

УТОЧНЕНИЕ СБОРКИ ГЕНОМА БАКТЕРИЙ *STAPHYLOCOCCUS WARNERI* 22-1

Е. В. Охремчук

Белорусский государственный университет, г. Минск;
katerinakolkovskaya@gmail.com;
науч. рук. – Л. Н. Валентович, канд. биол. наук, доц.

Бактерии *Staphylococcus warneri* являются представителями нормальной микробиоты человека. Ранее было выявлено, что данные микроорганизмы обладают высокой устойчивостью к присутствию органических растворителей в среде, что позволяет рассматривать их как потенциальных продуцентов биотоплива. В данной работе нами было осуществлено уточнение сборки генома бактерий *Staphylococcus warneri* 22-1. В частности, было осуществлено картирование межгенных спейсеров кластеров генов рНК, а также определено положение сдвоенного кластера генов рНК и установлено количество копий профага в геноме данного микроорганизма.

Ключевые слова: *Staphylococcus warneri*; сборка генома; органические растворители; профаг; кластеры генов рНК.

Бактерии *Staphylococcus warneri* являются представителями нормальной микробиоты кожи и слизистых покровов человека, других приматов и домашних животных. Ранее проведенные нами исследования показали, что *Staphylococcus warneri* обладают толерантностью к присутствию в среде различных органических растворителей, в том числе ацетону, бутанолу, этанолу и др. [1]. В настоящее время широко изучаются возможности получения биотоплива, в частности н-бутанола, в гетерологичных системах, что связано с низкой устойчивостью естественных продуцентов к присутствию бутанола в среде (до 2 % растворителя) [2]. Также особый интерес представляет изучение детерминант, которые обеспечивают микроорганизмам толерантность к органическим растворителям.

Нами был отобран устойчивый к бутанолу штамм *Staphylococcus warneri* 22-1 и секвенирован геном данного организма. Однако первичная сборка генома оставила нерешенными ряд задач. В ряде случаев, при сборке генома приходится сталкиваться с затруднениями, связанными с наличием повторяющихся последовательностей в геноме и невозможностью установить их точное расположение, основываясь лишь на данных, полученных при секвенировании. Данная проблема, в частности, существует с определением принадлежности межгенных спейсеров между генами 16S и 23S рНК в каждом из кластеров. Также в процессе сборки генома часто затруднительным является определение количества профагов, мобильных генетических элементов и др., что связано с их значительными размерами (могут достигать 100 т.п.н).

Целью работы являлось уточнение сборки генома *St. warneri* 22-1.

Были поставлены следующие задачи:

1. установить тип межгенного спейсера между 23S и 16S рДНК в каждом из кластеров;
2. установить точное количество копий профага в геноме с помощью количественной ПЦР с детекцией результатов в реальном времени;
3. установить положение двоянного кластера генов рРНК.

В геноме *St. warneri* 22-1 было выявлено 3 варианта межгенного пространства между генами 23S и 16S рРНК: NODE_95, NODE_75 и NODE_96.

Для установления типа межгенного пространства нами были разработаны праймеры, которые позволяют получить ампликоны, несущие часть последовательности 23S рДНК, а также часть последовательности, находящейся перед 16S рДНК. С помощью данных праймеров методом ПЦР нами были получены целевые ампликоны, которые затем были секвенированы по методу Сэнгера.

Выравнивание нуклеотидных последовательностей, полученных в результате секвенирования, позволяет заключить, что наиболее распространенным типом спейсеров между последовательностями генов 16S рРНК и 23S рРНК является NODE_95. Данный тип межгенного пространства встречается в геноме *St. warneri* 22-1 четыре раза (рис. 1).

