

Министерство науки и высшего образования РФ  
Федеральное государственное автономное  
образовательное учреждение высшего образования  
«**СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ**»  
Институт фундаментальной биологии и биотехнологии  
Кафедра геномики и биоинформатики

УТВЕРЖДАЮ:

Заведующий кафедрой

\_\_\_\_\_ И. Е. Ямских

«\_» \_\_\_\_\_ 2022 г.

**МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ**

«Выявление связей между структурой, функцией и таксономией в  
геномах митохондрий и хлоропластов»

06.04.01 Биология

06.04.01.06 Геномика и биоинформатика

Руководитель \_\_\_\_\_ проф., д.ф.-м.н. М.Г. Садовский  
подпись, дата должность, ученая степень

Выпускник \_\_\_\_\_ В.Д. Федотовская  
подпись, дата

Рецензент \_\_\_\_\_ проф., д.ф.-м.н. С.И. Барцев  
подпись, дата должность, ученая степень

Красноярск 2022

## РЕФЕРАТ

Магистерская диссертация по теме «Выявление связей между структурой, функцией и таксономией в геномах митохондрий и хлоропластов» содержит 30 страницы текстового документа, 9 иллюстраций, 4 таблицы и 33 использованных источника.

Ключевые слова: СТРУКТУРА, ФУНКЦИЯ, ТАКСОНОМИЯ, МЕТОД ДИНАМИЧЕСКИХ ЯДЕР, МЕТОД УПРУГИХ КАРТ, МИТОХОНДРИАЛЬНЫЙ ГЕНОМ, ХЛОРОПЛАСТНЫЙ ГЕНОМ.

Объект исследования — взаимосвязи между структурой нуклеотидных последовательностей (триплетным составом), кодируемой этими последовательностями функцией и таксономией носителей данных нуклеотидных последовательностей.

Цель работы — выявление связи между триплетным составом, функцией нуклеотидной последовательности и таксономией носителей генов органелл растений.

В результате проделанной работы были отобраны 168 митохондиальных и хлоропластных геномов 84 организмов, выделены гены субъединиц АТФ-синтаз и НАДН-дегидрогеназ, на основе этих последовательностей построены частотные словари триплетов и проведена их кластеризация методом упругих карт и методом динамических ядер. Состав кластеров был проверен с функциональной и таксономической точки зрения. С помощью метода динамических ядер были построены слоистые графы, показывающие перераспределение словарей по классам при увеличении классов. Оба метода кластеризации хорошо показывают, что кластеры формируются по функции, т. е. в кластер попадают словари, соответствующие одинаковым по функции генам, а не близким таксонам.

# СОДЕРЖАНИЕ

<b>ВВЕДЕНИЕ</b>	<b>4</b>
<b>1 Обзор литературы</b>	<b>9</b>
1.1 Митохондрии . . . . .	9
1.2 Геномы митохондрий . . . . .	10
1.3 Окислительное фосфорилирование . . . . .	11
1.4 Хлоропласты . . . . .	13
1.5 Геномы хлоропластов . . . . .	14
1.6 Фотосинтез . . . . .	14
1.7 Методы кластерного анализа . . . . .	17
<b>2 Материалы и методы</b>	<b>19</b>
2.1 Генетический материал . . . . .	19
2.2 Частотные словари . . . . .	20
2.3 Метод динамических ядер . . . . .	21
2.4 Метод упругих карт . . . . .	23
<b>3 Результаты и их обсуждения</b>	<b>25</b>
<b>ЗАКЛЮЧЕНИЕ</b>	<b>26</b>
<b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ</b>	<b>27</b>

## ВВЕДЕНИЕ

Изучение любого генетического материала требует его анализа с трёх сторон: анализ его структуры, его функции и таксономии его носителя. Эти стороны можно изучать по отдельности, но больший интерес вызывает их совместное изучение. Такому анализу генов хлоропластной и митохондриальной ДНК посвящена данная работа.

Актуальность работы связана с появлением большого количества данных, а также с быстро развивающимися технологиями секвенирования генетического материала. В данной работе изучаются гены, кодирующие субъединицы АТФ-синтазы и НАДН-дегидрогеназы митохондрий и хлоропластов. Эти гены, а также продукты, которые они кодируют, являются консервативными, т.к. выполняют наиболее важные функции в органеллах: синтез АТФ и перенос электронов в дыхательной цепи. ДНК могут быть использованы для изучения эволюции растений, построении филогении и проведении популяционного анализа. Также растения играют крайне важную роль в жизни человека и других живых организмов. Они осуществляют фотосинтез, являются источником питания, создают среду обитания для других организмов и т.д.

Результаты подобных исследований зависят от того, какой именно генетический материал изучается. В работах [19, 20] исследовалась связь между структурой геномов органелл (хлоропластов и митохондрий, соответственно) и таксономическим положением их носителей. Благодаря такому выбору генетического материала представляется возможным редуцировать задачу и рассматривать связь лишь между структурой и таксономией, исключая функцию, ведь функция у всех хлоропластов и митохондрий одна и та же. Таким же образом в работе [7] изучались генетические последовательности зрелых РНК генов 16S РНК бактерий и рассматривалась связь структура — таксономия.

Существует большое количество работ, в которых рассматривается структура генетического материала биологических объектов. Так, напри-

мер, в работе [25] было проведено сравнение таксономической структуры кормовых растений западного Забайкалья. В работах [7,8] изучалась классификация нуклеотидных последовательностей бактериальных РНК.

Также исследователи выделяют большое количество разных структур и методов анализа генетических последовательностей. Например, в работе [26] в качестве структуры рассматривался частотный словарь триплетов, а в качестве метода анализа использовался метод динамических ядер. В работе [27] также использовались частотные словари, но визуализация и кластеризация проводились с помощью метода главных компонент.

**Целью данной работы** является выявление связи между триплетным составом, функцией нуклеотидной последовательности и таксономией носителей генов органелл растений. Для проведения данного анализа были выбраны гены, кодирующие субъединицы АТФ-синтазы и НАДН-дегидрогеназы митохондрий и хлоропластов растений.

Для достижения данной цели были поставлены следующие **задачи**:

- 1) создание базы генов и их частотных словарей;
- 2) построение упругой карты для распределения генов субъединиц АТФ-синтаз и НАДН-дегидрогеназ органелл в пространстве частот триплетов;
- 3) проверка состава кластеров с точки зрения функции и таксономии.

**Объектом работы** являются взаимосвязи между структурой нуклеотидных последовательностей (триплетным составом), кодируемой этими последовательностями функцией и таксономией носителей данных нуклеотидных последовательностей.

