Contigs) визначені. Загалом в обох базах даних виявлено 78 стрептоміцетів, геноми яких містили повні LDM-кластери. Більшість стрептоміцетів у вибірці (65,4 %) належить до виду *S. globisporus*. Як відомо, LDM-кластери містяться на плазмідах SGLP1 (штам С-1027) та рTFHSG1 (штам TFH56). У базі даних Nucleotide Collection було знайдено 6 плазмід (complete genome). Послідовності виявлених плазмід та SGLP1 не є ідентичними — лише фрагменти з молекулярним розміром 104,5 тпн у їх послідовностях є подібними. Ідентифіковані плазмід були за формою молекули як лінійними (CP108547, CP109060, рTFHSG1, CP109147, CP108593), так і циклічною (рG-2-2). За молекулярним розміром усі 6 плазмід можна вважати великими плазмідами. Плазмід CP108593 та рG-2-2 відрізняються своєю послідовністю як від структур інших плазмід, так і від структур виявлених контигів стрептоміцетів. Крім того, у Whole-Genome Shotgun Contigs базі даних виявлено 72 стрептоміцети, геноми яких містили повні LDM-кластери. Однак, повні послідовності плазмід були виявлені тільки в послідовності кількох контигів. Повну послідовність плазмід рTFHSG1 виявили в контигах *S. badius* SF7B6 (LWMP01000002, 134520 пн), *S. badius* SP6C4 (LWMQ01000002, 140513 пн), *S. griseus* S4-7 (JYBE02000002, 141125 пн) штамів. Встановлено наявність повноцінних послідовностей LDM-кластерів у геномах 78 стрептоміцетів, представлених у базах даних на сервері NCBI. Було виявлено 6 плазмід, що містять LDM-кластери у своїх структурах.

Цитування: Поліщук Л.В. Стрептоміцети, в геномах яких наявні кластери генів біосинтезу протиракового антибіотика лідаміцину. *Фізіологія рослин і генетика*. 2025. 57, № 3. С. 268—276. <https://doi.org/10.15407/frg2025.03.268>

Ключові слова: *Streptomyces*, лідаміцин, LDM-кластер, плазміда, послідовність нуклеотидів, комп'ютерний аналіз, ідентичність, покриття запиту.

Встановлено, що антибіотик лідаміцин (інша назва C-1027) проявляє значну цитотоксичну дію за різних типів раку, включно рак печінки, молочної залози, підшлункової залози, товстої кишки, легенів, шлунка та головного мозку, а також при лімфомі та мієломі [1–3]. Також показано, що антибіотик має антимікробну активність проти більшості грампозитивних бактерій, але не проти *Mycobacterium sp.* або грамнегативних бактерій [2].

Лідаміцин (ЛДМ) — хромопротеїновий антибіотик, що продукується стрептоміцетами штамів C-1027 і TFH56 виду *Streptomyces globisporus* [1–8]. Встановлено, що ЛДМ-кластери в обох штамів локалізуються на плазмідах [6–8]. LDM-кластер (72676 пн) штаму *S. globisporus* C-1027 розташований на лінійній плазміді SGLP1 (167754 пн) [6–8]. LDM-кластер штаму *S. globisporus* C-1027 був ідентифікований, секвенований і гетерологічно експресований [7]. LDM-кластер штаму *S. globisporus* TFH56 також є плазмідним кластером (pTFSG1, 127820 пн) [4]. Кластер біосинтезу лідаміцину *S. globisporus* C-1027 включає 58 генів (WQO_33745 — WQO_34030). LDM-кластер штаму *S. globisporus* TFH56 містить гени DIJ_33905—DIJ_34195.

Встановлено, що стрептоміцети здатні продукувати значно більшу кількість і різноманітність біологічно активних метаболітів, ніж виявляється при лабораторному культивуванні чи при існуванні в умовах довкілля. Вважається, що це значною мірою пов'язано з тим, що продукти генів не потрібні в даних умовах або пошкоджені регуляторні механізми їх активації. Це так звані мовчазні гени чи кластери генів [8, 9].

Розвиток наукових технологій у молекулярній біології та біоінформатиці дає змогу виявити наявність таких кластерів у геномах стрептоміцетів [8, 9].