N10	-TAGCCCGTATCGGAAGGTGCGCGCTGGATCACCTCCCTTCTAAGGATATATTCGGAAACATCTTCGTAGAAGATGAG	74
N7	CTCA	47
N8	CTCA	0
NODE95	GTAGCCCGTATCGGAAGGTGCGCGCTGGATCACCTCCCTTCTAAGGATATATTCGGAAACATCTTCGTAGAAGATGAG	75
N5	GTAGCCCGTATCGGAAGGTGCGCGCTGGATCACCTCCCTTCTAAGGATATATTCGGAAACATCTTCGTAGAAGATGAG	75
N10	GAATAACATTGACATATTTGATTCAGTTTTGAATGTTTATATAAAGATTCAACATATCTTCTTTTTCAAAGAAAGC	149
N7	GAATAACATTGACATATTTGATTCAGTTTTGAATGTTTATATAAAGATTCAACATATCTTCTTTTTCAAAGAAAGC	121
N8	GAATAACATTGACATATTTGATTCAGTTTTGAATGTTTATATAAAGATTCAACATATCTTCTTTTTCAAAGAAAGC	0
NODE95	GAATAACATTGACATATTTGATTCAGTTTTGAATGTTTATATAAAGATTCAACATATCTTCTTTTTCAAAGAAAGC	150
N5	GAATAACATTGACATATTTGATTCAGTTTTGAATGTTTATATAAAGATTCAACATATCTTCTTTTTCAAAGAAAGC	150
N10	GATTCGTAGACAAATGCGTCGAGTTACTTTAAAGCGAAGTTTACTAGATAAATGAGTATTTAAAGTAATGATAAC	224
N7	GATTCGTAGACAAATGCGTCGAGTTACTTTAAAGCGAAGTTTACTAGATAAATGAGTATTTAAAGTAATGATAAC	196
N8	GATTCGTAGACAAATGCGTCGAGTTACTTTAAAGCGAAGTTTACTAGATAAATGAGTATTTAAAGTAATGATAAC	29
NODE95	GATTCGTAGACAAATGCGTCGAGTTACTTTAAAGCGAAGTTTACTAGATAAATGAGTATTTAAAGTAATGATAAC	225
N5	GATTCGTAGACAAATGCGTCGAGTTACTTTAAAGCGAAGTTTACTAGATAAATGAGTATTTAAAGTAATGATAAC	225
N10	AAATCAGTATGCAAAACGCTTGACTGGAAATAGAAGTTGTACATTGAAAACAGATAAGTAAGTAAATAGATTTTT	299
N7	AAATCAGTATGCAAAACGCTTGACTGGAAATAGAAGTTGTACATTGAAAACAGATAAGTAAGTAAATAGATTTTT	271
N8	AAATCAGTATGCAAAACGCTTGACTGGAAATAGAAGTTGTACATTGAAAACAGATAAGTAAGTAAATAGATTTTT	104
NODE95	AAATCAGTATGCAAAACGCTTGACTGGAAATAGAAGTTGTACATTGAAAACAGATAAGTAAGTAAATAGATTTTT	300
N5	AAATCAGTATGCAAAACGCTTGACTGGAAATAGAAGTTGTACATTGAAAACAGATAAGTAAGTAAATAGATTTTT	300
N10	ACCAAGCAAAAACCGAGTGAATTAGAGTTTTAAAGCTTGAATTCATAAGAAATAATCGCTAGTGTTCGAAAGAAC	374
N7	ACCAAGCAAAAACCGAGTGAATTAGAGTTTTAAAGCTTGAATTCATAAGAAATAATCGCTAGTGTTCGAAAGAAC	346
N8	ACCAAGCAAAAACCGAGTGAATTAGAGTTTTAAAGCTTGAATTCATAAGAAATAATCGCTAGTGTTCGAAAGAAC	179
NODE95	ACCAAGCAAAAACCGAGTGAATTAGAGTTTTAAAGCTTGAATTCATAAGAAATAATCGCTAGTGTTCGAAAGAAC	375
N5	ACCAAGCAAAAACCGAGTGAATTAGAGTTTTAAAGCTTGAATTCATAAGAAATAATCGCTAGTGTTCGAAAGAAC	375
N10	CTCACAAAGATTA	386
N7	CTCACAAAGATTA	358
N8	CTCACAAAGATTA	191
NODE95	CTCACAAAGATTA	387
N5	CTCACAAAGATTA	387