**Предметом работы** являются частотные свойства нуклеотидных последовательностей (структура частотных словарей триплетов) и отношения между этими свойствами и функцией, кодируемой исходными последовательностями, а также отношения между частотными свойствами нуклеотидных последовательностей и таксономией носителей нуклеотидных последова-

тельностью.

**Апробация работы.** По результатам магистерской диссертации опубликовано 14 работ, в том числе — 10 в международных изданиях:

- [1] Колесникова А.И. Выявление связи тринуклеотидного состава генов и таксономии их носителей на примере генов митохондрий некоторых грибов / Колесникова А.И., Федотовская В.Д., Шпагина Т.О. // Материалы 56-й Международной научной студенческой конференции (МНСК). — 2018. — Vol. 56. — Стр. 18;
- [2] Колесникова А.И. Влияние функциональных различий сильнее влияния таксономических различий для генов семейства *atp* митохондрий грибов / Колесникова А.И., Федотовская В.Д., Шпагина Т.О., Садовский М.Г. // Моделирование неравновесных систем. / Материалы XXI Всероссийского семинара — 2018. — Vol. 21. — Стр. 49–54;
- [3] Федотовская В. Д. О соотношении влияния функциональных и таксономических различий для генов семейства *atp* митохондрий некоторых грибов / Федотовская В.Д., Шпагина Т.О., Колесникова А.И., Садовский М.Г. // Нейроинформатика, ее приложения и анализ данных. / Материалы XXVII Всероссийского семинара — 2019. — Vol. 27. — Стр. 110–115;
- [4] Sadovsky M., Fedotovskaya V., Kolesnikova A., Shpagina T., Putintseva Y. Function vs. Taxonomy: The Case of Fungi Mitochondria *atp* Synthase Genes // Lecture Notes in Computer Science, Springer Verlag. — 2019. — Vol. 11465 — Pp. 335–345;
- [5] Fedotovskaya V. Kolesnikova A., Shpagina T., Putintseva Y., Sadovsky M. Function Overcomes Taxonomy: Case of *atp* Genes of Fungi Mitochondria // Tenth International Conference Dynamical Systems Applied to Biology and Natural Sciences: Book of Abstracts. — 2019. — Vol. 10. — Pp. 176–176.
- [6] Fedotovskaya V., Kolesnikova A., Shpagina T., Sadovsky M. The Distribution of Fungal Mitochondrial *atp* Genes in Amino Acids Space // 11<sup>th</sup> International Conference Dynamical Systems Applied to Biology and Natural

Sciences: Book of Abstracts. — 2020. — Vol. 11. — Pp. 215-216.

- [7] Fedotovskaya V., Sadovsky M., Kolesnikova A., Shpagina T., Putintseva Y. Function vs. taxonomy: further reading from fungal mitochondrial *ATP* synthases // Lecture Notes in Computer Science, Springer Verlag. — 2020. — Vol. 12108 — Pp. 438–444.
- [8] Федотовская В. Д. и др. Распределение генов митохондрий грибов в пространстве частот триплетов // Systems Biology and Bioinformatics (SBB-2020). — 2020. — С. 103–103.
- [9] Федотовская В. Д. и др. Распределение генов митохондрий грибов в пространстве аминокислот // Bioinformatics Research Projects. — 2020. — С. 68.
- [10] Федотовская В. Д., Садовский М. Г. Триплетный состав общих генов митохондрий и хлоропластов растений выявляет их дифференциацию // Институт биоинформатики. Сборник тезисов. — 2021. — С. 85.
- [11] Fedotovskaya V., Sadovsky M. Triplet composition of common mitochondrial and chloroplast genes of plants reveals their differentiation // Plant Genetics, Genomics, Bioinformatics, and Biotechnology. — 2021. — С. 68–68.
- [12] Федотовская В. Д. Распределение генов *atp* митохондрий и хлоропластов семенных растений // БИОЛОГИЯ. — 2021. — С. 21–21.
- [13] Sadovsky M., Fedotovskaya V. Triplet distribution of common mitochondrial and chloroplast genes of plants // 13<sup>th</sup> International Conference Dynamical Systems Applied to Biology and Natural Sciences: Book of Abstracts. — 2022. — Vol. 13. — Pp. 335-336.
- [14] Sadovsky M. et al. Function Overcomes Taxonomy for Organella Genes Triplet Composition. — 2022.

Основные положения диссертации были доложены на следующих конференциях:

- 1) 56-я Международная научная студенческая конференция, устный доклад, Новосибирск;
- 2) X международная конференция «Dynamical Systems Applied to Biology

- and Natural Sciences» (DSABNS), стендовый доклад, Неаполь, 2019;
- 3) Международная конференция студентов, аспирантов и молодых ученых «Перспектив Свободный — 2019», устный доклад, Красноярск;
  - 4) 7<sup>th</sup> International Work-Conference on Bioinformatics and Biomedical Engineering (IWBBIO), устный доклад, Гранада, Испания;
  - 5) XI международная конференция «Dynamical Systems Applied to Biology and Natural Sciences» (DSABNS), стендовый доклад, Тренто, 2020;
  - 6) Международная конференция студентов, аспирантов и молодых ученых «Перспектив Свободный — 2020», устный доклад, Красноярск, 2020;
  - 7) Конференция Онлайн-школы по биоинформатике, устный доклад, Санкт-Петербург, 2020.
  - 8) 12-я Международная школа молодых ученых «Системная Биология и Биоинформатика», стендовый доклад, Ялта и Севастополь, 2020;
  - 9) 6th INTERNATIONAL SCIENTIFIC CONFERENCE “PLANT GENETICS, GENOMICS, BIOINFORMATICS AND BIOTECHNOLOGY” (PlantGen2021), стендовый доклад, Новосибирск, 2021;
  - 10) Конференция Летней школы по биоинформатике, стендовый доклад, Санкт-Петербург, 2021;
  - 11) 59-я Международная научная студенческая конференция, устный доклад, Новосибирск, 2021;
  - 12) 13-я Международная школа молодых ученых «Системная Биология и Биоинформатика», устный доклад, Новосибирск, 2021;
  - 13) XIII международная конференция «Dynamical Systems Applied to Biology and Natural Sciences» (DSABNS), стендовый доклад, 2022.