Мета нашої роботи полягала в ідентифікуванні за допомогою біоінформаційних методів стрептоміцетів, геноми яких містять імовірні LDM-кластери.

Методика

Об'єктами дослідження були нуклеотидні послідовності хромосом і плазмід стрептоміцетів, депоновані в загальнодоступні Інтернет бази даних Nucleotide Collection (nr/nt) та Whole-Genome Shotgun Contigs (wgs) National Center for Biotechnology Information (NCBI). Первинні структури ДНК аналізували за допомогою Basic Local Alignment Search Tool (BLASTN: megablast). Послідовність LDM-кластера *S. globisporus* C-1027 використовувалася як запит (Query) для комп'ютеризованого аналізу нуклеотидних послідовностей. LDM-кластер — фрагмент послідовності (20150—92825 пн) плазміді SGLP1 (CP013739).

Визначали генетичну спорідненість стрептоміцетів-носіїв плазмід за значенням ANI (average nucleotide identity). Показник ANI визначали за допомогою програми CJ Bioscience's online ANI calculator (<https://www.ezbiocloud.net/tools/ani>).

Послідовність хромосоми *S. globisporus* C-1027 використали як запит (Query).

Результати та обговорення

Інтернет бази даних (nr/nt та wgs, NCBI) містять інформацію про первинні структури сотень тисяч хромосом і плазмід стрептоміцетів, які повністю або частково визначені. BLASTN-аналіз первинних структур ДНК у цих базах даних виявив послідовності, подібні LDM-кластеру *S. globisporus* C-1027 (Query) в тисячах геномів стрептоміцетів. Однак здебільшого це були невеликі фрагменти як плазмідної, так і хромосомної ДНК. Отже, показано досить значне поширення повноцінних LDM-кластерів у геномах стрептоміцетів (78 стрептоміцетів), що депоновані в базах даних nr/nt та wgs (NCBI) (рис. 1).

Біоінформативний аналіз послідовностей стрептоміцетів у базі даних Nucleotide Collection. BLASTN-аналіз первинних структур ДНК (complete genome) у базі даних nr/nt виявив послідовності, подібні Query у хромосомах і плазмідах сотень стрептоміцетів. На рис. 2 показана умовна поширеність і локалізація фрагментів послідовностей стрептоміцетів, аналогічних LDM -кластеру.

Комп'ютерний аналіз послідовностей ДНК стрептоміцетів у цій базі даних виявив лише шість генетичних елементів, структури яких

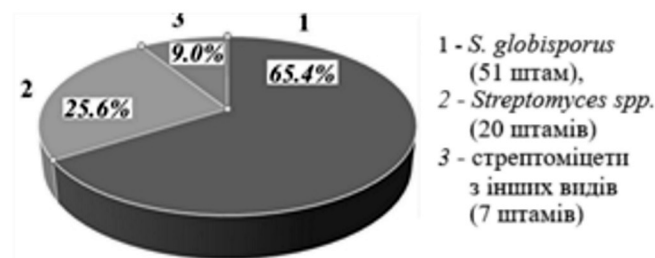


Рис. 1. Видовий склад сукупності виділених стрептоміцетів, які містять імовірні LDM-кластери (бази даних nr/nt та wgs, NCBI)

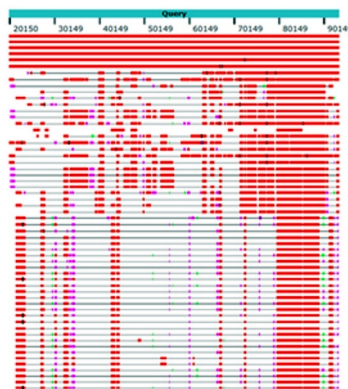


Рис. 2. Розподіл послідовностей, подібних до LDM-кластерної послідовності *S. globisporus* C-1027 (Query) у геномах стрептоміцетів у базі даних nr/nt (NCBI)

містили послідовності, ідентичні повній послідовності LDM-клас-
тера. Це були послідовності плазмід (табл. 1).

