

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ
Кафедра микробиологии

БАЛАШКО
Диана Сергеевна

***DE NOVO* СБОРКА МИТОХОНДРИАЛЬНОГО И
ХЛОРОПЛАСТНОГО ГЕНОМОВ ЯЧМЕНЯ
(*HORDEUM VULGARE*)**

Аннотация к магистерской диссертации

специальность 1-31 80 12 Микробиология

Научный руководитель:
к.б.н., заведующий сектором
биоинформатики Института генетики
и цитологии НАН Беларуси
Левданский О. Д.

Минск, 2023

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Ключевые слова: DE NOVO СБОРКА ХЛОРОПЛАСТНОГО ГЕНОМА, DE NOVO СБОРКА МИТОХОНДРИАЛЬНОГО ГЕНОМА, АНАЛИЗ ДАННЫХ СЕКВЕНИРОВАНИЯ, АССЕМБЛЕРЫ, ВЫРАВНИВАНИЕ, ХАРАКТЕРИСТИКА ПОЛУЧЕННОЙ СБОРКИ.

Цель работы — сборка хлоропластного и митохондриального генома *Hordeum vulgare* с использованием различных ассемблеров

Объекты исследования: данные секвенирования хлоропластного и митохондриального геномов *Hordeum vulgare*.

Полученные результаты. В результате исследования при сборке хлоропластных и митохондриального геномов *Hordeum vulgare* использовались такие ассемблеры как SPAdes, NOVOPlasty, GetOrganell и ABySS. Лучше всего, в качестве сборщиков хлоропластного и митохондриального генома *Hordeum vulgare*, справились SPAdes и ABySS. При сборке хлоропластного генома первому удалось восстановить последовательности, покрывающие референс на 81,61% (7_S15_L001) и 80,55% (10_S18_L001), а при сборке митохондриального: 83,90% и 86,73% соответственно. ABySS удалось собрать контиги, которые покрывают 88,17% (7_S15_L001) и 84,55% (10_S18_L001) референс хлоропластного генома *Hordeum vulgare*, а митохондриальный на 90,85% и 94,43% соответственно. GetOrganell неплохо зарекомендовал себя в сборке хлоропластных геномов 83,35% (7_S15_L001) и 78,49% (10_S18_L001) с образованием 6 и 5 контиг соответственно, но в ходе сборки митохондриального генома были собраны контиги, которые в общей сложности покрывают только 66,57% и 66,48% референса. NOVOPlasty неплохо справился со сборкой хлоропластного генома 80,37% (7_S15_L001) и 77,22% (10_S18_L001), при сборке митохондриального генома в процент совпадения с референсом равен 52,74% (7_S15_L001), а во втором 90,88% (10_S18_L001).

Структура магистерской диссертации: работа изложена на 68 страницах, состоит из разделов «Введение», «Общая характеристика работы», трех глав, разделов «Заключение», «Список использованных источников» из 42 наименований, а также «Приложение А», «Приложение Б» и «Приложение В». Работа содержит 18 иллюстрации и 4 таблицы.

АГУЛЬНАЯ ХАРАКТАРЫСТЫКА РАБОТЫ

Ключавыя словы: DE NOVO СБОРКА ХЛОРАПЛАСТНАГА ГЕНОМА, DE NOVO СБОРКА МІТАХАНДРЫЯЛЬНАГА ГЕНОМА, АНАЛІЗ ДАННЫХ СЕКВЕНІРАВАННЯ, АССЕМБЛЕРЫ, ВЫРАЎЛЕННЕ, ХАРАКТАРЫСТЫКА АТРЫМАНАЙ ЗБОРКІ.

Мэта работы — зборка хлорапластнага і мітахандрыяльнага генома *Hordeum vulgare* з выкарыстаннем розных ассэмблераў.

Аб'екты даследавання: дадзеныя секвеніравання хлорапластнага і мітахандрыяльнага геномаў *Hordeum vulgare*.