# 1. Обзор литературы

## 1.1 Митохондрии

Митохондрии — клеточные органеллы эукариотических клеток. Их число может достигать нескольких тысяч в одной клетке. Они представляют собой двумембранные органеллы эллипсоидной формы. Митохондрии содержатся в клетках эукариот. Внутренняя мембрана митохондрий образует складки (так называемы кристы), область ограниченная внутренней мембраной называется матриксом. У митохондрий есть собственный генетический материал, который локализован в матриксе. Митохондриальная ДНК представляет собой кольцевую молекулу. Главной функцией митохондрий является окисление органических веществ и синтез АТФ — источника энергии для клетки. Его осуществляет фермент АТФ-синтаза, который находится во внутренней мембране. Синтез АТФ проходит благодаря с использованием энергии окисляющихся веществ. Окисление происходит в кристах митохондрий, где локализованы ферменты, участвующие в системе переноса электронов.

Кроме описанных выше функций, недавно было выяснено, что митохондрии играют важную роль в регуляции апоптоза — запрограммированной клеточной смерти. Во время клеточного стресса разрыв митохондриальной мембраны или открытие высокопроницаемых каналов на внешней мембране митохондрий приводит к нарушению потенциала мембраны. Затем происходит высвобождению цитохрома с, белка AIF (apoptosis inducing factor — фактор, индуцирующий апоптоз) и последующую активацию каспаз. Эти белки участвуют в апоптозе. Кроме этого, митохондрии также участвуют в других видах регулируемой гибели клеток: некроптозе, ферроптозе и пироптозе [1].

Устройство генетического материала митохондрий позволяет выдвинуть гипотезу о их происхождении. Согласно эндосимбиотической гипотезе, митохондрия эволюционировала из бактерий, которые эндоцитировали более 1 млрд. лет назад. По одной из версий этой гипотезы, совре-

менные митохондрии — изначально бактерии, которые вступили в эндосимбиотические отношения с эукариотическими одноклеточными. Они, в свою очередь, приспособили систему окислительного фосфорилирования симбионтов к своим нуждам. Благодаря сравнению последовательностей митохондриальных геномов можно показать, что митохондрии произошли от определенного типа пурпурных фотосинтетических бактерий, потерявших способность к фотосинтезу [24].

## 1.2 Геномы митохондрий

Считается, что митохондриальная ДНК в основном имеет вид кольцевой двухцепочечной молекулы. Тем не менее, существуют данные показывающие, что митохондриальная ДНК может иметь различные альтернативные формы. Многие митохондриальные геномы растений могут иметь один или несколько пар повторов, выступающих в качестве сайтов для межмолекулярной или внутримолекулярной рекомбинации. Например, анализ митохондриальной ДНК *L. sativa* показал, что мтДНК может принимать линейную, разветвленную и кольцевую структуру [13]. Различие форм мтДНК, а также её рекомбинацию, подтверждают и другие исследования [17, 18]. Таким образом, мтДНК является динамичной системой с множеством различных форм.

Митохондриальная ДНК находится в матриксе митохондрий. Обычно она распределена по нескольким нуклеоидам — нуклеопротеиновым комплексам [15]. При этом один нуклеоид не обязательно должен содержать полный геном [21]. Аналогом нуклеоида в митохондриях является бактериальная хромосома. В нуклеоидах происходят процессы репликации и транскрипции [14], а также предполагается, что они являются сайтами процессинга рРНК и сборки рибосом [3].

Размер митохондриального генома цветковых растений довольно сильно отличается от размера митохондриального генома животных. Обычно размер митохондриального генома цветковых растений варьируется от 200 до 750 кб, но может достигать и 1556 кб у *Cucumis sativus* [10, 22]. При

этом, например, у человека его размер составляет 16,6 кб. Гены белков в митохондриальном геноме растений кодируют субъединицы комплексов цепей окислительного фосфорилирования, белки, участвующие в биогенезе этих комплексов, а также несколько рибосомных белков. Первым полностью секвенированным митохондриальным геномом растения был геном *Arabidopsis thaliana*. В геноме закодированы 32 белка, 3 рибосомальных рРНК (5S, 18S и 26S рРНК) и 22 тРНК [10].

Митохондриальные гены растений содержат некодирующие участки (интроны), находящиеся между кодирующими (экзонами). В процессе реализации генетического материала с каждого гена вместе с экзонами и интронами синтезируется пре-РНК. Затем интроны удаляются специальным белковым комплексом, а экзоны сшиваются в цельную молекулу иРНК. Описанный процесс называется сплайсингом. Между митохондриальными генами также находятся некодирующие участки. Часть этой ДНК может участвовать в экспрессии генов, играть роль регуляторных сигналов. Кроме интронов в митохондриальном геноме растений также присутствуют транспозоны — мобильные генетические элементы.

В данной работе изучаются митохондриальные гены:

- *atp1, atp4, atp6, atp8, atp9*, кодирующие субъединицы АТФ-синтазы;
- *nad, nad2, nad3, nad4, nad4L, nad5, nad6, nad7, nad9*, кодирующие субъединицы НАДН-дегидрогеназы.

### 1.3 Окислительное фосфорилирование

Окислительное фосфорилирование — это метаболический путь, в ходе которого энергия, образовавшаяся в ходе окисления органических веществ, запасается в виде молекул АТФ. В эукариотических клетках окислительное фосфорилирование протекает в митохондриях. В отличие от внешней мембраны, пропускающих большое количество низкомолекулярных соединений, внутренняя имеет низкую проницаемость. Благодаря этому, лишь некоторые молекулы могут попасть в матрикс митохондрии с помощью специальных белков-переносчиков. Избирательная проницаемость

внутренней мембраны позволяет разграничить метаболические процессы, которые протекают в цитозоле и матриксе митохондрии.

Окислительное фосфорилирование является последним этапом в процессе клеточного дыхания. Оно осуществляется посредством нескольких комплексов белков, расположенных во внутренней мембране. В эукариотических клетках их 5. Вместе они составляют электрон-транспортную цепь (ЭТЦ), по которой комплекс за комплексом перемещаются электроны, пока не достигнут конечного акцептора — кислорода. Электроны доставляются дегидрогеназами к универсальным акцепторам электронов — никотинамиднуклеотидам НАД<sup>+</sup> или НАДФ<sup>+</sup> и флавиннуклеотидам (ФМН или ФАД).

Кроме никотинамиднуклеотидов и флавиннуклеотидов в переносе электронов в дыхательной цепи также участвуют убихинон и железосодержащие белки (цитохромы и железо-серные белки). Процесс переноса электронов происходит следующим образом. Электроны от первичных доноров попадают на флавопротеины, убихинон и далее по цепи железо-серных белков и цитохромов в конце концов попадают на молекулярный кислород.

