

относительно референсного генома; данные по экспрессии представляются в табличном формате. Пайплайн (набор последовательно запускаемых алгоритмов) по анализу этих данных, выделению генов с измененной экспрессией, а также с наиболее интересными мутациями, их отображение на известные структуры белков будет реализован с помощью технологии Common Workflow Language (<http://www.commonwl.org/>). Это позволит сделать анализ полностью переносимым и воспроизводимым другими исследователями. Основные части пайплайна будут реализованы на языке python с применением библиотеки biopython[94].

=Методы оценки влияния мутаций на стабильность комплексов (для задачи 4.1.)=

Оценка влияния мутаций на стабильность комплексов будет осуществляться с помощью расчетов изменения свободной энергии связывания программой FOLDX [95]. Данная программа способна быстро оценивать свободную энергию на основании эмпирических силовых полей применяемых к атомистическим структурам биомолекул.

==== План работ и ожидаемые результаты ====

Диаграмма работ приведена на Рисунке 10.

=Пакет задач 1. Моделирование внутренней пластичности октамера гистонов в нуклеосоме. (Годы 1-3.)=

Задача 1.1. Построить модели конформационных перестроек димеров H3-H4 в нуклеосомах необходимых для взаимодействий с ремоделерами SNF2h и SWI/SNF. - Годы 1-2

Ожидаемые результаты (Год 1): С помощью методов атомистической МД (включая метадинамику, управляемую динамику и иные “продвинутое” методы) будет изучена внутренняя динамика димера гистонов H3-H4 в контексте имеющихся экспериментальных данных.

Ожидаемые результаты (Год 2): Будут построены модели пластичности нуклеосомы (включая ДНК) согласующиеся с динамикой димера H3-H4.

Задача 1.2. Установить влияние различия последовательности гистонов H2A и его варианта H2A.Z на конформацию и динамику H2A-H2B димеров и построить комплексную модель, объясняющую селективность связывания ремоделера Swr1 у дрожжей (аналог р400 и SRCAP человека) с нуклеосомами содержащими гистон H2A. - Год 2-3.

Ожидаемые результаты (Год 2): Проведено сравнительное моделирование димеров гистонов H2A/H2B и H2A.Z/H2B пекарских дрожжей методами атомистической молекулярной динамики, включая “продвинутое” методы МД. Установлен характер влияния аминокислотных различий на динамику.

Ожидаемые результаты (Год 3): Разработаны методы вычисления вероятности дейтеро-водородного обмена по данным динамики. Получены профили динамики и вероятности дейтеро-водородной обмена вдоль последовательности гистонов для сравнения с возможными экспериментальными данными.

=Пакет задач 2. Моделирование и анализ взаимодействий нуклеосом с пептидами. (Годы 1-3)=

Задача 2.1. Провести структурный и энергетический анализ известных взаимодействий пептидов/мотивов белков с кислотным лоскутом нуклеосомы (включая пептид LANA, белок CENP-C, антитело PL2-6). - Годы 1-2

Ожидаемые результаты (Год 1): Проведена молекулярная динамика комплексов пептидов с нуклеосомами, оценена динамика взаимодействий различных участков пептидов, построены карты контактов.

Ожидаемые результаты (Год 2): Оценена энергия взаимодействия различных известных пептидов с нуклеосомами, проведен *ab initio* докинг известных пептидов для оценки воспроизводимости результатов и качества скоринговых функций.

Задача 2.2. Провести дизайн искусственных пептидов с высокой аффинностью связывания с кислотным лоскутом. - Годы 2-3

Ожидаемые результаты (Год 2): Проведен дизайн пептидов на основе оценки изменения энергии взаимодействия известных пептидов при введении точечных мутаций в пептиды с целью оптимизации энергии связывания.

Ожидаемые результаты (Год 3): Проведен дизайн пептидов селективных к нуклеосомам с различными гистоновыми вариантами.

=Пакет задач 3. Биоинформатический анализ интерактома нуклеосом, разработка базы данных по взаимодействиям нуклеосом. (Годы 1-3).=

Задача 3.1. Анализ и классификация всех имеющихся в открытом доступе данных по взаимодействию нуклеосом с белками хроматина у человека (Годы 1-2).

Ожидаемые результаты (Год 1): Реализован программный пайплайн, который в полуавтоматическом режиме загружает и анализирует из открытых баз данных информацию о взаимодействиях нуклеосом с белками хроматина.

Ожидаемые результаты (Год 2): Проведена полуавтоматическая курация, аннотация и классификация собранного набора данных.

Задача 3.2. Разработать базу данных и веб-ресурс, представляющие в интерактивном виде информацию о известных взаимодействиях нуклеосом с белками хроматина, включая информацию по имеющимся трехмерным структурам. (Годы 2-3).

Ожидаемые результаты (Год 2): Разработан дизайн структуры базы данных, создана тестовая реализация на основе СУБД (Системы управления базы данных).

Ожидаемые результаты (Год 3): Реализован интерактивный интерфейс к базе данных.

=Пакет задач 4. Пакет задач 4. Биоинформатический анализ геномных и транскриптомных данных опухолей с точки зрения организации хроматина на нуклеосомном уровне (Годы 2-3)=

Задача 4.1. Проанализировать белки, взаимодействующие с гистонами и нуклеосомами, на предмет наличия в них повторяющихся (более чем у одного пациента) мутаций в образцах раковых опухолей по наборам данных международного консорциума раковых геномов и атласа раковых геномов (ICGC, TCGA). Провести структурную интерпретацию этих мутаций с расчетом их влияния на стабильность комплексов (при наличии соответствующий структур). -( Годы 2-3)

Ожидаемые результаты (Год 2): Список белков взаимодействующих с гистонами и нуклеосомами (включая сами гистоны), для которых обнаружены статистически значимые мутации в образцах раковых опухолей.

Ожидаемые результаты (Год 3): Анализ влияния мутаций в образцах раковых опухолей на стабильность и структуру комплексов нуклеосом и гистонов с белками хроматина.

Задача 4.2. Проанализировать белки, взаимодействующие с гистонами и нуклеосомами, на предмет аномалий экспрессии в образцах раковых опухолей по наборам данных международного консорциума раковых геномов и атласа раковых геномов (ICGC, TCGA). Провести интерпретацию влияния повышенной или пониженной концентрации белка на активность ассоциации/диссоциации известных структурных комплексов и возможные эффекты в работе хроматина. - (Год 3)

Ожидаемые результаты (Год 3): Список белков взаимодействующих с гистонами и нуклеосомами (включая сами гистоны), для которых обнаружены статистически значимые отклонения в уровнях экспрессии в образцах раковых опухолей по сравнению с нормальными тканями. Интерпретация влияния отклонений в экспрессии на активность ассоциации/диссоциации известных структурных комплексов и возможные эффекты в работе хроматина.

=== Ожидаемые результаты по публикациям ==

Ожидаемые результаты будут обнародованы в журналах с высоким импакт фактором в количестве не менее 8 единиц. Целевые журналы, для которых масштаб всех заявленных задач является адекватным для публикации в качестве отдельного исследования по нашему мнению следующие: задачи пакетов 1, 2 и 3 - Nucleic Acids Research (ИФ ~9), Journal of Molecular Biology(ИФ ~4), Epigenetics&Chromatin (ИФ ~5), Journal of Biomolecular Structure and Dynamics (ИФ ~4); задач пакета 3 - Nucleic Acids Research, Database Issue (ИФ ~9), Database (ИФ ~2). Задачи пакета 1 могут также войти как часть в совместные