

Семинары по иммуноинформатике

Антон С. Смирнов

October 20, 2025

Оглавление

Предисловие	1
Организация обучения	3
Цели и задачи	3
Ожидаемые результаты обучения	3
Приобретаемые знания	3
Приобретаемые навыки	4
Система оценивания	4
Учебные материалы	4
Основная литература	4
Лекции	7
I Медицина и иммуноинформатика	9
1 Интерпретация результатов NGS	11
1.1 Предварительная подготовка	11
1.2 Пример заключения	11
1.3 Пациент №1 (разминка)	13
1.4 Пациент №2	13
1.5 Пациент №3	15
1.6 Пациент №4 (на дом)	16
2 HLA-типирование	19
2.1 Генетические особенности	19
2.2 Биоинформатические особенности	19
II Анализ иммунных репертуаров	21
3 Анализ данных с помощью cellranger	23
4 Библиотека Seurat	25
III Компьютерное конструирование вакцин и антител	27
5 Проектирование мультиэпитопной вакцины	29
6 Защита проектов	31
7 Моделирование трехмерной структуры антитела	33
8 Моделирование взаимодействия антитело-антиген	35

9 Зачёт**37****References****39**

Предисловие

Материалы для семинаров для нового курса "Иммуноинформатика" для студентов 5 курса ФББ МГУ. Этот предмет включен в программу по настоятельной просьбе активных и мотивированных студентов факультета. Данный курс является краткой выжимкой курса иммуноинформатики, читаемого биологам-магистрам медико-биологического факультета Пироговского университета и рассчитан на 24 академических часа. Вследствие малого количества аудиторных часов, я рассчитываю на вашу активную самостоятельную работу. Предмет рассчитан на три модуля:

- Иммуноинформатика в медицине;
- Анализ данных иммуносеквенирования
- Компьютерное конструирование иммунопрепаратов.

Подглавы пособия соответствуют темам семинара, а главы - разделам, следовательно, 1 глава - 1 семинар. Ссылки на презентации лекций будут приложены отдельно к описаниям разделов в соответствующих главах.

Организация обучения

Цели и задачи

Цель курса – познакомить слушателя с источниками данных, рабочими инструментами в иммуноинформатике и принципами их работы. В курсе рассматриваются экспериментальные методы, применяемые в фундаментальной и клинической иммунологии вместе с методами анализа результатов экспериментов. Разбираются технические аспекты применения вычислительных методов. Особое внимание уделено вопросам компьютерного конструирования вакцин.

Ожидаемые результаты обучения

После изучения курса слушатель получит представление о применении методов биоинформатики в иммунологии и используемых и информационных ресурсах. Слушатель научится применять продвинутое технические инструменты для решения как иммуноинформатических, так и общих биоинформатических задач.

Приобретаемые знания

- Знание медицинской классификации генетических вариантов, алгоритма вынесения заключения по результатам высокопроизводительного секвенирования.
- Знание об алгоритмах, используемых в картировании прочтений локуса HLA.
- Понимание принципов работы программ-предсказателей
- Понимание особенностей открытых на момент преподавания курса иммунологических данных
- Знание о принципах и этапах компьютерного конструирования вакцин и иммунологических препаратов
- Понимание принципов анализа результатов single-cell RNA секвенирования.
- Понимание принципов моделирования трехмерных структур.

Приобретаемые навыки

- Навыки работы с общедоступными биоинформатическими базами данных
- Навыки работы с интерфейсом командной строки, работы в операционной системе Linux
- Навыки написания программ на языке программирования Python, R для решения иммуноинформатических задач
- Умение применять иммуноинформатические онлайн-инструменты, интерпретировать их результаты
- Способность решать иммуноинформатические задачи как прикладного (медицинского), так и фундаментального характера.