Нуклеотидні послідовності 6 плазмід і SGLP1 були частково ідентичні — всі вони містили послідовності, подібні до фрагмента 7747—112237 пн плазмиди SGLP1 (*S. globisporus* C-1027). Порівняння послідовностей 7 плазмід між собою показало, що плазмиди можна розділити на 2 групи: 1) CP109147, pTFHSG1, CP109060, CP108547, SGLP1 і 2) CP108593, pG-2-2. Рівень ідентичності послідовностей 5 плазмід (1 група) перевищує 98,5 %, тоді як ідентичність плазмід 2 групи становить менше 93 % порівняно з послідовностями інших плазмід.

Цікавим було визначити рівень подібності первинних структур хромосом штамів-носіїв плазмід (табл. 2).

ТАБЛИЦЯ 1. Інформація про виявлені плазмиди (послідовність LDM-клас-
тера *S. globisporus* C-1027 використали як запит (Query))

Плазмиди	Молекулярний розмір, пн	Посилання, GenBank	Штами-носії плазмід	Показники подібності LDM-клас-терів
Штами виду <i>S. globisporus</i>				
Unnamed1	216726	CP108547	NBC_01211	Qc = 100 %, I = 98,56 %
Unnamed1	294500	CP109060	NBC_01004	Qc = 100 %, I = 98,6 %
pTFHSG1	127820	CP029362	TFH56	Qc = 100 %, I = 98,56 %
Unnamed1	231877	CP109147	NBC_01745	Qc = 100 %, I = 98,54 %
Штам виду <i>Streptomyces</i> sp.				
Unnamed1	104367	CP108593	NBC_01178	Qc = 100 %, I = 92,96 %
Штам виду <i>S. parvulus</i>				
pG-2-2	214348	CP135078	F-G-2	Qc = 100 %, I = 93,56 %

Примітка: Qc (query coverage) — покриття, I (identity) — ідентичність.

ТАБЛИЦЯ 2. Подібність нуклеотидних послідовностей хромосом штамів-носіїв плазмід (послідовність хромосоми *S. globisporus* C-1027 використали як запит (Query))

Штам <i>Streptomyces</i>	Посилання в GenBank	Молекулярний розмір хромосоми, пн	ГЦ вміст, %	Показник ANI, %
Штами виду <i>S. globisporus</i>				
C-1027	CP013738	7608611	71,50	
NBC_01745	CP109149	7857971	71,49	95,97
NBC_01004	CP109059	7662562	71,53	95,98
NBC_01211	CP108546	7585586	71,61	95,96
TFH56	CP029361	7488586	71,6	96,05
Штам виду <i>Streptomyces</i> sp.				
NBC_01178	CP108592	6801306	72,8	89,48
Штам виду <i>S. parvulus</i>				
F-G-2	CP135079	7143830	71,77	94,36

Примітка: ANI (average nucleotide identity) — середня ідентичність видів.

Біоінформатичний аналіз секвенсису стрептоміцетів у базі даних Whole-Genome Shotgun Contigs. У базі даних wgs було виявлено 72 штами стрептоміцетів, ДНК яких містять послідовності, подібні послідовності запиту (Query). Рівні подібності послідовностей цих 72 імовірних LDM-кластерів до послідовності запиту були значущими — $Qc > 93 \%$, $I > 92,5 \%$. Більшість штамів (46 штамів) з бази даних wgs (NCBI) належали до виду *S. globisporus*.

Штами видів *S. griseus* (2 штами) та *S. mediolani* (1 штама) також належали до групи *S. griseus*, як і штами виду *S. globisporus*. У ДНК штамів *S. badius*, *S. bauhiniae*, *S. capillispiralis* виявлені LDM-кластери, які не мають відношення до групи *S. griseus*. У переважній більшості структур імовірних LDM-кластерів (63 штами) були локалізовані на одному контигу, послідовності з 9 імовірних кластерів локалізовані на 2 контигах.

Як відомо, в наборі контигів (база даних wgs (NCBI) послідовностей генома стрептоміцету одночасно присутні послідовності всіх генетичних чинників клітини стрептоміцету: і хромосоми, і плазмід.

