

## АНАЛИЗ МИКРОСАТЕЛЛИТНЫХ ЛОКУСОВ ГЕНОМА *PHOMA* НА ОСНОВАНИИ ПОЛНОГЕНОМНОГО СЕКВЕНИРОВАНИЯ

<sup>1</sup>Пантелеев С.В., <sup>1</sup>Баранов О.Ю., <sup>2</sup>Ярмолович В.А.,

<sup>2</sup>Середич М.О., <sup>1</sup>Рубель И.Э.

<sup>1</sup>Институт леса НАН Беларуси

(г. Гомель, Беларусь)

<sup>2</sup>Белорусский государственный технологический университет

(г. Минск, Беларусь)

На основании использования методов секвенирования нового поколения на базе Ion PGM Torrent проведено определение нуклеотидной последовательности генома одного из доминирующих патогенов в лесных питомниках Беларуси – возбудителя фомоза сеянцев хвойных пород. В ходе картирования выявлено и проанализировано 1111 микросателлитных локусов генома *Phoma*. Разработан набор праймеров для диагностики фомоза растений и видовой идентификации возбудителей инфекции с применением SSR-PCR метода.

### ВВЕДЕНИЕ

Фомоз сеянцев лесных древесных видов является одним из наиболее распространенных заболеваний в лесных питомниках Беларуси [1]. Возбудители данного заболевания – географически широко распространенные анаморфные грибы из рода *Phoma*. Большинство представителей рода оппортунистические и факультативные паразиты, обитающие, в основном, в почве. В настоящее время описано более 220 таксонов фомоподобных грибов, включающих около 200 фитопатогенов, 14 лихенизированных грибов, 9 видов, вызывающих болезни человека и животных, и несколько видов-возбудителей микозов рыб [2]. Фитопатогенные виды *Phoma* представлены в основном возбудителями болезней сельскохозяйственных культур: картофеля, свеклы, капусты и других. В последнее десятилетие имеются многочисленные сообщения о вредоносности грибов данного рода на хвойных (пихта, ель, сосна, лиственница, псевдотсуга) и лиственных (клен, ясень, орех, березу) древесных растениях. Однако диагностика фомоза затруднена, так как его возбудители являются таксономически сложной группой, что связано с бесполой природой большинства видов, высокой морфологической изменчивостью в естественных условиях и полифилитическим происхождением [3, 4].

В связи с этим актуальным является разработка и совершенствование методов диагностики возбудителей фомоза растений. Одним из эффективных методов изучения генетического разнообразия и паспортизации видов является SSR-анализ.

SSR-анализ, или микросателлитный анализ (с англ. Simple Sequence Repeat – простые короткие повторы). Значительная часть повторяющейся (сателлитной) ДНК состоит из tandemно повторяющихся копий различной длины. Наиболее часто используемые в ходе исследований SSR-локусы ха-

рактируются ди-, три- и тетрануклеотидными повторами. Мутации типа «инсерция/делеция», изменяют число повторов, т. е. общую длину такой многокопийной последовательности. Таким образом, аллели данных локусов отличаются друг от друга числом tandemно повторяющихся копий. Следует отметить преимущества данного типа маркеров: кодоминантный характер наследования; высокий уровень изменчивости – высокая разрешающая сила анализа (большое количество аллелей, высокий уровень гетерозиготности). Эти особенности позволяют использовать данный метод для видовой и внутривидовой паспортизации организмов [5].

Целью исследования является анализ микросателлитных локусов генома *Phoma* на основании полногеномного секвенирования для разработки набора молекулярно-генетических маркеров для диагностики фомоза.

## МАТЕРИАЛЫ И МЕТОДЫ

Объектом исследования являлся патогенный гриб *Phoma sp.1* из коллекции чистых культур УО «БГТУ». Данный вид был изолирован с пораженного посадочного материала хвойных пород микробиологическим методом [6]. Для выделения ДНК использовался мицелий гриба с применением СТАВ-модифицированного метода [5]. Этапы пробоподготовки парноконцевых ДНК-библиотек и полногеномное секвенирование выполнены на базе Ion PGM Torrent согласно инструкции фирмы-производителя.

## РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

В результате полногеномного секвенирования ДНК возбудителя фомоза растений получено порядка 3 млн. парноконцевых чтений общей длиной 850 млн. п.н., что с учетом размера генома *Phoma* в 40 млн. п.н. обеспечивает 21-кратное покрытие генома. Первичная биоинформатическая обработка результатов секвенирования осуществлялась при помощи программы Seqman NGen из пакета программного обеспечения DNASTAR Lasergene 12.1. Параметры обработки были выбраны по умолчанию, включая функцию удаления последовательностей праймеров и адаптеров из сиквенса.