Система оценивания

Дисциплина предполагает получение «зачета». Для «зачета» необходимо достичь порогового показателя в 60% по каждому виду деятельности:

Виды деятельности	Итого, баллы	Пороговое значение, баллы	Комментарии
Работа на очных занятиях	70	42	7 занятий по 10 баллов
Индивидуальная домашняя работа	40	24	4 домашних задания по 10 баллов
Проект по конструированию вакцины	20	12	
Итоговое устное собеседование	30	18	
Итого	160	96	

Учебные материалы

Основная литература

1. Мерфи, К. Иммунобиология по Джанвэю / К. Мерфи, К. Уивер; пер. с англ.; под ред. Г.А. Игнатъевой, О.А. Свитич, И.Н. Дьякова. - М.: Логосфера, 2020. - 1184 с. : ил. : 21,3 см. ISBN 978-5-98657-070-9
2. Руководство по интерпретации данных последовательности ДНК человека, полученных методами массового параллельного секвенирования (MPS) (редакция 2018, версия 2) <https://ngs.med->

gen.ru/mgngs19/%D0%A0%D1%83%D0%BA%D0%BE%D0%B2%D0%BE%D0%B4%D1%81%D1%82%D0%B2%D0%BE%20NGS%202019%20%D0%B2%D0%B5%D1%80.2.pdf

3. Namrata Tomar. Immunoinformatics. Springer Protocols 3 ed., 2020 DOI: <https://doi.org/10.1007/978-1-0716-0389-5>
4. Pedro A. Reche. Computational Vaccine Design 1 ed., 2023 DOI: <https://doi.org/10.1007/978-1-0716-3239-0>
5. Курс «Анализ транскриптомных данных», Сергей Исаев. <https://github.com/serjisa/transcriptomics>.msu

Лекции

1. Интерпретация результатов NGS. HLA-типирование.

Part I

Медицина и иммуноинформатика

Chapter 1

Интерпретация результатов NGS

1.1 Предварительная подготовка

Вам потребуется для работы:

1. Интернет - для доступа к статьям, базам данных и UCSC Genome Browser
2. Геномный браузер igv. Скачать можно по ссылке

О Broad Institute

Поскольку британское правительство - нехорошие люди, институт блокирует российские IP-адреса. Проблема легко решаема самым простым VPN. С помощью VPN можно скачать самую последнюю версию IGV. Также могут возникнуть проблемы с загрузкой референсного генома, т.к. igv осуществляет их загрузку с серверов Amazon. Чтобы решить эту проблему, добавьте в скрипт запуска igv-launcher.bat (он лежит в папке, куда установили igv) опции, указанные ниже.

```
start %JAVA_CMD% ... --genomeServerURL=https://tools.epigenetic.ru/igv/genomes.txt --genome=https://tools.epigenetic.ru/igv/genomes.txt
```

3. DB Browser for SQLite

1.2 Пример заключения

Пример

1. Паспортная информация о документе
 - Название организации, контактная информация
 - Название лаборатории:

- Карта:
- Номер исследования:
- Фамилия:
- Дата рождения:
- Регион:
- Направившее ЛПУ:
- Направивший врач:
- Метод исследования: Массовое параллельное секвенирование; биоинформатический анализ ДНК
- Материал: Жидкая кровь с ЭДТА, не подвергавшаяся замораживанию
- Направляющий диагноз:
- Пункт преискуранта:

2. Находки

Table 1.1: Патогенные варианты нуклеотидной последовательности, являющиеся вероятной причиной заболевания

Ген	Положение (hg38)	Генотип	Экзон/ интрон	Положение в кДНК	Эффект	Частота аллеля*	Глубина Транскриптпрочтения
CD81	chr11:23955250	G>A	5	c.459+5G>A p.?		0.00001972	NM_004356.3 8

Table 1.2: Вероятно патогенные варианты нуклеотидной последовательности, являющиеся возможной причиной заболевания

Ген	Положение (hg38)	Генотип	Экзон/ интрон	Положение в кДНК	Эффект	Частота аллеля*	Глубина Транскриптпрочтения
IVD	chr15:4041544	C>T	9	c.932C>T	p.(Ala311Val)	0.00072930	NM_002225.4 10

