S rep om ce .

Min Xu,[a] Yemin Wang,[a] Zhilong Zhao,[b] Guixi Gao,[a] Sheng-Xiong Huang,[c] Qianjin Kang,[a] Xinyi He,[a] Shuangjun Lin,[a] Xiuhua Pang,[b] Zixin Deng,[a] Meifeng Tao[a]

State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China[a]; State Key Laboratory of Microbial Technology, School of Life Sciences, Shandong University, Jinan, China[b]; Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China[c]

ABSTRACT

Genome sequencing projects in the last decade revealed numerous cryptic biosynthetic pathways for unknown secondary metabolites in microbes, revitalizing drug discovery from microbial metabolites by approaches called genome mining. In this work, we developed a heterologous expression and functional screening approach for genome mining from genomic bacterial artificial chromosome (BAC) libraries in *Streptomyces* spp. We demonstrate mining from a strain of *Streptomyces rochei*, which is known to produce streptothricins and borrelidin, by expressing its BAC library in the surrogate host *Streptomyces lividans* SBT5, and screening for antimicrobial activity. In addition to the successful capture of the streptothricin and borrelidin biosynthetic gene clusters, we discovered two novel linear lipopeptides and their corresponding biosynthetic gene cluster, as well as a novel cryptic gene cluster for an unknown antibiotic from *S. rochei*. This high-throughput functional genome mining approach can be easily applied to other streptomycetes, and it is very suitable for the large-scale screening of genomic BAC libraries for bioactive natural products and the corresponding biosynthetic pathways.

IMPORTANCE

Microbial genomes encode numerous cryptic biosynthetic gene clusters for unknown small metabolites with potential biological activities. Several genome mining approaches have been developed to activate and bring these cryptic metabolites to biological tests for future drug discovery. Previous sequence-guided procedures relied on bioinformatic analysis to predict potentially interesting biosynthetic gene clusters. In this study, we describe an efficient approach based on heterologous expression and functional screening of a whole-genome library for the mining of bioactive metabolites from *Streptomyces*. The usefulness of this function-driven approach was demonstrated by the capture of four large biosynthetic gene clusters for metabolites of various chemical types, including streptothricins, borrelidin, two novel lipopeptides, and one unknown antibiotic from *Streptomyces rochei* Sal35. The transfer, expression, and screening of the library were all performed in a high-throughput way, so that this approach is scalable and adaptable to industrial automation for next-generation antibiotic discovery.

M

(1).

3 ,

( ) (-

),

METHODS AND MATERIALS

Bacterial strains, plasmids, and culture conditions.

Metabolic analysis of *S. rochei* Sal35 and BAC-containing *Streptomyces* strains.

Construction of the genomic BAC library of *S. rochei* Sal35.

1 , ,

, ,

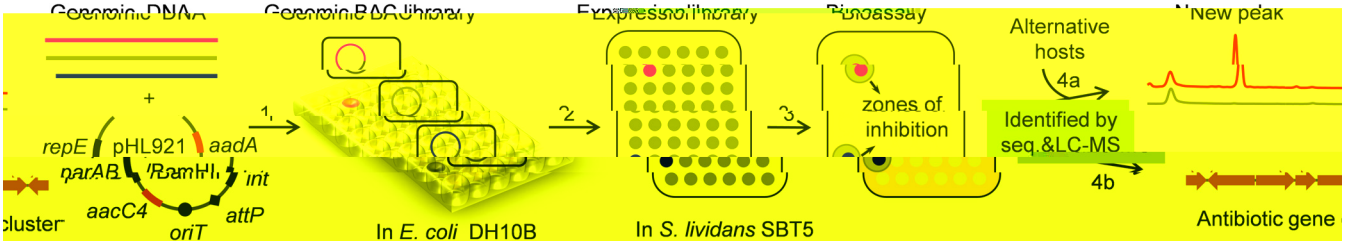
E *l*		
10		
12567	, , , *M* , ,	50
35	-	
l , ,		
5	Δ Δ *KL* Δ , 3 *l* *L*( )	20
1	Δ Δ *KL* Δ , 3 *l* *l* *L*( )	26
l *l*		
1152	145 Δ Δ Δ Δ *B*(12)	1
1154	145 Δ Δ Δ Δ *B*(12) *L*(262)	1
11	145 Δ Δ *L*	51
l 65	-	
l 10201		
B *ll*		
M ^{2}155		52
		,
21	*K2* - Φ 31 (3)-*I* , *A* *A* *B* *C* *E.* *l* -	
307	*K2*- -	23

1	-	1
2	-	2
6	-	2
2 6	-	3
3 5	-	3
5	-	3
10	-	3
1	-	3
2 3	4	
6 1	4	
6 11	4	
11	4	

, .

