



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
федеральное государственное бюджетное образовательное учреждение

высшего образования

«ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ФГБОУ ВО «ИГУ»

Кафедра физико-химической биологии, биоинженерии и биоинформатики

УТВЕРЖДАЮ

Декан биолого-почвенного факультета

А.Н. Матвеев

« 21 »

февраля 2025 г.

ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ

для проведения текущего контроля и промежуточной аттестации по дисциплине:

Б1.В.19 «ДНК МЕТАБАРКОДИНГ»

Специальность: 06.05.01 «Биоинженерия и биоинформатика»

Специализация: «Биоинженерия и биоинформатика»

Квалификация выпускника: биоинженер и биоинформатик

Форма обучения: очная с элементами электронного обучения и дистанционных образовательных технологий

Согласовано с УМК биолого-почвенного факультета

Протокол № 5 от 24 марта 2025 г.

Председатель Матвеев А.Н. Матвеев

Рекомендовано кафедрой физико-химической биологии, биоинженерии и биоинформатики

Протокол № 12 от 19 марта 2025 г.

Зав. кафедрой Саловарова В.П. Саловарова

Иркутск 2025 г.

ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ

Разработан для учебной дисциплины Б1.В.19 «ДНК метабаркодинг» специальности 06.05.01 «Биоинженерия и биоинформатика», специализация «Биоинженерия и биоинформатика». Фонд оценочных материалов (ФОМ) включает оценочные материалы для проведения текущего контроля, промежуточной аттестации в форме зачета.

Оценочные материалы соотнесены с требуемыми результатами освоения образовательной программы 06.05.01 «Биоинженерия и биоинформатика», в соответствии с содержанием рабочей программы учебной дисциплины Б1.В.19 «ДНК метабаркодинг» с учетом ОПОП.

Нормативные документы, регламентирующие разработку ФОМ:

- статья 2, часть 9 Федерального закона «Об образовании в Российской Федерации», ФЗ-273, от 29.12.2012 г.;

- ФГОС ВО по специальности 06.05.01 «Биоинженерия и биоинформатика», утвержденный приказом Министерства науки и высшего образования Российской Федерации 12 августа 2020 г. № 973.

1. Компетенции, формируемые в процессе изучения дисциплины (5 курс, 9 семестр)

ПК-1: Способен осуществлять научно-исследовательскую деятельность по решению фундаментальных и прикладных задач в области физико-химической биологии, биотехнологии, биоинформатики и смежных дисциплин

Компетенция	Индикаторы компетенций	Результаты обучения	Формы и методы контроля и оценки
ПК-1 Способен осуществлять научно-исследовательскую деятельность по решению фундаментальных и прикладных задач в области физико-химической биологии, биотехнологии, биоинформатики и смежных дисциплин	<i>ИДК ПК-1.1</i> Знает перспективы междисциплинарных исследований, основные понятия, идеи, достижения и современные направления развития физико-химической биологии, биотехнологии, биоинформатики и смежных дисциплин, основные методологические подходы и методы решения задач по тематике научных исследований	Знать: литературу по теме, владеть навыками, анализа информации сети «интернет» для поиска и освоения новых методов в эволюционной биологии ДНК-баркодинге и ДНК-метабаркодинге. Уметь: выбирать оптимальные методы и программы и базы данных в области ДНК-баркодинга и ДНК-метабаркодинга. Владеть: методами построение сложных алгоритмов, в составлении конвейеров программ для сравнительного анализа ДНК-баркодинга и ДНК-метабаркодинга.	Текущий контроль: - письменная работа (решение самостоятельных заданий) - Промежуточная аттестация: зачет
	<i>ИДК ПК-1.2</i> Умеет использовать в профессиональной	Знать: классификацию алгоритмов, и баз данных для анализа результатов ДНК-	Текущий контроль: - письменная работа (решение самостоятельных

	<p>деятельности современные представления о процессах жизнедеятельности на всех уровнях организации биологических систем, правильно ставить задачи исследования, обосновывать актуальность, новизну, теоретическую и практическую значимость исследования, выбирать и применять классические и современные методы, прогнозировать перспективы дальнейших исследований</p>	<p>баркодинга и ДНК-метабаркодинга. Уметь: анализировать входные и выходные данные в ходе исследований, проводимых с применением методов ДНК-баркодинга и ДНК-метабаркодинга. Владеть: навыками комбинации методов и информации из баз данных для практической реализации освоенных алгоритмов и программных комплексов.</p>	<p>заданий)</p> <p>- Промежуточная аттестация: зачет</p>
	<p><i>ИДК ПК-1.3</i> Владеет логикой и терминологическим аппаратом научного исследования, приемами организации работы по сбору, анализу, проведению научных исследований биосистем с использованием соответствующих методов, прикладного ПО и баз данных</p>	<p>Знать: спектр молекулярных маклеров их классификацию и ограничения при планировании исследований с применением методов ДНК-баркодинга и ДНК-метабаркодинга. Уметь: осуществлять автоматизированный поиск информации в баз данных при анализе данных и результатов ДНК-баркодинга и ДНК-метабаркодинга. Владеть: методами анализа комплексных биологических данных с использованием методов ДНК-баркодинга и ДНК-метабаркодинга.</p>	<p>Текущий контроль: - письменная работа (решение самостоятельных заданий)</p> <p>- Промежуточная аттестация: зачет</p>