Біоінформатичний пошук повних послідовностей плазмід у послідовностях контигів стрептоміцетів (база даних Whole-Genome Shotgun Contigs). Послідовності з 7 плазмід, що містять LDM-кластери, були використані як запити в аналізі бази даних wgs (NCBI). Виявлено, що в послідовностях контигів низки стрептоміцетів наявні послідовності, повністю схожі зі структурами деяких плазмід. Це переважно були послідовності плазмиди pTFHSG1. Так послідовності, подібні послідовності всієї плазмиди pTFHSG1, містили структури контигів штамів *S. badius* SF7B6 (LWMP01000002, 134520 пн), *S. badius* SP6C4 (LWMQ01000002, 140513 пн), *S. griseus* S4-7 (JYBE02000002, 141125 пн). Однак такі послідовності становили лише фрагмент структур контигів (рис. 3). BLASTN-аналіз структур цих контигів показав, що послідовності LWMP01000002, LWMQ01000002, JYBE02000002 майже ідентичні ($Qc > 97 \%$, $I = 99,99 \%$).

Послідовність, подібна до фрагменту структури плазмиди CP135078 (214348 пн), виявлена на контигу штамів *Streptomyces sp.* PBSH9 (JAKZO01000002, 207599 пн). Проте послідовність контигу є меншою за молекулярним розміром, ніж послідовність плазмиди CP135078, і становить лише 91 % послідовності плазмиди.

Первинні структури геномної ДНК штамів *S. badius* SF7B6, *S. badius* SP6C4, *S. griseus* S4-7, *Streptomyces sp.* PBSH9 натеper частково секвеновані й внесені в базу даних wgs (NCBI) як набори контигів. Після секвенування геномів штамів і побудови їх генетичних карт буде точно визначено, чи є згадані вище контиги послідовностями плазмід.

База даних GenBank містить нуклеотидні послідовності 770 стрептоміцетів як повністю визначених геномів (197 штамів), так і



Рис. 3. Візуалізація локалізації послідовностей контигів, подібних до послідовності pTFHSG1 плазмиди (complete genome)

наборів контигів (573 штами). Всього в бази даних GenBank внесені послідовності 81 стрептоміцету з виду *S. glodisporus*, зокрема геноми шести стрептоміцетів, які повністю визначені. Імовірні LDM-кластери були виявлені в стрептоміцетів з 51 штаму цього виду.

У геномі штаму *S. globisporus* NBC_01078 (CP108664) не виявлено послідовності LDM-кластера, незважаючи на те, що клітини штаму містять 4 плазмиди: гігантську (CP108665), велику (CP108666) та 2 малі (CP108667, CP108668). Імовірні LDM-кластери були виявлені в наборах контигів геномної ДНК низки добре вивчених штамів видів *S. globisporus*, наприклад, штам JCM4378 (type strain) та NRRL B-2709. Раніше було показано, що існує значна ідентичність геномної ДНК *S. globisporus* TFH56 (Qc = 89 %, I = 97 %), *S. globisporus* NRRL B-2709 (Qc = 81 %, I = 95 %) та *S. globisporus* 1912 4Crt (Qc = 82 %, I = 96 %) з послідовністю хромосом *S. globisporus* C-1027 штамів [12].

Цікаво, що раніше була показана значна різниця (Qc = 38 %, I = 95 %) у структурі штаму ДНК *S. globisporus* NRRL B-2293 з послідовністю хромосом штаму *S. globisporus* C-1027 [12]. Але LDM-кластери не були ідентифіковані в послідовності ДНК штамів *S. globisporus* 1912 4Crt і *S. globisporus* NRRL B-2293, які генетично споріднені між собою.

Згадані вище стрептоміцети були виділені зі зразків ґрунту (здебільшого); з поверхні рослин (*S. globisporus* TFH56, *S. globisporus* NBC_01211). Однак штам *S. globisporus* 4-3 був виділений з тіла мурахи. Стрептоміцети видів *S. globisporus* були виділені зі зразків, взятих із різних країн: Вірменії, Данії, Південної Кореї, Китаю, США, Пакистану.

Кластери генів для біосинтезу переважної більшості антибіотиків зазвичай локалізовані в хромосомах стрептоміцетів [10]. Проте в деяких випадках такі скупчення розташовуються на плазмідах. Переважно на так званих великих плазмідах. Наприклад, молекулярний розмір плазмиди pKSL становить 520 тпн [11]. Деякі з плазмідних кластерів біосинтезу антибіотиків детально вивчені: SCP1 та pSV1, що детермінують синтез метиленоміцину, pSLA2-L — ланкацидину й ланкаміцину, pPZG103 — окситетрацикліну, SGLP1 та pTFH56 — лідаміцину, pKSL — саліноміцину [4, 5, 9, 12, 13]. Більшість цих плазмід є лінійними, за винятком плазмиди pSV1 [7, 8].