(0.04 - , 1). 21 (. 1 (,30 , 5 6) 4 2×
) , 13.7- - . 10 / .
, , , 10 50 ,
. - , (-
0.025-μ - 10%) 200 μ / . - -
( ) 800 2 4 *E.* *l* 10 - 30 . *E.* *l* -
. 1 37 4 - . . *l* , ,
50 μ / 5
. 60

LEXAS screening of the *S. rochei* Sal35 BAC genomic library. 6- . *E.* *l* -
E. *l* , *K2*- 307 12567
- , 10 /
E. *l* 12567/ 307 ( ), 10 / (- *K2*- . *l* , 5. -
), . *l* , 5 ( ). (12 16 30 ,
) 6- (130 μ) 50 μ /
4 6 600 ($_{600}$) 0.4 0.6. -
12567/ 307 (100) 50 (50 μ /) *E.* *l*
μ / $_{600}$ 0.4 0.6, 30 4 6 . , -
, 15 . , 20 , 4
μ 12567/ 307 10 / (50 μ /) (25 μ /) -
, / 1 1, *E.* *l* . *l* , ,
(- - , 5.
,) 200 5 . , . *l* , 5 -
5 - , , 3,



1 . (1) . 100
21. (2) , 21- E. l 10
.l 5. .(3) 5
. .(4) -

6- 30 .
24 4 60
. (/ - , 3- - - , / -)
.

Identification of streptothricins and borrelidin. -
-
.

DNA sequence.
/454 ( ) , -
(// . /) (32).
16 -
. 35 1 ( ) 1 ( ).

Isolation and purification of linear lipopeptides 8D1-1 and 8D1-2.
. l l 1152/ 1 30
3 6 30 . -
-
- - 20 5 .
40 (-
20), 3 ,
20 2 -
1-1 1-2 ,
- , 2 -
1 50% 100% . 60% 70% 1-1
1-2 , -
(1 5 , .4 250),
2 (0.1%) ( ). 15-
4 % 2 / . 1-1 1-2
5.7 , . ,20 1-1 15
1-2 . 1-1 1-2
- , (/
), - (1) 2
( ) .

Marfey assay for chiral analysis of amino acids.
4- 6 (6) 3- - ,
.0.1 1-1 2 6
120 24 .
50 210 , 20 - -
(2), 1.5- , 40
1% *N*-(5- -2, 4-)- -
(,) 1 3
.
40 1
. 4
2 200 - (33).
3, 0.1 1-1 -
2- - 1 2 (5% -
) 166 0.5 (34).

LC-MS analysis for FDAA-derivatized amino acids. - -
(2) 6530 -
- (-)
1260 1 -
- (1 3.5 , 2.1 150).
0 20 10 % , 20 25
100% , 26 35 10% , 2 (0.1%
) ( ) 0.2
/ .
, / 50 1,700. 457.1 ± 0.2 ,
420.1 ± 0.2 , 402.1 ± 0.2 - ,
- , -3- - , ,
.

Accession number(s).
.
440 52 (16), 3602 (21), 362046 (
), 36204 ( ), 362047 (
), 346560 ( ).

RESULTS

LEXAS procedure to discover antibiotics and biosynthetic gene clusters in *Streptomyces*.
- (. 1)
- .l , 5,
- E. l
( )
.
-
(27)
E. l .
, -
. -
- -
. -
- -
. -
, -
E. l
6- , .
. 1 -
E. l 10 - -
- . E.

l 10 , . *E. l* - - 21 - (.1 . 1). 21, , *C4* , *A*, *K2* (22). 2, - . 10 / 21 6- . . *l* 5 , . *l l* . *l* . . *l* 5 - . , . *l* 5 , , - - , (20). 3 . , . *l* 5 , . 4 (4) (4). 4 . *l l* 1152, . *l l* 1154, . *l* 1 21- . *l* 5 . 4 , , , , - .