В ходе первичной обработки данных получено 45651 контиг. С использованием программы SeqmanPro из пакета программного обеспечения DNASTAR Lasergene проведена более качественная пересборка парноконцевых чтений, в результате которой получены более длинные последовательности. Из полученного пула данных произведена выборка и картирование рибосомального оперона (7524 п.н.) и митохондрия (31916 п.н.) на основании сравнительного анализа с геномом близкородственного вида *Leptosphaeria maculans*, изолят JN3. Полученные сборки частично депонированы в международном геномном банке NCBI (ID KM387394.1) [7]. Сборка полного генома затрудняется из-за большого количества повторяющихся последовательностей.

На следующем этапе проводился анализ микросателлитных локусов (SSRs). Для этого в программе SeqmanPro из контигов был создан модельный

скаффолд размером 24117660 п.н. Выявление tandemных повторов в геноме осуществлялось с помощью программы SciRoKo [8] в режиме поиска идеальных повторов (PerfectRepeatsMode).

В модельном скаффолде выявлено 1111 микросателлитных локусов. Типы повторов указаны в таблице 1.

Таблица 1 – Микросателлитные локусы, выявленные в геноме *Phoma*

Тип повтора	Количество локусов	Средняя длина локуса, п. о.	Встречаемость на 1 млн. п. о.
Мононуклеотидный	65	20,03	2,70
Динуклеотидный	40	20,38	1,66
Тринуклеотидный	286	21,28	11,86
Тетрануклеотидный	102	19,72	4,23
Пентануклеотидный	288	18,93	11,94
Гексануклеотидный	330	23,86	13,68

Аллели микросателлитных локусов отличаются друг от друга числом tandemно повторяющихся копий. Анализ данного количества повторов осуществляют с помощью ПЦР-амплификации, используя праймеры, комплементарные уникальным последовательностям (доменам), которыми фланкирован каждый микросателлитный локус.

При картировании геномов предпочтение отдают последовательностям с повторами небольшой длины. Микросателлиты с мононуклеотидным типом повторов характеризуются высокой скоростью изменчивости, обусловленной «проскальзыванием» фермента ДНК-полимеразы при репликации ДНК и точечными мутациями (SNP). В свою очередь более длинные последовательности амплифицируются с помощью ПЦР с меньшей эффективностью, чем короткие мотивы. Поэтому для исследования полиморфизма аллелей микросателлитных локусов генотипа *Phoma* в программах разработки праймеров использовались динуклеотидные мотивы.

С использованием специального программного обеспечения Primer3web version 4.0.0 [9] и Primer Blast [10] проведен дизайн праймеров для специфической диагностики видов *Phoma*.

Для увеличения производительности и повышения качества анализа в программной среде проводилась оптимизация термодинамических параметров ПЦР. Так как при несоответствии показателей амплификации степени специфичности, сконструированные праймеры могут послужить причиной возникновения неспецифических продуктов и образованию димеров, снижающих или подавляющих образование искомого продукта. Оптимально праймер должен иметь размер от 15 до 30 нуклеотидов, так как его длина пропорциональна эффективности отжига. Пара праймеров должна иметь примерно одинаковую температуру плавления (разница не должна превышать 4-6 °C). В противном случае функциональность их будет некорректной. Праймер должен обладать 100%-ной комплементарностью по отношению к сайту-мишени и не распозна-

вать близкие по нуклеотидному составу последовательности других видов. В структуре праймеров должна отсутствовать внутренняя гомология и гомология между ними, чтобы предотвратить образование «шпилек», димеров и прочих частично двухцепочечных структур [11].

С учетом вышеперечисленных критериев условия подбора праймеров с помощью программного обеспечения были следующие: температура плавления праймеров ( $T_m$ ),  $^{\circ}\text{C}$  – 60; различия в температурах плавления праймеров, не более – 5  $^{\circ}\text{C}$ ; максимальное количество GC-нуклеотидов на 3'-конце праймера – 2; длина праймера – 17-23; GC-состав праймеров – 30,0-70,0%; максимальное количество GC-нуклеотидов на 3'-конце праймера – 2; наибольшее количество G-повторов – 3; максимальная стабильность 3'-конца,  $\Delta G$  – 9,0 ккал/моль; максимально допустимое сходство с не комплементарными последовательностями – 12; максимально допустимое сходство с эктопическими сайтами в последовательности-мишени – 12; максимально допустимая сумма сходства пары праймеров (по одному для каждого праймера) с не комплементарными последовательностями – 20; максимально допустимое суммированное сходство обоих праймеров с эктопическими сайтами в последовательности-мишени – 24; максимальная комплементарность – 8; максимальная комплементарность 3'-конца – 3; максимальная длина моонуклеотидных повторов – 5; Длина ампликона – 100-190 п.н.