Рис. 1. Выравнивание нуклеотидных последовательностей, полученных в результате секвенирования гибридных конструкций (N7-23S, N8-23S, N5-23S, N10-23S) с нуклеотидной последовательностью спейсера типа NODE_95

Спейсер, который примыкает к контигу NODE_6, относится к типу NODE_75 (рис. 2).

```

NODE7 TGGGGTGAAGTCGTAAACAAGGTAGCCGTATCGGAAAGGTGCGGCTGGATCA 50
N6-23 TGGGGTGAAGTCGTAAACAAGGTAGCCGTATCGGSAAGGTGCGGCTGGATCA 50

NODE7 CCTCC TTTCTAAGGATA TATTCGGAACA TCTT CGTAG AAGA TGAGGAATA 100
N6-23 CCTCC TTTCTAAGGATA TATTCGGAACA TCTT CGTAG AAGA TGAGGAATA 100

NODE7 ACATTGACATATTTGATATTCAGTTTTGATTGTTTTGTTTTACAAATAATTAA 150
N6-23 ACATTGACATATTTGATATTCAGTTTTGATTGTTTTGTTTTACAAATAATTAA 150

NODE7 TGGGCCTATAGCTCAGCTGGTTAGAGCGCACGCCTGATAAGCGTGAGGTC 200
N6-23 TGGGCCTATAGCTCAGCTGGTTAGAGCGCACGCCTGATAAGCGTGAGGTC 200

NODE7 GGTGGTT CGAGTCCACTTAGGCCCA CCA TTAATTAATATTTATTGGGGG 250
N6-23 GGTGGTT CGAGTCCACTTAGGCCCA CCA TTAATTAATATTTATTGGGGG 250

NODE7 CTTAGCTCAGCTGGGAGAGCGCCTTCTTTGCA CGCAGGAGGTCAGCGGTT 300
N6-23 CTTAGCTCAGCTGGGAGAGCGCCTTCTTTGCA CGCAGGAGGTCAGCGGTT 300

NODE7 CGATCCCGCTAGTCTCCACCACTTAT AATTGTA CATTGAAAACTAGATAAG 350
N6-23 CGATCCCGCTAGTCTCCACCACTTAT AATTGTA CATTGAAAACTAGATAAG 350

NODE7 TAAGTAAAA TAGATTTTCCAAAGCAAAA CCGAGTGAATTAGAGTTTTAAA 400
N6-23 TAAGTAAAA TAGATTTTCCAAAGCAAAA CCGAGTGAATTAGAGTTTTAAA 400

NODE7 GCTTGAATTCATAAGAAATAATCGCTAGTGTTCGAAAGAACA CT CACAAG 450
N6-23 GCTTGAATTCATAAGAAATAATCGCTAGTGTTCGAAAGAACA CT CACAAG 450

NODE7 ATTA 454
N6-23 ATTA 453

```

Рис. 2. Выравнивание нуклеотидных последовательностей, полученных в результате секвенирования ампликона N6-23S с нуклеотидной последовательностью спейсера типа NODE_75

По данным геномного сиквенса *St. warneri* 22-1, в геноме присутствует сдвоенный кластер генов рРНК, соединенный контигом NODE_85. Для определения положения данного кластера нами были проанализированы близкородственные геномы, секвенированные с помощью технологии PacBio из БД GenBank WGS [3] и было установлено, что сдвоенный кластер располагается между контигами NODE_2 и NODE_10. Кроме того, так как ранее нами было установлено, что кластер генов рРНК, прилегающий к контигу NODE_10, имеет спейсер типа NODE_95, нами было заключено, что спейсер кластера рДНК, прилежащий к NODE_2, относится к типу NODE_96.

Таким образом, нами было осуществлено физическое картирование межгенных спейсеров. Четыре из них относились к типу NODE_95, один – к NODE_96 и один – к NODE_75.

Далее нами были разработаны праймеры для установления точного количества кластеров рДНК и копий профага в геноме *St. warneri* 22-1.