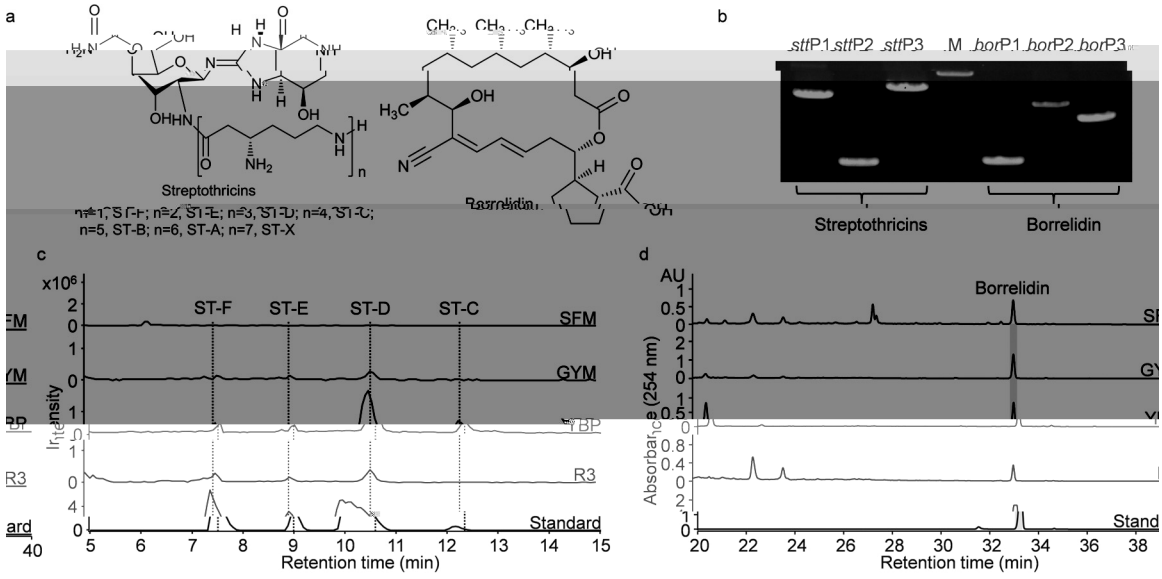
Production of streptothricins and borrelidin by *S. rochei* Sal35 and the LEXAS clones. , . 35, ( ) . . (30), (35), (36), - (37). . 35. ( ) - - *5* *19* *A* 1.5, 0.7, 1.7 , (30). ( ) *E*, - ( ) *A2*, *J* 0.7, 1.5, 1.2 (35) (2 . 2). . 35. (*l*) *l C l F l E* (*l*) 450 - *l F l K* β- *l H l l* . 35 (2). , (*l*)

2 , , *l* , *l*

(5′→3′)	( )
1-	743
1-	
2-	1,4
2-	
3-	1,201
3-	
l 1-	1,42
l 1-	
l 2-	1, 5
l 2-	
l 3-	1,271
l 3-	
l 1-	1,612
l 1-	
l 2-	675
l 2-	
l 3-	1, 7
l 3-	
1-	1,536
1-	
2-	727
2-	
3-	1,710
3-	

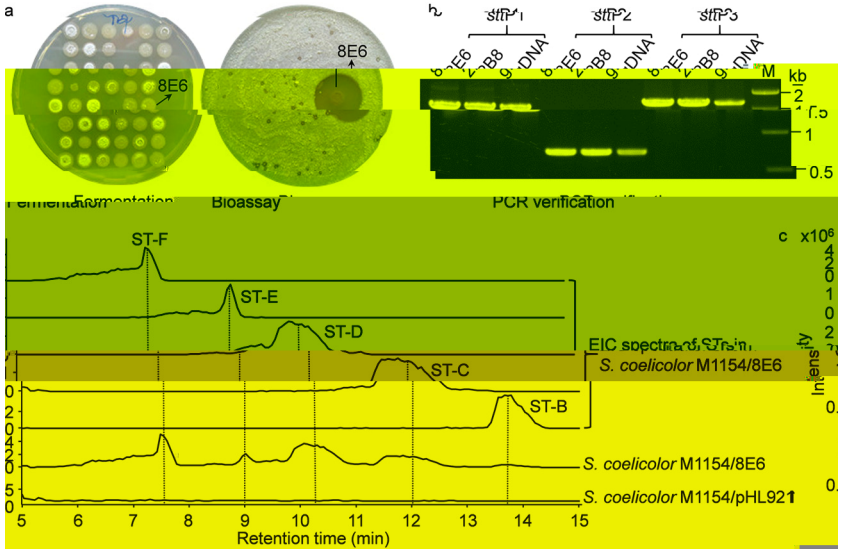
, , .