Атрыманыя вынікі. У выніку даследавання пры зборцы хлорапластных і мітахандрыяльнай геномаў *Hordeum vulgare* выкарыстоўваліся такія ассэмблеры як SPAdes, NOVOPlasty, GetOrganell і ABySS. Лепш за ўсё, у якасці зборшчыкаў хлорапластнага і мітахандрыяльнай геному *Hordeum vulgare*, справіліся SPAdes і ABySS. Пры зборцы хлорапластнага геному першаму атрымалася аднавіць паслядоўнасці, якія пакрываюць рэферэнс на 81,61% (7_S15_L001) і 80,55% (10_S18_L001), а пры зборцы мітахандрыяльнага: 83,90% і 86,73% адпаведна. ABySS атрымалася сабраць кантыгі, якія пакрываюць 88,17% (7_S15_L001) і 84,55% (10_S18_L001) рэферэнс хлорапластнага геному *Hordeum vulgare*, а мітахандрыяльнай на 90,85% і 94,43 GetOrganell нядрэнна зарэкамендаваў сябе ў зборцы хлорапластных геномаў 83,35% (7_S15_L001) і 78,49% (10_S18_L001) з адукацыяй 6 і 5 кантыг адпаведна, але падчас зборкі мітахандрыяльнай геному былі сабраны кантыгі, якія % і 66,48% рэферэнсу. NOVOPlasty нядрэнна справіўся са зборкай хлорапластнага геному 80,37% (7_S15_L001) і 77,22% (10_S18_L001), пры зборцы мітахандрыяльнай геному ў працэнт супадзення з рэферэнсам роўны 52,74% 1, 7 0_S18_L001).

Структура магістарскай дысертацыі: праца выкладзена на 68 старонках, складаецца з раздзелаў «Уводзіны», «Агульная характарыстыка працы», трох раздзелаў, раздзелаў «Заклучэнне», «Спіс выкарыстаных крыніц» з 42 найменняў, а таксама «Дадатак А», «Дадатак Б» і «Дадатак У». Праца змяшчае 18 ілюстрацыі і 4 табліцы.

GENERAL DESCRIPTION OF WORK

Key words: *DE NOVO* ASSEMBLY OF THE CHLOROPLAST GENOME, *DE NOVO* ASSEMBLY OF THE MITOCHONDRIAL GENOME, SEQUENCING DATA ANALYSIS, ASSEMBLERS, ALIGNMENT, CHARACTERIZATION OF THE COMPLETED ASSEMBLY.

The aim of the work is assembly of the chloroplast and mitochondrial genome of *Hordeum vulgare* using different assemblers.

Objects of the study: sequencing data from the chloroplast and mitochondrial genomes of *Hordeum vulgare*.

Research results. As a result of the study, when assembling the chloroplast and mitochondrial genomes of *Hordeum vulgare*, such assemblers as SPAdes, NOVOPlasty, GetOrganell, and ABySS were used. SPAdes and ABySS performed best as assemblers of the chloroplast and mitochondrial genome of *Hordeum vulgare*. When assembling the chloroplast genome, the first managed to restore sequences covering the reference by 81.61% (7_S15_L001) and 80.55% (10_S18_L001), and when assembling the mitochondrial: 83.90% and 86.73%, respectively. ABySS managed to collect contigs that cover 88.17% (7_S15_L001) and 84.55% (10_S18_L001) of the *Hordeum vulgare* chloroplast reference genome, and 90.85% and 94.43% of the mitochondrial genome, respectively. GetOrganell performed well in the assembly of chloroplast genomes 83.35% (7_S15_L001) and 78.49% (10_S18_L001) with the formation of 6 and 5 contigs, respectively, but during the assembly of the mitochondrial genome, contigs were assembled, which in total cover only 66.57 % and 66.48% of the reference. NOVOPlasty coped well with the assembly of the chloroplast genome 80.37% (7_S15_L001) and 77.22% (10_S18_L001), when assembling the mitochondrial genome, the percentage of coincidence with the reference is 52.74% (7_S15_L001), and in the second 90.88% (10_S18_L001).

Structure of the master's thesis: the work is set out on 68 pages, consists of sections "Introduction", "General Characteristics of Work", three chapters, sections "Summary", "List of references" of 42 titles, as well as "Appendix A", "Appendix B" and "Appendix C". The work contains 18 illustrations and 4 tables.