Основной смысл работы ЭТЦ заключается в том, что каждый белковый комплекс может перекачивать протоны из матрикса в межмембранное пространство. Белки, выполняющие такую функцию называются протонными помпами. Таким образом протоны накапливаются в межмембранном пространстве, что создает протонный градиент и разность электрических потенциалов. Впоследствии накопленная энергия используется АТФ-синтазой для синтеза АТФ из молекулы АДФ.

Перенос электронов вызывает градиент Н<sup>+</sup>. Этот электрохимический градиент называется протон-движущей силой. Таким образом в процессе переноса электронов по ЭТЦ высвобождается большое количество свободной энергии, которая используется для синтеза АТФ из АДФ и неорганического фосфата.

## 1.4 Хлоропласты

Хлоропласты — зеленые пластиды, органоиды фотосинтезирующих клеток эукариот. Они встречаются в клетках высших растений, водорослей и некоторых фотосинтезирующих простейших. Хлоропласты высших растений имеют овальную форму, их размер около 5 мкм. Как и митохондрии, хлоропласты имеют высокопроницаемую наружную мембрану и менее проницаемую внутреннюю. Эти мембраны разделены узким межмембранным пространством. Область, окруженная внутренней мембраной, называется стромой. Она является аналогом матрикса митохондрий. В строме хлоропласта локализованы тилакоиды — мембранные компартменты. Мембрана тилакоида окружает его просвет или, по-другому, люмен. Граной называется стопка тилакоидов. В хлоропласте может находиться от 10 до 100 гран. Более того, часто грани соединены между собой ламеллами или вытянутыми тилакоидами. Благодаря такому соединению, грани образуют в строме единое пространство.

Важными молекулами, участвующими во многих процессах жизнедеятельности живых организмов, являются пигменты. В зависимости от своего строения и, следовательно, спектра поглощения света, пигменты делятся на несколько групп. Хлоропласты содержат в себе множество различных пигментов. Среди них:

- Зеленые пигменты (хлорофиллы);
- Оранжевые пигменты (каротиноиды);
- Красные и синие пигменты (фикобилины).

В хлоропластах разных организмов содержатся разные пигменты. Например, хлорофилл а содержится в хлоропластах большинства организмов, способных осуществлять фотосинтез, хлорофилл b у высших растений и зеленых водорослей, хлорофилл c и d у бурых и красных водорослей соответственно, каротиноиды — у всех организмов, кроме прокариот, фикобилины — у красных водорослей.

Как и митохондрии, хлоропласты имеют собственную ДНК и рибосомы, размножаются в клетке делением, сходны по строению с цианобак-

териями. Кроме того, многие участки генома хлоропластов гомологичны участкам бактериальной хромосомы. Все это может говорить нам о происхождении хлоропластов. Считается, что хлоропласты, аналогично митохондриям, произошли от древних фотосинтезирующих организмов (бактерий), которые эндоцитировали по разным оценкам от 1 до 1,5 млрд. лет назад в древнюю клетку. Также исследования показывают, что хлоропласты произошли от прокариот и имеют общего предка с цианобактериями.

## 1.5 Геномы хлоропластов

Как и митохондрии, хлоропласты имеют свой собственный геном, распределенный по нуклеоидам. Нуклеоиды имеют сферическую или яйцевидную форму. Зрелый хлоропласт содержит около 20 нуклеоидов со средней плоидностью 4-5 пластидных геномов на нуклеоид [9]. В нуклеоидах локализованы многие факторы, участвующие в репликации, транскрипции, обработки транскриптов (сплайсинг и редактирование мРНК), организации и репарации хлоропластной ДНК [16].

Размер генома хлоропластов составляет около 150 кб (*Nicotiana tabacum* — 156 кб, *Arabidopsis* — 154 кб). В геноме закодировано около 100 белок-кодирующих генов, а также гены рРНК и тРНК [12, 23]. Согласно данным, полученным на гель-электрофорезе в пульсирующем поле, большая часть хлДНК принимает линейные или разветвленные формы и лишь небольшая часть существует в виде кольцевых молекул. При этом разветвленные формы, по-видимому, имеют сайты для осуществления генетической рекомбинации.

Существуют доказательства, основанные на ряде исследований, показывающие, что в большинстве хлоропластов нуклеоид связан с тилакоидными мембранами в области инвертированного повтора [2, 11].

## 1.6 Фотосинтез

Фотосинтез — процесс, в результате которого энергия видимого света превращается в энергию химических связей органических веществ. Фото-

синтез могут осуществлять растения и некоторые бактерии. В отличие от животных клеток, растительные клетки имеют:

- жесткую клеточную стенку;
- вакуоли, выполняющие функции накопления запасных веществ и продуктов метаболизма, а также поддержание тургора;
- хлоропласты.

Именно благодаря хлоропластам в растительных клетках могут происходить биохимические реакции фотосинтеза. Хлоропласты используют энергию солнечного света для синтеза глюкозы и кислорода из углекислого газа и воды. Кроме прочего, в хлоропластах содержатся различные пигменты. Два из них играют центральную роль в световой фазе фотосинтеза: хлорофилл а и хлорофилл b. Разные пигменты поглощают свет определенной длины волны, имея при этом максимум поглощения. Так, максимальные длины волн поглощения для хлорофилла а и хлорофилла b лежат красной и сине-фиолетовой области соответственно. Хлорофиллы отражают зеленый свет, что придает характерную окраску растениям.

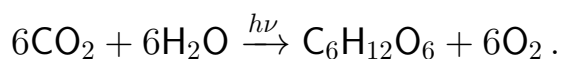
Хлорофиллы имеют следующую структуру: к порфириновому кольцу с помощью сложноэфирной связи присоединён длинный углеводородный хвост. Порфириновое кольцо состоит из 4 колец, образующих плоскую квадратную структуру. Каждое из колец содержит атом азота, благодаря которому возможно взаимодействие с атомом магния. Единственный атом магния локализован в середине порфиринового кольца. Разные хлорофиллы имеют разные боковые цепи. Именно это влияет на спектры их поглощения.

Процесс фотосинтеза разделяют на две фазы: световую и темновую. Световая фаза проходит в гранах хлоропластов, для неё необходим солнечный свет. Его поглощают молекулы хлорофилла, которые вместе с определенными белками входят в состав светособирающего комплекса. Этот комплекс называется фотосистемой I. Поглощение света приводит к тому, что электроны хлорофилла переходят в возбужденное состояние. Затем энергия электронного возбуждения передается в реакционный центр фо-

тосистемы I. Электрон проходит по цепи компонентов фотосистемы I и попадает к акцептору — НАДФ<sup>+</sup>. В результате образуется НАДФН. Положительно заряженный хлорофилл фотосистемы I получает электрон от фотосистемы II, электроны которой также возбуждаются солнечным светом. Благодаря фотолизу воды (разложение на протоны, электроны и кислород) восполняется нехватка электронов хлорофилла. Электроны попадают в электрон-транспортную цепь, аналогичную электрон-транспортной системе митохондрий. Возникает протонный градиент, который используется для синтеза АТФ. Таким образом, по завершении световой фазы образуются следующие продукты: молекулы АТФ, НАДФН и кислород. Первые два вещества необходимы для протекания темновой фазы, а последний является побочным продуктом.