Відомо, що штам *S. globisporus* C-1027 та *S. globisporus* TFH56 продукують протипухлинний агент лідаміцин [4—8]. ЛДМ-кластери в обох штамів локалізуються на плазмідах [4—8].

Серед плазмід, відібраних у наших дослідженнях, представлена плазміда pTFH56 (див. табл. 1), яка, як відомо, містить LDM-кластер [4, 5]. Це доводить вірогідність результатів, отриманих застосуванням методом. Як зазначалося вище, всі 6 ідентифікованих плазмід (включно й pTFH56), що містять LDM-кластери і SGLP1, є великими плазмідами (див. табл. 1).

Відомо, що синтез антибіотика метиленоміцину визначають за допомогою mmy/mmh-кластерів, локалізованих на плазмідах SCP1 (*S. colicolor* A3(2)), pSV1 (*S. violaceus ruber*). Встановлено, що плазміда SCP1 є лінійною, а плазміда pSV1 має циклічну форму. Серед відіб-

раних 6 плазмід є плазміди обох форм. Було ідентифіковано 5 лінійних плазмід (як SGLP1) та 1 циклічну плазмиду (pG-2-2). Рівень ідентичності послідовностей 5 плазмід (CP109147, pTFHSG1, CP109060, CP108547 та CP013739) перевищує 98,5 %, тоді як ідентичність CP108593 та pG-2-2 плазмід становить менше 93 % порівняно з послідовностями всіх 7 плазмід.

Визначення подібності послідовностей хромосомних ДНК стрептоміцетів-носіїв плазмід показало, що штами *Streptomyces sp.* NBC_01178 та *S. parvulus* F-G-2 генетично неспоріднені зі штамми виду *S. globisporus* — це демонструє розмір показника спорідненості ANI. Порогове значення для розмежування видів становить 95 % [13]. Цікаво, що плазміди стрептоміцетів *Streptomyces sp.* NBC_01178 і *S. parvulus* F-G-2 складають групу 2.

У базі даних Whole-Genome Shotgun Contigs виявлено 72 штами стрептоміцетів, ДНК яких містять послідовності подібні LDM-кластеру *S. globisporus* C-1027 ($Q_c > 93 \%$, $I > 92,5 \%$). Більшість штамів (49 штамів), виявлених в базі даних Whole-Genome Shotgun Contigs, належали до групи *S. griseus*. Значущими були рівні подібності послідовностей їх вірогідних LDM-кластерів до послідовності запитів — $Q_c > 93 \%$, $I > 92,5 \%$.

Як відомо, послідовності геномів більшості стрептоміцетів визначені фрагментарно й депонуються в даних Whole-Genome Shotgun Contigs у вигляді наборів контигів. У наборі контигів штаму одночасно представлені послідовності як хромосом, так і плазмід. Цікавим було визначити імовірні плазмідні послідовності серед відібраних контигів, що містять LDM-кластери.

Послідовності з 7 плазмід, що містять LDM-кластери, були використані як запити в аналізі бази даних Whole-Genome Shotgun Contigs. Послідовності, аналогічні всій послідовності плазміди pTFHSG1 (127820 пн), наявні на контигах штамів *S. badius* SF7B6 (LWMP01000002, 134520 пн), *S. badius* SP6C4 (LWMQ01000002, 140513 пн), *S. griseus* S4-7 (JYBE02000002, 141125 пн). Однак такі послідовності становили лише фрагменти структур контигів (див. рис. 3). Було показано, що структури цих контигів (LWMP01000002, LWMQ01000002, JYBE02000002) є майже ідентичними між собою послідовностями ($Q_c > 97 \%$, $I = 99,99 \%$).

Після повного секвенування послідовностей геномів штамів *S. badius* SF7B6, *S. badius* SP6C4, *S. griseus* S4-7, *Streptomyces sp.* PBSH9 і побудови генетичних карт буде точно визначено, чи є зазначені вище контиги плазмідними послідовностями.