Table 1.3: Варианты нуклеотидной последовательности неопределенного значения, имеющие возможное отношение к фенотипу

Ген	Положение (hg38)	Генотип	Экзон/ интрон	Положение в кДНК	Эффект	Частота аллеля*	Глубина Транскриптпрочтения

3. Интерпретация

4. Описание методики
5. Сведения о качестве исследования
6. Литература
7. Подписи специалиста и заведующего лабораторией

i Волшебные фразы

1. Данные секвенирования могут быть предоставлены по запросу лечащего врача.
2. Все варианты, указанные в заключении, необходимо подтверждать методом прямого секвенирования по Сенгеру.
3. Клиническое заключение по результатам данного исследования может быть дано только врачом-генетиком.
4. Результат анализа может быть интерпретирован только врачом.

1.3 Пациент №1 (разминка)

Данные + выписки

1.4 Пациент №2

Новорожденный мальчик с неонатального скрининга

Данные. Скрины из выписок представлены ниже

Анамнез жизни: Ребенок от 4 беременности (1 беременность - 2012 г., срочные роды; 2 - 2013 г., выкидыш; 3 - 2022 г., выкидыш), 2 самостоятельных родов на сроке 39-40 нед. Течение беременности: на фоне токсикоза в 1-м триместре, варикозного расширения вен, ОРВИ - 13 и 38 нед., ЖДА в 3-м триместре. Вес при рождении 3502 г. Рост 53 см. Закричал сразу. По шкале Апгар 7/9 баллов. К груди приложен сразу. БЦЖ-М, гепатит В - проведено в роддоме, далее медицинский отвод от профилактических прививок. Естественное вскармливание до 2 нед. Перенесенные заболевания: перелом ключицы при родовой травме; гипопаратиреоз.

Семейный анамнез: по онкологическим, гематологическим заболеваниям неотягощен.

Анамнез заболевания: На первом месяце жизни (08.02.23) - повторные эпизоды судорог, госпитализирован в ОРИТ по месту жительства, назначены антиконвульсанты и миоплегическая

терапия, пациент переведен на ИВЛ. По данным ЭЭГ - убедительных данных за истинную эпилептиформную активность нет, лабораторно - снижение ионизированного кальция до 0,8 ммоль/л, снижение паратгормона до 0,3 пг/мл. При исследовании ликвора - незначительное повышение белка (0,5 г/л), в остальном норма. В крови - виремия (ЦМВ до 2300 к/мл), при исследовании ликвора ЦМВ, ЭБВ, ВГЧ6, ВПГ1-2 не детектированы. Пациенту проводилась противосудорожная терапия, комбинированная противомикробная терапия (сульперазон, амикацин, ванкомицин, бисептол, флуконазол, ганцикловир), ВВИГ, коррекция гипокальциемии (препараты кальция вв и per os, оксидевит). На фоне проводимой терапии отмечена стабилизация состояния, пациент экстубирован 21.02.23.

	Результат	
	%	10 ⁶ /мл
WBC	-	9,42
Granulocytes	70,0	6,59
Monocytes	15,0	1,41
Lymphocytes	15,0	1,41
	%	кл/мкл
T-cells CD3+ Lym (% om Lym)	0,18	3
CD3+CD4+ Lym (% T cells)	67,0	2
T-naïve cells CD4+CD45RA+CD197+ (% om CD4)	0,0	0
T-central memory cells CD4+CD45RA-CD197+ (% om CD4)	0,0	0
Effector memory cells CD4+CD45RA-CD197- (% om CD4)	0,0	0
TEMRA CD4+CD45RA+CD197- (% om CD4)	0,0	0
CD3+CD8+ Lym (% T cells)	14,0	0
T-naïve cells CD8+CD45RA+CD197+ (% om CD8)	0,0	0
T-central memory cells CD8+CD45RA-CD197+ (% om CD8)	0,0	0
Effector memory cells CD8+CD45RA-CD197- (% om CD8)	0,0	0
TEMRA CD8+CD45RA+CD197- (% om CD8)	0,0	0
B-cells CD19+ Lym (% om Lym)	72,0	1017
NK-cells CD3-CD16+CD56+ Lym (% om Lym)	27,8	393
CD56+high NK (% om NK-cells)	1,0	4
T-NK-cells CD3+CD56+ Lym (% om Lym)	0,0	0