В темновую фазу с помощью АТФ и НАДФН образуются простые сахара (глюкоза). Этот процесс имеет циклический характер и называется циклом Кальвина. Ключевую роль в нем играет крайне важный фермент — рубиско (рибулозо-1,5-дифосфаткарбоксилаза). Он катализирует присоединение молекул СО<sub>2</sub> к производному рибозы — рибулозо-1,5-бифосфату — на начальном этапе цикла Кальвина. Считается, что рубиско — наиболее распространенный фермент на планете Земля. Молекулы углекислого газа поступают в строму хлоропластов из воздуха через устьица растения. На второй этап цикла Кальвина затрачивается энергия в форме молекул АТФ и восстановительная сила молекул НАДФН. Третий этап цикла — восстановление рибулозо-1,5-бифосфата. После этого он снова участвует в фиксации углекислого газа. Таким образом, в результате темновой фазы образуются следующие продукты: глюкоза, НАДФ<sup>+</sup>, АДФ, Ф и вода. Молекулы НАДФ<sup>+</sup>, АДФ, Ф снова используются в световой фазе.

Резюмируя, в результате фотосинтеза из молекул воды и углекислого газа под действием света образуются молекула глюкозы и кислорода. Упрощенное уравнение фотосинтеза выглядит следующим образом:





## 1.7 Методы кластерного анализа

Для изучения различного рода данных необходимо выделить их структурированность. Для этого используются многочисленные методы кластерного анализа. В настоящий момент не существует универсального метода кластеризации. Их необходимо подбирать исходя из задач, которые нужно решить и природы данных. При этом каждый метод кластеризации имеет как достоинства, так и недостатки. Сейчас не существует единственной полной классификации методов кластеризации. Часто тот или иной метод можно отнести к нескольким группам. Далее приведем одну из классификаций методов кластеризации.

Первая группа методов кластеризации — иерархические алгоритмы. Они основаны на иерархическом выделении кластеров и как правило представляются в виде графов или дендрограмм. Среди иерархических методов выделяют две разновидности: дивизивные и агломеративные методы. Дивизивные методы основаны на выделении все более мелких кластеров из более крупных. Так, на первом шаге все множество точек образует единственный большой кластер. Затем наиболее непохожие данные разделяются по разным кластерам до тех пор, пока каждая точка не станет отдельным кластером. Кластеризация данных до этого момента не имеет смысла, поэтому встает вопрос: на каком шаге необходимо остановить кластеризацию? Универсального ответа на этот вопрос нет: он зависит от того какие данные берутся в рассмотрение, какие задачи необходимо решить и какие результаты получить. Агломеративные же методы работают в обратном порядке. Сначала каждая точка является отдельным кластером. Затем эти точки объединяются в большие кластеры на основе признаков, делающих их похожими. Примерами таких алгоритмов могут быть филогенетические деревья, построенные на основе сравнения генетических последовательностей разных организмов.

Вторая группа — статистические алгоритмы. Они основаны на том, что существует некая вероятность тех или иных данных попасть в тот или иной кластер. Другими словами, состав кластера должен описывать-

Таблица 1 – Некоторые метрики.

Название	Определение
Линейное расстояние	$d(I^{(i)}, I^{(j)}) = \sum_{k=1}^m  I_k^{(i)} - I_k^{(j)} $
Евклидово расстояние	$d(I^{(i)}, I^{(j)}) = \sqrt{\sum_{k=1}^m (I_k^{(i)} - I_k^{(j)})^2}$
Квадрат евклидова расстояния	$d(I^{(i)}, I^{(j)}) = \sum_{k=1}^m (I_k^{(i)} - I_k^{(j)})^2$
Обобщенное степенное расстояние Минковского	$d(I^{(i)}, I^{(j)}) = \left[ \sum_{k=1}^m (I_k^{(i)} - I_k^{(j)})^p \right]^{\frac{1}{p}}$

ся некоторым вероятностным распределением, а идея методов состоит в том, чтобы оптимизировать функцию этого распределения. Примеры статистических алгоритмов кластеризации: метод К-средних, метод К-медиан, дискриминантный анализ и т.д.

Очень часто результат кластеризации может меняться при изъятии некоторого количества точек. Можно подобрать такой набор данных, что изъятие единственной точки результат кластеризации будет значительно отличаться от первоначального. Кроме этого, существуют такие методы кластеризации, которые при каждом новом запуске будут выдавать разный результат. Поэтому часто необходимо вводить такое понятие как устойчивость кластеризации — долю тех точек, которые всегда попадают в один и тот же кластер. Также результат кластеризации может зависеть от того, какая выбрана функция расстояния между объектами. В Таблице 1 приведены некоторые из них [4].

В данной работе методы кластерного анализа использовались для выделения групп похожих по структуре генов, а также для выявления признака, по которому они попадают в один и тот же кластер: по принадлежности к одному и тому же гену, органелле или к близкому таксону. Для такого выделения групп в качестве методов кластеризации были использованы метод динамических ядер (для линейной кластеризации) и метод упругих карт (для нелинейной кластеризации). Описание их алгоритмов см. в Главе 2.

## 2. Материалы и методы

### 2.1 Генетический материал

Генетический материал был взят из открытой базы данных, NCBI GenBank<sup>1</sup>. В качестве исходного генетического материалы использовались полногеномные последовательности органелл растений. Последовательности были выбраны таким образом, чтобы для каждого организма были депонированы геномы обеих органелл одновременно. С помощью программы CLC Genomic Workbench из исходных последовательностей были выделены последовательности изучаемых генов. Эти гены приведены в Таблице 2.

Таблица 2 – Список исследуемых генов.