(*l*) , *l* *l* . 35. *l* *l* 210- 2- . 7434 4 (36, 37), . 35, - *l* *l* . - , . 35 - - . , (- , - , - , - , -) - , , - , . 35 - (.2 . 2). , . 35 76 *E. l* .100 , . (. 3). - . *l* - 3% (.3). , 12 (1.7%) . *l* 5 - (. 3 3 . 4).



2 . 35. ( ) . 35. ( ) . 35 . 1- , 2, 1.5, 1 . 2. ( ) . 35 , , , 3 . - (/ , 503.4), - (/ , 631.6), - (/ , 75 .7), - (/ , . 2 . ( ) . 35 , , , 3 . , .

12 , (.3 . 6). , 4 (, . 35, . *l l* 1154/ 6 - 3 .4 . 5). (.3 . 2). 1 (1) 2 . (2 6) (.3 . 6). , 3 4 . 6 - 1 - *E. l*



. 35 . ( ) . 6 . *M.* ^{2}155. ( ) 2 6. ( ) . ( ) . *l l* 1154/ 6 . *l l* 1154/ 21 3 . , . .

35

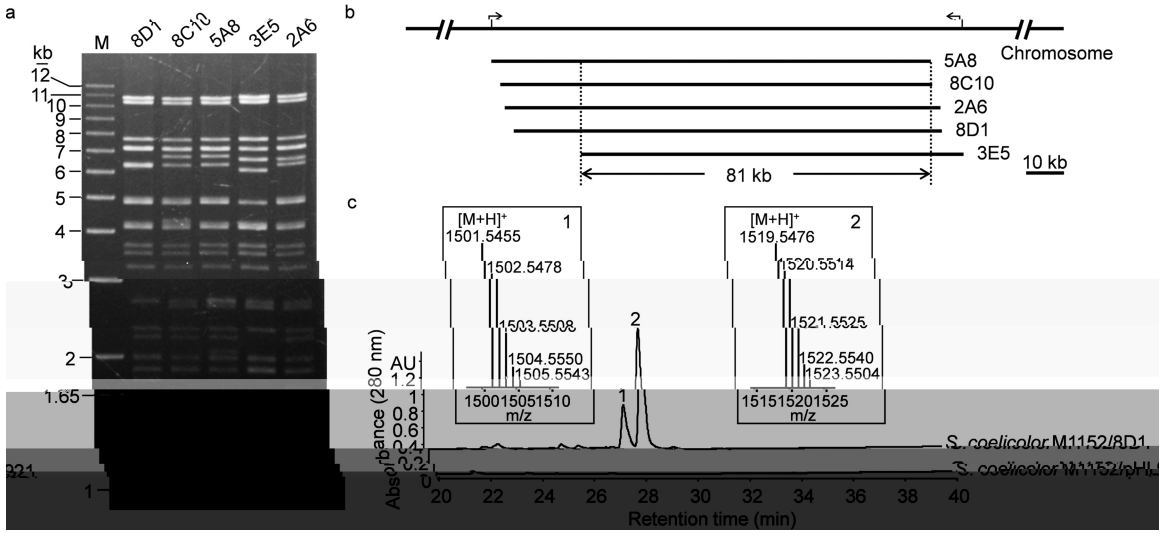
( ) ( )