Название набора	ТК-SMA - Набор реагентов для выделения и количественного определения ДНК TREC, KREC и качественного выявления гомозиготной делеции экзона 7 гена SMN1 методом ПЦР-РВ
Амплификатор	ДТ-прайм (ДНК-Технология)

РЕЗУЛЬТАТЫ МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКОЙ ДИАГНОСТИКИ

	Показатель пробы	Интерпретация	Референсные значения
SMN1	21,5	Не обнаружена	менее 25 - не обнаружена
TREC	0	Патология	менее 100 - патология от 100 до 450 - сомнительный результат более 450 - норма
KREC	2028	Норма	менее 100 - патология от 100 до 200 - сомнительный результат более 200 - норма
<i>Комментарии.</i> <i>Результат лабораторного исследования не является диагнозом.</i> <i>Интерпретация результата проводится врачом с учетом клинических проявлений и данных анамнеза.</i>			

1.5 Пациент №3

Новорожденный мальчик с неонатального скрининга

Данные

Table 1.4: TREC/KREC

TREC	KREC	SMN1
3.1 [2.0-3.5]	1.1 [2.0-3.5]	N

Table 1.5: Субпопуляции лимфоцитов

Показатель	Результат	Ед. изм.
WBC	8.54	10/мл
Гранулоциты, % от WBC	21.0	%
Гранулоциты	1.79	10/мл
Моноциты, % от WBC	11.0	%
MONO#	0.94	10/мл
LYM	5.81	10/мл
LYMPH%	68.0	%

Показатель	Результат	Ед. изм.
T-cells CD3+ Lym	5052	кл/мкл
CD3 (Т-лимфоциты), % от LYM	87.0	%
CD3+CD4+ (Th cells) (% от CD3)	63.7	%
CD3+CD4+ Lym	3218	кл/мкл
CD45RA+CD197+ (T-naïve cells)	2095	кл/мкл
CD45RA+CD197+ (T-naïve cells) (% от CD4)	65.1	%
CD45RA-CD197+ (T-central memory cells)	933	кл/мкл
CD45RA-CD197+ (T-central memory cells) (% от CD4)	29.0	%
CD45RA-CD197- (Effector memory cells)	103	кл/мкл
CD45RA-CD197- (Effector memory cells) (% от CD4)	3.2	%
CD45RA+CD197- (TEMRA)	90	кл/мкл
CD45RA+CD197- (TEMRA) (% от CD4)	2.8	%
CD3+CD8+ (Tc cells) (% от CD3)	30.4	%
CD3+CD8+ Lym	1536	кл/мкл
CD45RA+CD197+ (T-naïve cells)	1193	кл/мкл
CD45RA+CD197+ (T-naïve cells) (% от CD8)	77.7	%
CD45RA-CD197+ (T-central memory cells)	177	кл/мкл
CD45RA-CD197+ (T-central memory cells) (% от CD8)	11.5	%
CD45RA-CD197- (Effector memory cells)	71	кл/мкл
CD45RA-CD197- (Effector memory cells) (% от CD8)	4.6	%
CD45RA+CD197- (TEMRA)	97	кл/мкл
CD45RA+CD197- (TEMRA) (% от CD8)	6.3	%
CD19+ Lym (B-cells)	0.3	%
CD19 (В-лимфоциты), % от LYM	17	кл/мкл
CD3/CD16+56 (NKT), % от LYM	0.2	%
CD3+CD16+CD56+ Lym (T-NK-cells)	12	кл/мкл
CD56+high NK, % от NK-cells	2.0	%
CD56+high NK	12	кл/мкл
CD16+56+ (NK-клетки), % от LYM	10.0	%
CD3-CD16+CD56+ Lym (NK-cells)	581	кл/мкл

1.6 Пациент №4 (на дом)

Данные

Из выписок. Подозрение на первичный иммунодефицит. Иммунологическими тестами не проводились.