ПЦР в реальном времени позволила установить, что количество копий профага в образце находится в диапазоне от 7 до 8. Точное количество затруднительно определить в связи с различной эффективностью работы праймеров. Расширенный эксперимент с использованием дополнительных пар праймеров к NODE9 (профаг) и NODE13 (гены рРНК, которые находятся в 6 копиях) показал схожие с ранее полученными результаты в отношении количества профага, 8 копий.

Также нами была проведена ПЦР в реальном времени с образцом тотальной ДНК, выделенным на несколько недель позже, чем ДНК, которая использовалась в ранее описанных экспериментах. Результаты данной ПЦР показали, что количество копий профага в данном случае рав-

но двум. Полученные результаты позволяют говорить о том, что данный фаг может находиться в виде эписомы, а в ряде условий часть копий может исключаться из генома и переходить в литический цикл. Таким образом, исходя из коэффициента встречаемости, мы можем заключить, что в пробе ДНК, которая использовалась для полногеномного секвенирования, действительно могло присутствовать 4 копии профага.

Подводя итог можно заключить, что нами была уточнена первая версия сборки генома *St. warneri* 22-1 (рис. 3).

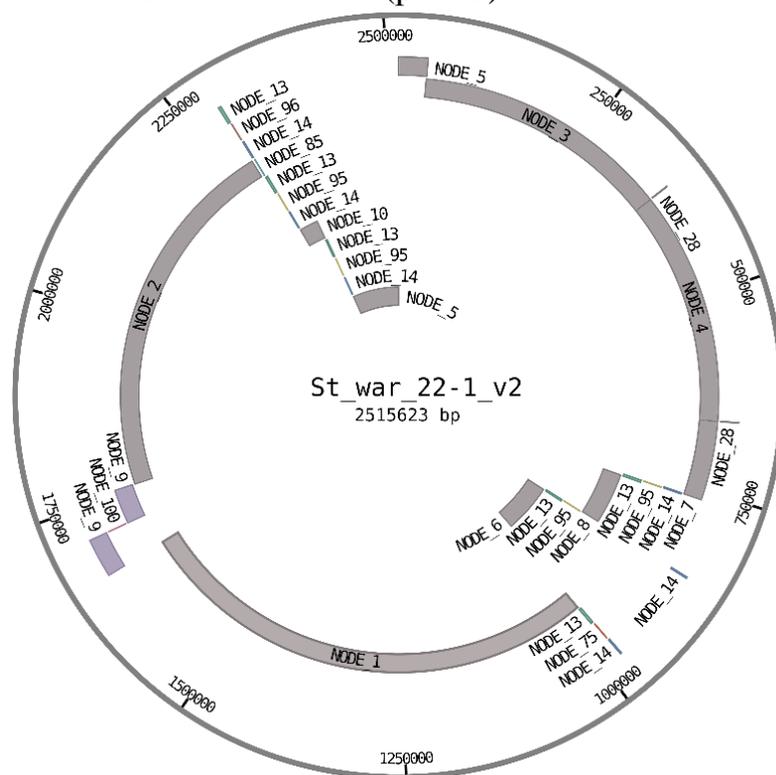


Рис. 3. Карта уточненной сборки генома *St. warneri* 22-1

Библиографические ссылки

1. Zhang, J. Isolation and characterization of butanol-tolerant *Staphylococcus aureus* / J.Zhang [et al.] // Bio-technol. Lett. – 2016. – Vol. 38, № 11 – P. 1929-1934.
2. Цинкевич, С. В., Евдокимова О.В., Валентович Л.Н. Исследование физиолого-биохимических характеристик бактерий рода *Staphylococcus* – потенциальных продуцентов органических растворителей (биотоплива) // Микробные биотехнологии: фундаментальные и прикладные аспекты: тез. докл. X Междунар. науч. конф., Минск, 5- 9 июня 2017 г. / Институт микробиологии НАН Беларуси – Минск: Беларуская навука. – Минск, 2017. С. 103–105.
3. Altschul, S.F. Basic local alignment search tool / S.F. Altschul [et al.] // J. Mol. Biol. – 1990. – Vol. 215, № 3. – P. 403–410.