Таким чином, встановлено наявність повноцінних послідовностей LDM-кластерів у геномах 78 стрептоміцетів, представлених у базах даних (nr/nt, wgs) на сервері NCBI. Штами з вибраної сукупності стрептоміцетів здебільшого (65,4 %) належать до виду *S. globisporus*. У базі даних nr/nt було знайдено 6 плазмід (CP108547, CP109060, pTFHSG1, CP109147, CP108593, pG-2-2). Ідентифіковані плазміди та SGLP1 не ідентичні — вони містять у своїх послідовностях лише фрагменти з 104,5 тпн, які мають подібну структуру. Виявлені плазміди можуть бути як лінійними (CP108547, CP109060, pTFHSG1,

CP109147, CP108593), так і циклічними (pG-2-2) за формами молекул. За молекулярними розмірами плазмідних ДНК, всі 6 плазмід можна віднести до великих плазмід. Плазмиди CP108593 та pG-2-2 вирізняються відмінністю своїх послідовностей як від плазмідних структур, так і від структур згаданих вище контигів. Повну послідовність pTFHSG1 плазмиди (127820 пн) виявили в межах контигів штамів *S. badius* SF7B6 (LWMP01000002, 134520 пн), *S. badius* SP6C4 (LWMQ01000002, 140513 пн), *S. griseus* S4-7 (JYBE02000002, 141125 пн).

REFERENCES

1. Hu, J.L., Xue, Y.C., Xie, M.Y., Zhang, R., Otani, T., Minami, Y., Yamada, Y. & Marunaka, T. (1988). A new macromolecular antitumor antibiotic, C-1027. I. Discovery, taxonomy of producing organism, fermentation and biological activity. The Journal of antibiotics, 41, No. 11, pp. 1575-1579. <https://doi.org/10.7164/antibiotics.41.1575>
2. Shi, Y.K., Wu, S.Y., Huang, Y.H. & Zhen, Y.S. (2006). Chemosensitivity of *mdr1* gene overexpressed multidrug resistant cancer cells to lidamycin. Acta Pharmac. Sinica, 41, No. 12, pp. 1146-1151.
3. Chen, Y., Sun, W., He, R., Zhang, F., Wang, H., Li, P., Shao, R.G. & Xu, X. (2017). Lidamycin decreases CD133 expression in hepatocellular carcinoma via the Notch signaling pathway. Oncol. let., 14, No. 6, pp. 7889-7895. <https://doi.org/10.3892/ol.2017.7248>
4. Cho, G. & Kwak, Y.S. (2019). Evolution of Antibiotic Synthesis Gene in Clusters the *Streptomyces globisporus* TFH56, Isolated from Tomato Flower. G3 Genes[Genomes]Genetics, 9, No. 6, pp. 1807-1813. <https://doi.org/10.1534/g3.119.400037>
5. Aguilar, A. & Hopwood, D.A. (1982). Determination of methylenomycin A synthesis by the pSV1 plasmid from *Streptomyces violaceus-ruber* SANK95570. J. General Microbiol., 128, No. 8, pp. 1893-1901. <https://doi.org/10.1099/00221287-128-8-1893>
6. Li, X., Lei, X., Zhang, C., Jiang, Z., Shi, Y., Wang, S., Wang, L. & Hong, B. (2016). Complete genome sequence of *Streptomyces globisporus* C-1027, the producer of an enediyne antibiotic lidamycin. J. Biotechnol., 222, pp. 9-10. <https://doi.org/10.1016/j.jbiotec.2016.02.004>
7. Liu, W., Christenson, S.D., Standage, S. & Shen, B. (2002). Biosynthesis of the Eneidyne Antitumor Antibiotic C-1027. Science, 297, No. 5584, pp. 1170-1173. <https://doi.org/10.1126/science.1072110>
8. Li, M., Liu, W. & Li, Y. (2003). Cloning, expression and characterization of gene *sgcD* involved in the biosynthesis of novel antitumor lidamycin. Science in China Series C-Life Sciences. 46, pp. 310-319. <https://doi.org/10.1360/03yc9033>
9. Gram, L. (2015). Silent clusters — speak up! Microbial Biotechnol., 8, No. 1, pp. 13-14. <https://doi.org/10.1111/1751-7915.12181>
10. Mao, D., Okada, B.K., Wu, Y., Xu, F. & Seyedsayamdost, M.R. (2018). Recent advances in activating silent biosynthetic gene clusters in bacteria. Current Opinion Microbiol., 45, pp. 156-163. <https://doi.org/10.1016/j.mib.2018.05.001>
11. Kinashi, H. & Shimaji, M. (1987). Detection of giant linear plasmids in antibiotic producing strains of *Streptomyces* by the OFAGE technique. J. Antibiotics, 40, No. 6, pp. 913-916
12. Polishchuk, L.V. (2020). Similarity of genomic sequences of five *Streptomyces globisporus* strains. Mikrobiol. Zhurn., 82, No. 1, pp. 45-50. <https://doi.org/10.15407/microbiolj82.01.043>
13. Yoon, S.H., Ha, S.M., Lim, J., Kwon, S. & Chun, J.A. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie van Leeuwenhoek, 110, No. 10, pp. 1281-1286. <https://doi.org/10.1007/s10482-017-0844-4>