НАДН-дегидрогеназа		АТФ-синтаза	
Хлоропластные	Митохондриальные	Хлоропластные	Митохондриальные
<i>ndhA</i>	<i>nad1</i>	<i>atpA</i>	<i>atp1</i>
<i>ndhB</i>	<i>nad2</i>	<i>atpB</i>	<i>atp4</i>
<i>ndhC</i>	<i>nad3</i>	<i>atpE</i>	<i>atp6</i>
<i>ndhD</i>	<i>nad4</i>	<i>atpF</i>	<i>atp8</i>
<i>ndhE</i>	<i>nad4L</i>	<i>atpH</i>	<i>atp9</i>
<i>ndhF</i>	<i>nad5</i>	<i>atpI</i>	
<i>ndhG</i>	<i>nad6</i>		
<i>ndhH</i>	<i>nad7</i>		
<i>ndhI</i>	<i>nad9</i>		
<i>ndhJ</i>			
<i>ndhK</i>			

Выделенные последовательности представляли собой последовательности генов, содержащие и интроны, и экзоны. Выбор именно такого варианта последовательностей, а не, например, последовательностей CDS, объясняется тем, что последовательности находятся в геноме именно в таком виде. Эти и так консервативные последовательности являются более вариабельными по сравнению с белковыми последовательностями или по-

<sup>1</sup><https://www.ncbi.nlm.nih.gov>

следовательностями CDS. Это позволяет учитывать генетические различия близких видов.

Исследуемые организмы являются представителями семенных растений *Spermatophyta*. Всего было выбрано 84 растения, у которых были депонированы оба генома: и митохондриальный, и хлоропластный. Среди них 83 растения являются представителями отдела *Magnoliophyta* или Цветковых растений, 1 — представитель отдела *Gymnospermae* или Голосеменные растения. Распределение семейств представлено на Рисунке 1.

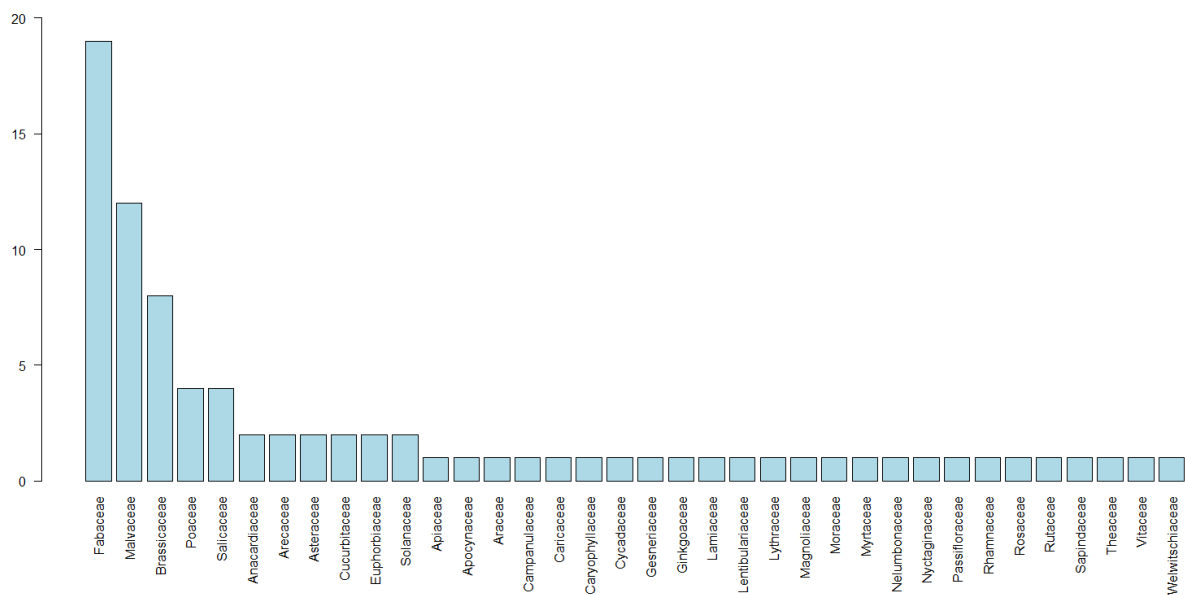


Рисунок 1 – Распределение исследуемых организмов по семействам.

## 2.2 Частотные словари

Целью данной работы является выявление связей и закономерностей между структурой функцией и таксономией. Как правило в подобных исследованиях четко определяется функция, кодируемая последовательностью, и таксономическое положение носителя этой последовательности. Функция нуклеотидной последовательности определяется функцией продукта, экспрессируемого с нее. Это может быть как РНК, так и белок. Таксономическое положение — это положение организма в иерархической системе, основанной на общих признаках или общем родстве. Однако, структура нуклеотидной последовательности может истолковываться по-

разному. В настоящей работе под структурой будем понимать частотный словарь триплетов, построенный на последовательностях генов.

Любая нуклеотидная последовательность — это символьная последовательность из алфавита, содержащего 4 буквы:  $\aleph = \{A, C, G, T\}$ . Все исследуемые последовательности разделяются на подпоследовательности длиной  $q$  следующим образом. Рамка считывания, длиной  $q$  помещается на последовательность и перемещается по ней с шагом  $t \geq 1$ . При этом ведется список всех попавших в рамку считывания  $q$ -плетов  $\omega$ , а также подсчитывается количество их копий  $n_\omega$ . Частота конкретного  $q$ -плета рассчитывается следующим образом:

$$f_\omega = \frac{n_\omega}{M}, \quad (1)$$

где  $M$  — количество всех встретившихся триплетов. Все возможные подпоследовательности (их количество равно  $4^q$ ), а также информация о частоте их встречаемости в исходной последовательности называется частотным словарем  $W_{(q,t)}$ . Далее в работе исследовался частотный словарь триплетов  $W_{(3,1)}$ .

Частотные словари являются точками в многомерном пространстве размерности 64, в котором координатами являются частоты триплетов. Для кластеризации частотных словарей в этом пространстве необходимо задать метрику расстояния. В данной работе использовалась Евклидова метрика, расстояние в которой определяется следующим образом:

$$d(W^{(1)}, W^{(2)}) = \sqrt{\sum_{\omega=AAA}^{\Gamma\Gamma\Gamma} (f_{\omega^{(1)}} - f_{\omega^{(2)}})^2} \quad (2)$$

Здесь  $W^{(1)}$  и  $W^{(2)}$  — частотные словари,  $f_{\omega^{(1)}}$  и  $f_{\omega^{(2)}}$  — частоты триплета  $\omega$  в первом и во втором частотном словаре, соответственно.