1	1	.
2	2 6	. , *B.* , *M.* ^{2}155, .
3	2 6, 3 5, 5 10, 1	.
4	2 3, 6 1, 6 11, 11	. , *B.*

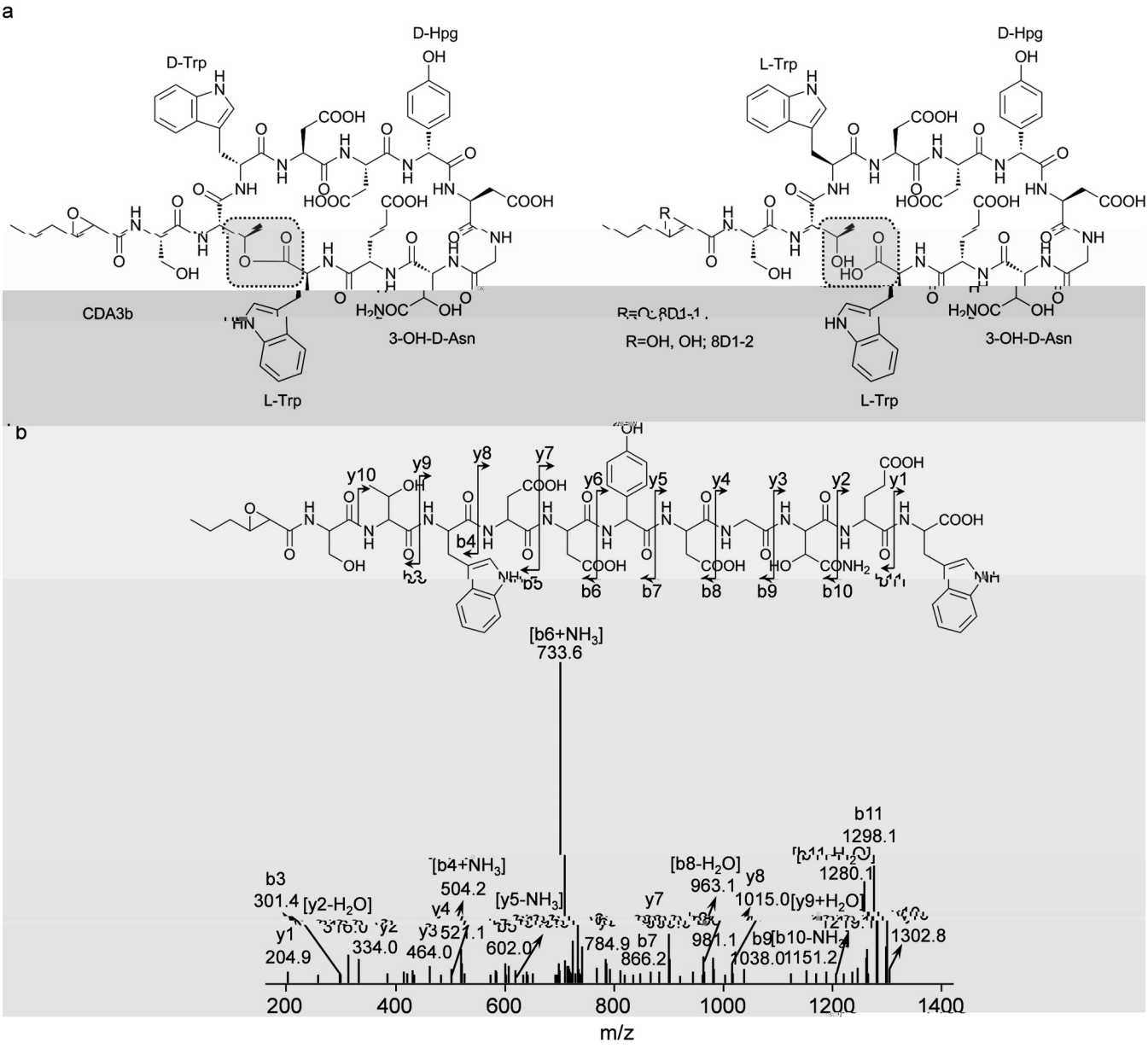
. 4 .

. *l* , 5, - . 3 4 - .

Identification of two linear lipopeptides from BAC clone of contig 3. , 1 3, . *l* 1 . *l l* 1152, . - *l l* 1154. 3 , , - - . , . *l l* 1154/ 1 . *l l* 1152/ 1 (.4). - - 1 (1-1 / , 1,501.5455 ^{66} ^{14} ^{27} + $^{+}$) 2 (1-2 / , 1,51 .5476 ^{66} ^{14} ^{2} + $^{+}$) (.4), - (3). 1-1 1-2 . *l* 1 1 (. 7). , - ( ) 3 , 3 , 4 , 4 . *l* 1 1 - (. 7), 1 . 1 *l* , 1 (.). 35 , - , (.)). , 20 1-1 15 1-2 - - . *l l* 1152/ 1 . - 1-1 / , ^{1} - , ^{13} - , 2 , 1- 2- 3- 4- 5- 6- 7- - 10- 11, ^{2} 3- -2- ( ) , 4- - 3- (3- -) (. 5 1). 1-2 - - 1-1 . ., 2,3- ( ) (.5 2). 1-1 1-2 - 1 2. 1-1 - 3 - . *l l* .*l* , , 3, 6, 3- - - - ( ) (3). - , 1-1 , , . 3, 6, 3- - 1-1 -, -, - , (.6 . 10 -