Анализ MLPA:

Локус 22q11: DGS-регион / Число копий LCR22-A n=2; LCR22-B n=2; LCR22-C n=2; LCR22-D n=2; LCR22-E n=2; LCR22-F n=2; LCR22-G n=2

методом MLPA с использованием набора реактивов P250-B2 DiGeorge probemix (MRC-Holland) проведен анализ числа копий генов локуса 22q11 (DGS-регион), изменения в котором приводят к развитию синдрома Диджорджи и велокардиофациального синдрома. Делеций/дупликаций не выявлено.

Примечание:

- LCR22-A: гены CLTCL1, HIRA, CDC45, CLDN5, GP1BB, TBX1, TXNRD2, DGCR8.
- LCR22-B: гены ZNF74, KLHL22, MED15 (PCQAP).
- LCR22-C: гены SNAP29, LZTR1.
- LCR22-D: гены HIC2, PPIL2, TOP3B.
- LCR22-E: гены RTDR1, GNAZ, RTDR1, RAB36.
- LCR22-F: гены SMARCB1.
- LCR22-G: ген SNRPD3.

Chapter 2

HLA-типирование

2.1 Генетические особенности

Главный комплекс гистосовместимости - комплекс тесно связанных генетических локусов, расположенных на коротком плече 6-ой хромосомы (6p), а также их белковых продуктов, отвечающих за развитие иммунного ответа и синтез трансплантационных антигенов. Этот комплекс содержит больше 200 генов. Сцепленные группы генов МНС называются гаплотипами. Функции генов МНС заключаются в следующем (жирным выделены функции, обозначающие цель существования такой системы)

- **Обеспечение процессинга и презентации антигенных пептидов - индукторов и мишеней иммунного ответа;**
- Обеспечение взаимодействия клеток
- Распознавание собственных, измененных собственных и чужеродных клеток; запуск и реализация иммунного ответа против носителей генетической информации
- Поддержание иммунологической толерантности (в том числе во время беременности);
- Участие в селекции Т-лимфоцитов;
- **Создание генетического разнообразия и обеспечение выживаемости вида.**

Исторически сложилось, что главный комплекс гистосовместимости у человека называют человеческим лейкоцитарным антигеном (Human Leukocyte Antigen, HLA). Поэтому процесс определения аллелей главного комплекса гистосовместимости у человека называется HLA-типированием. Генетическая карта HLA представлена

2.2 Биоинформатические особенности

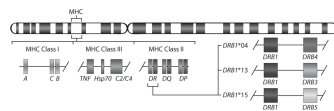


Figure 2.1: Генетическая карта HLA

Part II

Анализ иммунных репертуаров

Chapter 3

Анализ данных с помощью cellranger

Chapter 4

Библиотека Seurat

Part III

Компьютерное конструирование вакцин и антител

Chapter 5

Проектирование мультиэпитопной вакцины

Chapter 6

Защита проектов

Chapter 7

Моделирование трехмерной структуры антитела

Chapter 8

Моделирование взаимодействия антитело-антиген

Chapter 9

Зачёт

Примерный перечень вопросов к зачету.

1. Клиническая интерпретация результатов NGS. Алгоритм фильтрации вариантов
2. Заключение клинического биоинформатика
3. HLA-типирование. Область применения. Используемые программы и принципы их работы
4. Пайплайн cellranger.
5. Анализ результатов scRNA-seq. Визуализация данных
6. Иммуноинформатические ресурсы и базы данных
7. Этапы разработки медицинских препаратов
8. Типы иммунобиологических препаратов
9. Моделирование взаимодействия антитело-антиген
10. Моделирование трехмерной структуры антитела
11. Понятие об обратной вакцинологии
12. Этапы компьютерного конструирования вакцин

References

//... {#refs} //...