### 2.3 Метод динамических ядер

Метод динамических ядер (также известный как метод  $K$ -средних или  $K$ -means) является одним из наиболее известных методов кластеризации без учителя. Это итерационный метод, основанный на переопределении

центров масс кластеров, полученных на предыдущем шаге. Для использования этого метода нужно выбрать число классов, на которые следует поделить множество исследуемых точек; пусть оно будет равно  $K$ . На первом шаге в пространстве данных случайным образом определяется  $K$  начальных точек, представляющих собой динамические ядра классов. Динамическое ядро класса — центр масс всех точек, попавших в определенный класс. Затем определяется принадлежность каждой точки к какому-либо классу. Вычисляется расстояние от каждой точки до каждого динамического ядра, а затем переопределяется принадлежность точек к классу: каждая точка должна оказаться в том классе, к динамическому ядру которого она ближе всего. Затем на полученной конфигурации данных перевычисляются динамические ядра кластеров и все повторяется до тех пор, пока состав кластеров не будет изменяться.

Важно отметить, что метод динамических ядер не всегда останавливается за конечное число шагов. В то же время, реализация метода динамических ядер на одних и тех же данных не всегда приводит к одному и тому же результату. Это происходит из-за того, что начальное разбиение на кластеры осуществляется случайным образом. Если же разделение точек по классам всегда принимает один и тот же вид независимо от начального распределения точек, то такая классификация является устойчивой, но в реальных данных встречается это крайне редко. Поэтому важно ответить на вопрос об устойчивости классификации. Для этого необходимо провести достаточно длинную серию реализаций метода и рассчитать долю тех точек, которые попадают в один и тот же кластер. Изучение точек, меняющих свою принадлежность к тому или иному классу при проведении кластеризации методом динамических ядер, может иметь важное значение для исследования. В таком неустойчивом поведении точек могут наблюдаться свои закономерности, которые, возможно, будут иметь свое биологическое значение.

Строго говоря, после получения классификации стоит проверить различимость классов. Существуют следующие критерии различимости клас-

сов. Сильный: для различимости любых двух классов сумма их радиусов должна быть не меньше расстояния между их центрами. Слабый: радиус большего класса должен быть не больше расстояния между классами.

## 2.4 Метод упругих карт

Для выполнения задач, поставленных в этом исследовании, использовались два метода кластеризации. Первый метод — метод динамических ядер — описан в Разделе выше. Кроме него был использован второй метод кластеризации — метод упругих карт. В отличие от метода динамических ядер, который является линейным методом кластеризации, метод упругих карт — нелинейный метод. Он используется для визуализации и кластеризации многомерных данных. Суть метода заключается в том, чтобы уменьшить размерность пространства, потеряв как можно меньше информации. Это достигается тем, что данные проецируются на не слишком изогнутую поверхность и отображаются на ней как на карте. Метод упругих карт является похож на метод главных компонент. Отличие заключается в том, что в первом используется упругая поверхность, а во втором — абсолютно жесткая плоскость [6].

Упругие карты строятся следующим образом. Как и при использовании метода главных компонент, прежде всего необходимо определить две первые главные компоненты. Первая главная компонента показывает направление в пространстве, вдоль которого наблюдается максимальная выборочная дисперсия данных. Вторая главная компонента — это направление, вдоль которого наблюдается следующая по величине выборочная дисперсия. Важно, что вторая главная компонента должна быть перпендикулярна первой. Выбор главных компонент сводится к нахождению собственных векторов и собственных значений ковариационной матрицы исследуемых данных.

Затем данные проецируются на плоскость, которую образуют две первые главные компоненты. Каждая точка соединяется со своей проекцией математической пружиной. Она обладает бесконечной растяжимостью.

После чего на плоскости с проекциями определяется квадрат минимального размера, в который должны попасть все проекции. Стороны этого квадрата должны быть параллельны главным компонентам. Поверхность квадрата заменяется гибкой и растяжимой мембраной.

На следующем шаге система отпускается и релаксирует до состояния с минимумом энергии. Важно отметить, что при этом необходимо сохранить топологию: поверхность квадрата нельзя разрывать и склеивать. Затем определяются ортогональные проекции для каждой точки на деформированной поверхности. После освобождения системы от пружин, она снова становится квадратом или упругой картой во внутренних координатах [6].

Если данные имеют какую-либо структурированность, часть из них на упругой карте будет образовывать скопления. Для выделения этих скоплений используется локальная плотность точек. Для определения локальной плотности на упругой карте каждой точке сопоставляется какая-либо куполообразная функция распределения с максимумом в этой точке, например, функция Гаусса [5]

$$f_j(r) = A \cdot \exp \left\{ -\frac{(r - r_j)^2}{\sigma^2} \right\}. \quad (3)$$

Здесь  $r_j$  — координата  $j$ -ой точки,  $A$  — множитель, одинаковый для всех точек,  $\sigma$  — полуширина этой функции, полностью аналогичная стандартному отклонению для случая нормального распределения случайной величины, подгоночный параметр, определяющий контрастность картины локальной плотности. Затем производится суммирование всех функций

$$F(r) = \sum_{j=1}^N \exp \left\{ -\frac{(r - r_j)^2}{\sigma^2} \right\}. \quad (4)$$

Здесь  $N$  — количество точек. Метод упругих карт и метод динамических ядер реализованы в свободно распространяемом программном обеспечении *ViDaExpert*<sup>2</sup> [6].

---

<sup>2</sup><http://bioinfo-out.curie.fr/vidaexpert/>



### **3. Результаты и их обсуждения**

В связи с авторскими правами была изъята 21 страница.

## ЗАКЛЮЧЕНИЕ

В ходе выполнения данной работы были выполнены все задачи, а поставленная цель достигнута:

1. Были построены частотные словари генов;
2. С помощью программы *ViDaExpert* была проведена кластеризация частотных словарей методом упругих карт;
3. Была проведена кластеризация с помощью метода динамических ядер, построены графы перехода точек между кластерами при увеличении числа кластеров  $k$ .
4. Был проведен анализ состава кластеров с точки зрения функции и таксономии.

Анализ результатов данной работы показал, что в кластерах находятся словари, принадлежащие одному и тому же гену. Точнее,

- гены субъединиц АТФ-синтаз и НАДН-дегидрогеназ митохондрий растений чётко кластеризуются, при этом кластеры включают в себя гены субъединиц, а не таксоны;
- гены субъединиц АТФ-синтаз и НАДН-дегидрогеназ хлоропластов растений чётко кластеризуются, при этом кластеры включают в себя гены субъединиц, а не таксоны.