2. Оценочные материалы текущего контроля

В рамках дисциплины «ДНК метабаркодинг» используются следующие формы текущего контроля - письменная работа по решению самостоятельных заданий (все формулировки заданий для самостоятельного решения с необходимыми сопроводительными материалами выложены на образовательном портале ИГУ в темах курса «ДНК метабаркодинг»);

Перечень письменных работ для самостоятельного выполнения по разделам – темам дисциплины.

Задание по теме 1:

Необходимо определить таксономическую принадлежность приведенной ниже последовательности через BLAST на сервере NCBI:

Последовательность:

```
TTTCTGGAATGATAGGAACCGCTTAAGTATGTTAATTAGAATTGAACCTTCAGCAC
CTGGAAGAATAATAGGAGATGATCATTATATAATGTTAGTAACTGCACATGCTT
TTGTTATGATATTTTTAGTAATACCAGTTAATAGGAGGTTATGGAAACTGATT
TGTACCAATTATATAGGAGCACCAGATATGGCTTCCCAAGACTAAATAACCTAAG
TTTTGATTACTCCCCCTGCACTAATTACTCTAACTTCTTTAGTAGAACAA
GGAGCTGGAACAGGTTGAACGTGTTATCCACCCCTATCTGGGCCATTAGCTCATTG
GGGGGTTCAGTTGACTATCTTAGCTGCATTGTGCTGGATTCTCTTCTATTG
CAGGAGCTATAAATTATCACAACCCTTAATATGAGAACACCAGGTTGACAT
TTGATAAAACTCCTTATTGTATGATCAGTTAATTACAGCATTAACTATTATTATA
TCATTGCCTGTATTAGCAGGAGCAATAACTATGCTTTAACTGATAGAAATTAAAT
ACTACTTTTTGATCCAGCTGGAGGG
```

Ответ:

>GU722877.1 Hydra oxycnida strain 1784b cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial

```
ACTTTATATATAATTGGAGCTTTCTGGAATGATAGGAACCGCTTAAGTATGT
TAATTAGAATTGAACCTTCAGCACCTGGAAGAATAATAGGAGATGATCATTATATA
ATGTTATAGTAACTGCACATGCTTTGTTATGATATTTTTAGTGATGCCAGTTT
AATAGGAGGTTATGGAAACTGATTGTTCCAATTATATAGGAGCACCAGATATGGC
TTTCCCAAGACTAAATAACCTAAGTTTGATTACTCCCCCTGCACTAATTAACTT
CTAAGTTCTCTTAGAGAACAGGAGCTGGAACAGGTTGAACGTGTTATCCACCC
TTATCTGGCCATTAGCTCATTGGGGGTTGGACTTAGCTATCTTAGCTTG
ATTGTGCTGGATTCTCTTCTATTGCAGGAGCCATAAATTATCACAACCCTTTAA
TATGAGATCACCAGGTTGACATTGATAAACTCCTTATTGTATGATCAGTTTA
ATTACAGCATTAACTATTATCATTGCCTGTATTAGCAGGAGCAATAACTATGC
TTTTAACTGATAGAAATTAAATAACTACTTTTGATCCAGCTGGAGGGGTGATCC
TGTACTATCACACATTATTT
```

Задание по теме 2:

В вашем распоряжении имеется файл CO1_5_S45.fasta, с нуклеотидным последовательностями фрагмента гена CO1 полученных при высокопроизводительном секвенировании ампликона из содержимого желудка рыбы выловленном в Иркутском водохранилище. Файл для самостоятельного анализа.

Вам необходимо сопоставить последовательности CO1_5_S45.fasta с данными по маркеру CO1 базы данных BOLD с помощью алгоритма BLAST выравнивания. По результатам выравнивания необходимо сгенерировать таблицу (текстовый файл с разделителем столбцов знаком табуляции) содержащую 10 ближайших гомологов для

каждой последовательности из массива CO1_5_S45.fasta с последовательностями из базы BOLD с указанием таксономических идентификаторов BOLD последовательностей.

