

DOI: 10.25205/978-5-4437-1843-9-12

АНАЛИЗ РАСПРЕДЕЛЕНИЯ МЕТАБОЛИЧЕСКИХ ПУТЕЙ В ПАНГЕНОМЕ БАКТЕРИЙ, НЕСУЩИХ ГЕН *LUXR*^{*}

ANALYSIS OF THE DISTRIBUTION OF METABOLIC PATHWAYS IN THE PANGENOME OF BACTERIA CARRYING THE *LUXR* GENE

В. Б. Виниченко¹, К. С. Согомонян¹, Ф. М. Шматов^{1,2}, Л. Г. Данилов¹, К. С. Антонец^{1,2}

¹Санкт-Петербургский государственный университет

²Всероссийский НИИ сельскохозяйственной микробиологии, Санкт-Петербург

V. B. Vinichenko¹, K. S. Sogomonian¹, F. M. Shmatov^{1,2}, L. G. Danilov¹, K. S. Antonets^{1,2}

¹Saint Petersburg State University

²All-Russia Research Institute for Agricultural Microbiology, Saint Petersburg

✉ vinichenko.nika@yandex.ru

Аннотация

Квorum-сенсинг регулирует вирулентность бактерий через белок LuxR. Проанализированы 3358 бактериальных геномов из NCBI RefSeq с проверкой качества и выявлением генов *luxR*. Обнаружено различие в метаболических путях между геномами с *luxR* и без него, что указывает на роль *luxR* в регуляции метаболизма и квorum-сенсинга.

Abstract

Quorum sensing regulates bacterial virulence through the LuxR protein. A total of 3,358 bacterial genomes from NCBI RefSeq were analyzed with quality control and identification of *luxR* genes. Differences in metabolic pathways were found between genomes with and without *luxR*, indicating the role of *luxR* in metabolism regulation and quorum sensing.

Квorum-сенсинг — это система регуляции экспрессии генов бактерий, которая отвечает за ключевые факторы вирулентности, включая формирование биопленок и секрецию токсинов. Исследования метаболизма бактерий позволяет изучить биохимические процессы, осуществляющие регуляцию работы квorum-сенсинга, что может быть использовано для подавления развития и снижения вирулентности патогенных бактерий.

Ген *luxR*, кодирующий одноименный белок LuxR, играющий центральную роль в регуляции экспрессии генов посредством системы квorum-сенсинга [1], был выбран в качестве объекта для дальнейшего изучения.

Нами были выгружены геномные последовательности всех бактерий, представленных в открытой базе данных NCBI RefSeq Assembly (на 09.04.2025), после чего для обеспечения надежности дальнейшего анализа был проведен многоэтапный контроль качества. Сначала была выполнена фильтрация датасета на основе характеристик, указанных в NCBI, включая статус Assembly level “Complete Genome”, представленность генов по CheckM (версия: 1.2.3) не менее 97,5 %, а также подтверждение принадлежности организмов к царству бактерий (NCBI TaxID: 2).

Для оценки качества геномных сборок использовалась программа QUAST (версия: 5.3.0), которая позволила определить ключевые параметры: N50 — показатель непрерывности сборки, отражающий длину контига, при достижении суммарной длины которого достигается половина общей длины генома; GC-состав, важный для выявления возможных ошибок секвенирования или контаминации; общую длину генома и количество контигов. Дополнительно с помощью программы BUSCO (версия: 5.8.3) была проведена оценка качества аннотации, включающая определение числа потенциально кодирующих участков (CDS) и полноты сборки геномов на основе сравнения с бактериальным набором данных из базы OrthoDB версии 12.0. Для фильтрации геномных сборок были установлены следующие критерии качества: значение N50 должно составлять не менее 90 % от общей длины собранного генома, количество кодирующих генов — не менее 500, а полнота сборки по BUSCO — не ниже 85 %.

В результате фильтрации было оставлено 3358 геномов. Поскольку белок LuxR характеризуется наличием С-концевого ДНК-связывающего (идентификатор PFA: PF00196) и N-концевого атоиндуктор-связывающего доменов (идентификатор PFA: PF03472), в этих геномных сборках с помощью программы HMMER (версия: 3.4)

* Исследование выполнено в рамках государственного задания по теме «Изучение генетических основ внутривидовой изменчивости, надорганизменных взаимодействий и таксономического разнообразия с использованием биоресурсных коллекций» (шифр 124032000041-1).

проводилась проверка наличия генов, кодирующих белки с обоими доменами LuxR. В отфильтрованном датасете выявлено, что 1838 геномов содержат белки с сохраненным С-концевым доменом, 1317 — гены, кодирующие белки с обоими доменами LuxR, и 203 генома не содержат генов, кодирующих белки с доменами LuxR. Для оценки аннотации метаболических путей использовалась программа GapSeq (версия: 1.3.1).

Среднее количество метаболических путей на геном составило 3,1 тысячи для всех исследуемых групп. Число полноценных метаболических путей, общих для всех геномов в каждой группе, составило 5 для геномов с *luxR* (из которых 1 — с белками, содержащими только С-домен, и 4 — с белками, обладающими обоими доменами) и 14 для геномов без *luxR*. Из 12 тысяч реакций, выявленных в геномах с *luxR*, лишь 4 реакции были закодированы во всех исследуемых геномах (2 — в геномах с белками, содержащими только С-домен, и 2 — с белками, обладающими обоими доменами), тогда как в геномах без *luxR* таких реакций оказалось 191.

Наличие гена *luxR*, вероятно, указывает на его возможную роль в регуляции метаболических процессов. В условиях экспрессии данного гена могут активироваться дополнительные метаболические пути, которые, возможно, функционируют как регуляторные системы для *luxR*, либо наоборот — участвуют в синтезе лигандов, необходимых для сигнальной активности. Полученные результаты будут использованы для дальнейшего анализа регуляции кворум-сенсинга.

Литература

1. Rajput A., Kumar M. In silico analyses of conservational, functional and phylogenetic distribution of the LuxI and LuxR homologs in Gram-positive bacteria // Sci. Rep. 2017. Vol. 7. Art. 6969. URL: <https://doi.org/10.1038/s41598-017-07241-5>.