Таким образом было показано преобладание функции, кодируемой исследованными последовательностями, над таксономией их носителей двумя методами кластеризации.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Florian J Bock and Stephen WG Tait. Mitochondria as multifaceted regulators of cell death. *Nature reviews Molecular cell biology*, 21(2):85–100, 2020.
2. Stephen A Boffey, Stephen A Boffey, and David Lloyd. *Division and segregation of organelles*, volume 35. Cambridge University Press, 1988.
3. Alexandra-Viola Bohne. The nucleoid as a site of rRNA processing and ribosome assembly. *Frontiers in Plant Science*, 5:257, 2014.
4. Michel Marie Deza and Elena Deza. Encyclopedia of distances. In *Encyclopedia of distances*, pages 1–583. Springer, 2009.
5. K. Fukunaga. *Introduction to statistical pattern recognition*. Academic Press, London, 1990.
6. Alexander N. Gorban and Andrei Zinovyev. Principal manifolds and graphs in practice: From molecular biology to dynamical systems. *International Journal of Neural Systems*, 20(03):219–232, 2010. PMID: 20556849.
7. AN Gorban, TG Popova, and MG Sadovsky. Classification of symbol sequences over their frequency dictionaries: towards the connection between structure and natural taxonomy. *Open Systems & Information Dynamics*, 7(1):1–17, 2000.
8. A.N. Gorban, T.G. Popova, M.G. Sadovsky, and D.C. Wunsch. Information content of the frequency dictionaries, reconstruction, transformation and classification of dictionaries and genetic texts. In *Intelligent Engineering Systems Through Artificial Neural Networks*, pages 657–663. American Society of Mechanical Engineers (ASME), 2001.
9. Stephan Greiner, Hieronim Golczyk, Irina Malinova, Tommaso Pellizzer, Ralph Bock, Thomas Börner, and Reinhold G Herrmann. Chloroplast

- nucleoids are highly dynamic in ploidy, number, and structure during angiosperm leaf development. *The Plant Journal*, 102(4):730–746, 2020.
10. José M Gualberto, Daria Milesina, Clémentine Wallet, Adnan Khan Niazi, Frédérique Weber-Lotfi, and André Dietrich. The plant mitochondrial genome: dynamics and maintenance. *Biochimie*, 100:107–120, 2014.
  11. Sun Yong Jeong, Annkatrin Rose, and Iris Meier. MFP1 is a thylakoid-associated, nucleoid-binding protein with a coiled-coil structure. *Nucleic Acids Research*, 31(17):5175–5185, 2003.
  12. Samir Kaul, Hean L Koo, Jennifer Jenkins, Michael Rizzo, Timothy Rooney, Luke J Tallon, Tamara Feldblyum, William Nierman, Maria-Ines Benito, Xiaoying Lin, et al. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 408(6814):796–815, 2000.
  13. Alexander Kozik, Beth A Rowan, Dean Lavelle, Lidija Berke, M Eric Schranz, Richard W Michelson, and Alan C Christensen. The alternative reality of plant mitochondrial DNA: One ring does not rule them all. *PLoS genetics*, 15(8):e1008373, 2019.
  14. Karin Krupinska, Joanna Melonek, and Kirsten Krause. New insights into plastid nucleoid structure and functionality. *Planta*, 237(3):653–664, 2013.
  15. Martin Kucej and Ronald A Butow. Evolutionary tinkering with mitochondrial nucleoids. *Trends in cell biology*, 17(12):586–592, 2007.
  16. Wojciech Majeran, Giulia Friso, Yukari Asakura, Xian Qu, Mingshu Huang, Lalit Ponnala, Kenneth P Watkins, Alice Barkan, and Klaas J van Wijk. Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant physiology*, 158(1):156–189, 2012.
  17. Delene J Oldenburg and Arnold J Bendich. Size and structure of replicating mitochondrial DNA in cultured tobacco cells. *The Plant Cell*, 8(3):447–461, 1996.


18. Delene J Oldenburg and Arnold J Bendich. The structure of mitochondrial DNA from the liverwort, *marchantia polymorpha*. *Journal of molecular biology*, 276(4):745–758, 1998.
19. Michael Sadosky, Yulia Putintseva, Anna Chernyshova, and Vaselina Fedotova. Genome structure of organelles strongly relates to taxonomy of bearers. In *International Conference on Bioinformatics and Biomedical Engineering*, pages 481–490. Springer, 2015.
20. Michael G Sadosky, Natalia A Zaitseva, and Yulia A Putintseva. System biology on mitochondrion genomes. In *The Third International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies*, pages 61–66, 2011.
21. M. Satoh, Y Nemoto, S. Kawano, T. Nagata, H. Hirokawa, and T. Kuroiwa. Organization of heterogeneous mitochondrial DNA molecules in mitochondrial nuclei of cultured tobacco cells. *Protoplasma*, 175:112–120, 1993.
22. Daniel B Sloan. One ring to rule them all? genome sequencing provides new insights into the ‘master circle’ model of plant mitochondrial DNA structure. *New Phytologist*, 200(4):978–985, 2013.
23. Masahiro Sugiura. The chloroplast genome. *10 Years plant molecular biology*, pages 149–168, 1992.
24. Б Альбертс, Д Брей, К Хопкин, А Джонсон, Дж Льюис, М Рэфф, К Робертс, and П Уолтер. *Основы молекулярной биологии клетки*. М.: Бином, 2015.
25. Б. Б. Бадмаев. Сравнение таксономической структуры кормовых растений двух видов наземных беличьих Западного Забайкалья и степной флоры региона. *Вестник Красноярского государственного аграрного университета*, (9), 2009.

26. М.Г. Садовский and А.И. Чернышова. Выявление связи структуры и таксономии геномов хлоропластов методом динамических ядер. *Фундаментальные исследования*, 3(11), 2014.
27. М.Ю. Сенашова and М.Г. Садовский. Семикластерная структура геномов хлоропластов отражает филогению их носителей. *Международный журнал прикладных и фундаментальных исследований*, 12(7):1167–1173, 2016.

Министерство науки и высшего образования РФ  
Федеральное государственное автономное  
образовательное учреждение высшего образования  
**«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»**  
Институт фундаментальной биологии и биотехнологии  
Кафедра геномики и биоинформатики

УТВЕРЖДАЮ:

Заведующий кафедрой

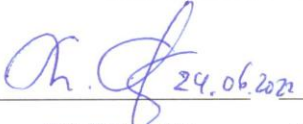
 И. Е. Ямских  
«24» июня 2022 г.

**МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ**

«Выявление связей между структурой, функцией и таксономией в  
геномах митохондрий и хлоропластов»

06.04.01 Биология

06.04.01.06 Геномика и биоинформатика

Руководитель  24.06.2022 проф., д.ф.-м.н. М.Г. Садовский  
подпись, дата должность, ученая степень

Выпускник  24.06.2022 В.Д. Федотовская  
подпись, дата

Рецензент  24.06.2022 проф., д.ф.-м.н. С.И. Барцев  
подпись, дата должность, ученая степень

Красноярск 2022