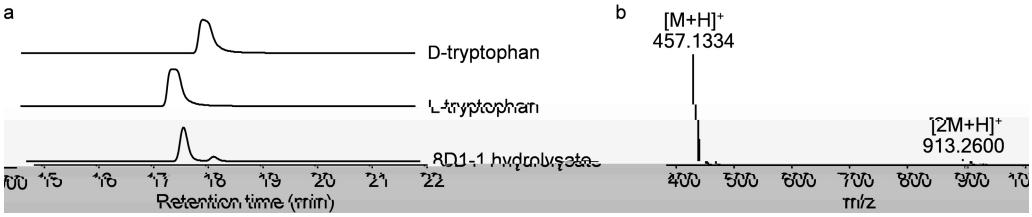


▲ . 35 .( ) 3 .( ) 3. . 35 . *l l* 1152/ 1 . *l l* 1152/ 21 . 2 , 1-1 (/ , 1,501.5455) 1-2 (/ , 1,51 .5476) . , .



1-1 1-2. ( ) 3 . l l 3(2) 1-1 1-2 . . 35. 3 , 1-1 1-2 ( ). 1-1 3 , 1-2. ( ) / 1-1 / .

). , 1-1 - 3 . l l 3(2) (40), (2 1, 2 2, 2 3), -2,3- Sal35 for the biosynthesis of 8D1-1/2. - - ( ), 3 1-1 -2 . l l 1152. 4- - , , - . 1 (.4 (. 11). ,). - 2 1, 2 2, 2 3 - . l l . , - . l l 3(2), . l l . 11 - 3(2) 2 1 2 2, . , 2 (41). . . 35. 2 (1 2) - - 3,



3 1-1. ( ) ’ - - 1-1 . 3 11
1 - , 3 11 3 - - , .
, 1 1 - - 3 (3). - 1-1,
1 . ( ) - / - .

6, 3- - (40). , -
- 3 1-1/2 ,
1 -
. - , ,
- . ,
1 .*l* 5 -
2 1, , 15 ,
15 (42) (. 12 - 1-1 -2 .
). 15 ,
- ,
6- .*l* 5 -
2 1 1 - 3 - - , ,
. *E.*

Bioinformatic analysis of contig 4. .*l* 5, *l* 6- . ,
4 (2 3, 6 1, 6 11, 11) , ,
. *B.* . , .*l* 1
11, -
. , -
,
100 (11) ,
.*l* 1 21.
B ll ,
(~40), (~70
), 1-1 -2 (~). ,
4 6 , -
(. 5), ,
- (3) 4 -
(32). .
25- , - 1.7% 3 4
- ( ), β- - (35)
( ) , β- - , ,
- , - , . ,
/ -
- ( )- .

3 .

DISCUSSION

(3 5, 43, 44). - . , *E. l* -12 -
, .
, , , ,
- , , , , -
- . , -
(45).
, .
. /

( ) - ( ) -
- -
(46), ( ) -
-
(47), ( ) -
-
(42), -
(4). , 2 3 1-1/2
-
3 ., - - - - - - .
3 / 4 / .
1 1 2 3
3- (4).
-
1-1 -2 - 3 - 3,
-
2 2
- 3- - -
, 2 3- , -
,
. . .
, 1-1 1-2
-
(47).
,
,
, ( ), -
( ), (1-1
-2), ,
.

ACKNOWLEDGMENTS

, ,
, , ,
2012 -
, ,
, ,
.

(73 200 11 05 63 2010 10 201),
(31370134), (-
0 077 533-0 001) , -
(15 1400401).

FUNDING INFORMATION

, , -
, (31370134),
, ( )
(-0 077 533-0 001),
( ) (15 1400401),
, ( )
(200 11 05 2010 10 201).
, ,
.