На основании использования пула микросателлитных мотивов было сконструировано 10 пар праймеров, функциональность которых проверена в базе данных международного генного банка Национального центра биотехнологической информации (США) с применением программ nucleotide Blast и Primer Blast. Установлено, что разработанный набор характеризуется высокой степенью специфичности к фомоподобным грибам и неспецифичен к другим организмам. В качестве примера приводится структура нескольких пар праймеров: F-5'-AAGAATACCCGCATGGTGAC-3', R-5'-TGGAGGGGAGCGTGTAGATA-3' (размер видоспецифического продукта 124 п.н.); F-5'-TCATGCTATGCCTAGTGCTGT-3', R-5'-GACAGAGCGTGTGGTTCC-3' (размер специфического продукта 139 п.н.). В перспективе апробация разработанных праймеров с коллекцией ДНК возбудителей фомоза с использованием SSR-PCR метода.

## ЗАКЛЮЧЕНИЕ

С использованием методов секвенирования нового поколения на базе Ion PGM Torrent проведено определение нуклеотидной последовательности генома одного из доминирующих патогенов в лесных питомниках Беларуси – возбудителя фомоза сеянцев хвойных пород. На основании полученных данных выявлено и проанализировано 1111 микросателлитных локусов генома *Phoma*. Разработан набор праймеров для диагностики фомоза растений и видовой идентификации возбудителей инфекции с применением SSR-PCR метода.

## СПИСОК ЛИТЕРАТУРЫ

1. Молекулярно-генетические аспекты диагностики и идентификации возбудителей фомоза / О.Ю. Баранов [и др.] // Труды БГТУ. № 1: Лесное хозяйство. – Минск, 2014. – С. 198-201.
2. Phoma identification manual. Differentiation of specific and infra-specific taxa in culture / G. H. Voerema [et al.] // CAB International publishing. – 2004. – P.1-448.
3. Aveskamp M. Phylogeny and DNA-based identification in Phoma and related genera. Wageningen. 2014. pp. 18-24.
4. DNA phylogeny reveals polyphyly of Phoma section Peyronellaea and multiple taxonomic novelties / M. M. Aveskamp [et al.] // Mycologia. – 2009. – Vol. 101. – P. 363-382.
5. Падутов В.Е. Методы молекулярно-генетического анализа / В.Е. Падутов, О.Ю. Баранов, Е.В. Воропаев. – Мн.: Юнипол, 2007. – 176 с.
6. Павлович С.А. Медицинская микробиология. Практикум / С.А. Павлович, К.Д. Пяткин. – Минск: Вышэйшая школа, 1993. – 200 с.
7. The National Center for Biotechnological Information, NCBI [Electronic resource]. Mode of access: <http://www.ncbi.nlm.nih.gov/>. – Date of access: 05.03.2015.
8. SciRoKo SSR-search module [Electronic resource]. Mode of access: <http://kofler.or.at/bioinformatics/SciRoKo/index.html/>. – Date of access: 05.03.2015.
9. Online tool to design and analyze primers for PCR: Primer3web version 4.0.0 [Electronic resource]. Mode of access: <http://primer3.ut.ee/> (date of access: 06.05.2015).
10. The National Center for Biotechnology Information (NCBI) [Electronic resource]. Mode of access: <http://www.ncbi.nlm.nih.gov/tools/primer-blast/> (date of access: 07.05.2015).
11. Dieffenbach C.W. General Concepts for PCR Primer Design / C.W. Dieffenbach, T.M. Lowe. J., G.S. Dveksler // PCR Methods App. – 1993. – Vol. 3. – P. 30-37.

### THE ANALYSIS OF MICROSATELLITE LOCI OF THE *PHOMA* GENOME ON THE BASIS OF FULL-GENOMIC SEQUENCING

*Panteleev S.V., Baranov O.Yu., Yarmolovich V.A.,  
Seredich M.O., Rubel I.E.*

*Based on the use of next-generation sequencing methods on the basis of Ion PGM Torrent carried out the determination of the nucleotide sequence of the genome of one of the dominant pathogens in forest nurseries of Belarus – causative agent of Phoma Blight. During the genome mapping identified and analyzed 1111 microsatellite loci of Phoma. Developed a set of primers for the diagnosis of Phoma Blight and causative agent of infectious disease using SSR-PCR method.*

Статья поступила в редколлегию 12.03.2015 г.