Received 30.04.2025

STREPTOMYCETES, IN THE GENOMES OF WHICH THERE ARE CLUSTERS OF GENES FOR THE BIOSYNTHESIS OF THE ANTI-CANCER ANTIBIOTIC LIDAMYCIN

L.V. Polishchuk

D.K. Zabolotny Institute of Microbiology and Virology, National Academy of Sciences of Ukraine
154 Akademika Zabolotnogo St., Kyiv, 03143, Ukraine
e-mail: LVPolishchuk@ukr.net

Lidamycin is an antitumor and antibacterial antibiotic synthesized by microorganisms of the species *Streptomyces globisporus* (strains C-1027 and TFH56). The aim of the work was to identify streptomycetes whose genomes contain probable clusters of lidamycin biosynthetic genes (LDM-cluster). The objects of the study were the nucleotide sequences of streptomycetes deposited in online databases (Nucleotide Collection, Whole-Genome Shotgun Contigs) on the server of the National Center for Biotechnology Information. The primary structures of the DNA were analyzed using the Basic Local Alignment Search Tool. The sequence of the LDM-cluster from *S. globisporus* C-1027 was used as a query for the computational analysis of nucleotide sequences. The online databases contain information about the primary structures of hundreds of thousands of chromosomes and plasmids of streptomycetes that are complete (Nucleotide Collection) or partially (Whole-Genome Shotgun Contigs) defined. In total, 78 streptomycetes were identified in both databases whose genomes contained complete LDM-clusters. The majority of streptomycetes in the set (65.4 %) belong to the species *S. globisporus*. As is known, LDM clusters are located on the plasmids SGLP1 (strain C-1027) and pTFHSG1 (strain TFH56). Six plasmids (complete genome) were found in the Nucleotide Collection database. The sequences of the identified plasmids and SGLP1 are not identical — only fragments with a molecular size of 104.5 kbp in their sequences are similar. The identified plasmids were found in the shape of molecules as both linear (CP108547, CP109060, pTFHSG1, CP109147, CP108593) and circular (pG-2-2). By molecular size, all 6 plasmids can be classified as large plasmids. Plasmids CP108593 and pG-2-2 differ in their sequence both from the structures of other plasmids and from the structures of the identified streptomycet contigs. Additionally, 72 streptomycetes whose genomes contained complete LDM clusters were discovered in the Whole-Genome Shotgun Contigs database. However, complete plasmid sequences were identified only in the sequences of several contigs. The complete sequence of plasmid pTFHSG1 was found in the contigs of *S. badius* SF7B6 (LWMP01000002, 134520 bp), *S. badius* SP6C4 (LWMQ01000002, 140513 bp), *S. griseus* S4-7 (JYBE02000002, 141125 bp) strains. The presence of complete LDM cluster sequences in the genomes of 78 streptomycetes was established, as represented in databases on the NCBI server. Six plasmids containing LDM clusters in their structures were identified.

Key words: *Streptomyces*, lidamycin, LDM-cluster, plasmid, nucleotide sequence, computerized analysis, identity, query coverage.

ORCID

Л.В. ПОЛІЩУК — L.V. Polishchuk <https://orcid.org/0000-0002-3159-5022>