REFERENCES

1. , . 2012. 30 1 2010. 311 335. // . . /10.1021/ 200 06 .
2. , , - , . 2002. 3(2). 141 147. // . . /10.103 417141 .
3. , , , , , , , , , , , . 2014. . 1 412 421. // . . /10.1016/ . .2014.06 .034.
4. , . 200 . 2 161 165. // . . /10.1126 / .116 243.
5. , , . 200 . . 2 1362 13 . // . . /10.103 / 706 .
6. . 2014. - . 111 7266 7271. // . . /10.1073/ .140001 111.
7. , , . 2005. . 1 265 26 . // . . /10.103 731.
8. , , , . 2014. - , - . 7 56 7 . // . . /10.1002/ .201402632.
9. , , , , , . 2011. 51- . 10 625 6263. // . . /10.1073/ .101 07710
10. , . 2013. - . 2 11063 11067. // . . /10.1002/ .20130510 .
11. , . 1 7. -

1… - , . 2011. *l* - *l* 207 215. // . . /10.1111/ .1751-7 15 .2010.0021 . .

1… . , , , , , , , , - , , . 2013. *l* - 2 3 3 6. // . . /10.1021/ 3001003.

20. , , , . 2014. *l* , 5 . 1 6.

21. . 2004. - . 66 6 . // . . /10.112 .6 .66 -6 .2004.

22. , , , , , . 2004. - - - . 0 2452 2463. // . . /10.112 .70.4.2452-2463.2004.

23. , , . 1 7. *E* *l* - . 1 223 22 . // . . /10.1111/ .1574-6 6 7. 13 . .

24. , , . 1 5. *E* *l* . 2 1 0 1 6. // . . /10.10 3/ /23.11.1 0.

25. , . . 1 . *E.* *l* . 13 143. // . . /10.1016/016 525( ) 0054-1.

26. , , , , , , , . 2016. - . 1 572 575. // . . /10 .1021/ . .5 0360 .

27. , , , , . 2000. . , , .

2… , , , , . 1 6. *L* (12) *l* , *l* *l* 3(2). 1 7276 72 .

2… , , , , . 200 . *A*, *l* *l* . 215 2165. // . . /10.112 .0220 -0 .

30. , , , , - , , , . 2012. - . 7 1 7 7. // . . /10.103 .1040.

31. , . 2003. . 2 3 20.

32. , , , , , . 2013. 2.0 . 1 204 212. // . . /10.10 3/ / 44 .

33. , . 2004. ' . 2 231 247. // . . /10.1007 / 00726-004-011 0.

34. , , . 1 . 25- . 1 2 245 24 . // . . /10.1016/0003-26 7( ) 0405-7.

35. , , , , , , , , , , , . 2004. *l* 4055 - . 11 7.

36. , , . 2007. - . ( ) 0 700 70 // . . /10 .103 .2007. 0.

37. , , , , . 2006. - . 0 1 46 1 52. // . . /10.112 .00016-06.

3… , , , , , , , , , , , , , . 2016. 1202 1213. // . . /10.10 3/ . / 51.

3… , , , , , . 1 7. , , - *l* *l* 3(2) . 4 501. // . . /10.1002 / .1 704 .

40. , , , , , , , , , , . 2002. , , *l* *l* . 1175 11 . // . . /10.1016/ 1074-5521(02)00252-1.

41. , . 200 . . 1 217. // . . /10.1016/ 0076-6 (0)04 3.

42. , . 2000. . - 5775 57 . // . . /10.1021/ 2 002.

43. , , , , , , , . 2014. - . 10 63 6 . // . . /10.103 .165 .

44. , . 2014. 1 175 1 . // . . /10 .1007/ 102 5-013-13 - .

45. . 2007. . 2 125 131.

46. , , , . 1 7. . 2631 2650. // . . /10.1021 / 6001 .

47. ' , , , , , , . 2012. . 757 764. // . . /10.112 .05441-11.

4… , , , , . 2002. - . 10 301 310. // . . /10.1016/ 0 6 -2126(02)00716-5.

4… , , . 2004. - - - - *l* *l* . 2 15 2 25. // . . /10.1021/ 036140 .

50. , , , , , . 1 2. *l* . 111 61 6 // . . /10.1016/037 111 (2) 0603- .

51. , , , , , , , . 2012. - - *E* *l* - *l* *l* . 1 1 0. // . . /10.1007/ 002 -011-004 5.

52. , , , , , . 1 0. - *M* . 1 11 1 1 . // . . /10.1111/ .1365-2 5 0. 02040. .