Отдельно необходимо создать текстовые файл Word в который в качестве отчета о результатах анализа необходимо поместить таблицу, содержащую информацию о совпадениях последовательности CO1_5_S45.fasta с BOLD последовательностями, имеющим степень сходства после выравнивания 90% и более. Таблицы должны содержать 4 столбца: 1 столбец - идентификатор последовательности в файле CO1_5_S45.fasta; 2 - идентификатор последовательности базы данных BOLD; 3 - степень сходства после выравнивания в %; 4 - таксономический идентификатор базы данных BOLD (филум, порядок и так далее).

В качестве результатов выполнения задания необходимо прикрепить отчет с таблицей в формате документ Word.

Ответ:

process_id	quid	den	ax_id	ingdom	hylum	lass	rder	o	amily	ubfamily	rib	enu	pecies	ubspecies
SBAR6 09-06	NO DE_126_length_144	6.899	879	nimalia	hordata	ctinopterygii	corpaeniformes	S	ottidae	one	one	one	one	one
SBAR6 09-06	NO DE_97_length_210	6.517	879	nimalia	hordata	ctinopterygii	corpaeniformes	S	ottidae	one	one	one	one	one
SBAR6 09-06	NO DE_142_length_136	6.429	879	nimalia	hordata	ctinopterygii	corpaeniformes	S	ottidae	one	one	one	one	one
SBAR6 09-06	NO DE_140_length_137	5.536	879	nimalia	hordata	ctinopterygii	corpaeniformes	S	ottidae	one	one	one	one	one
SBAR6 09-06	NO DE_156_length_131	5.349	879	nimalia	hordata	ctinopterygii	corpaeniformes	S	ottidae	one	one	one	one	one
SBAR6 09-06	NO DE_119_length_149	5.302	879	nimalia	hordata	ctinopterygii	corpaeniformes	S	ottidae	one	one	one	one	one
SBAR6 09-06	NO DE_159_length_130	5.276	879	nimalia	hordata	ctinopterygii	corpaeniformes	S	ottidae	one	one	one	one	one
SBAR6 09-06	NO DE_172_length_128	5.2	879	nimalia	hordata	ctinopterygii	corpaeniformes	S	ottidae	one	one	one	one	one
SBAR6 09-06	NO DE_78_length_244	5.082	879	nimalia	hordata	ctinopterygii	corpaeniformes	S	ottidae	one	one	one	one	one
SBAR6 09-06	NO DE_132_length_141	5.035	879	nimalia	hordata	ctinopterygii	corpaeniformes	S	ottidae	one	one	one	one	one

Задание по теме 3:

На вычислительном сервере в директории /home/isu/DNA_Barcoding_School/fastq_data имеется серия парных прочтений ампликона 16S rRNA региона V3-V4 для 8 образцов микробиома фотического слоя воды из озера Байкал.

Необходимо каждому участнику школы подключиться к удаленному вычислительному серверу через программу WinScp (смотрите информацию по доступу к высокопроизводительному серверу). Зайти в директорию /home/isu/DNA_Barcoding_School и создать там папку с именем Имя_Фамилия, используя только латинские символы.

Далее необходимо скопировать в созданную директорию все файлы формата fastq.gz из директории /home/isu/DNA_Barcoding_School/fastq_data.

Затем необходимо создать в созданной вами директории Имя_Фамилия папку qual_data для сохранения отчетов по анализу качества секвенирования.

Затем необходимо проанализировать качество прочтения каждой пары fastq.gz файлов с помощью программы fastqc (смотрите команду для запуска fastqc в файле

/home/isu/DNA_Barcoding_School/fastq_data/fastqc.sh), запустив команды на сервере через приложение PuTTY. Все отчеты контроля качества автоматически сохраняются в созданную вами директорию quol_data в вашей личной папке Имя_Фамилия.

Далее необходимо загрузить все отчеты с сервера на ваш локальный компьютер и проанализировать их. Необходимо создать отчет по анализу в формате файла MS Word в котором должна размещаться таблица из двух колонок. Первая колонка наименование образца (наименование образца заключено в файлах fastq.gz до символа "_"), вторая колонка должна содержать информацию о количестве прочтении на образец.

Загрузите отчет MS Word на образовательный портал.

Ответ:

Name	Total Sequences
Apr12	45885
Jul10	52798
Jul19	62429
Jun6	53998
Mar29	44496
May27	43190
Sep3	71363
Sep14	85166

Задание по теме 4:

Задание по анализу данных ДНК-метабаркодинга с помощью программы USEARCH.

На вычислительном сервере в директории /home/isu/DNA_Barcoding_School/fastq_data_euk представлен набор данных в виде файлов fastq формата, представляющих результаты исследования сезонной динамики сообщества эукариотических микроорганизмов из фотического слоя воды из озера Байкал на основе ампликона фрагмента V4 гена 18S rRNA.

Необходимо в собственной папке «Фамилия_Имя», расположенной в рабочей директории /home/isu/DNA_Barcoding_School/ школы создать директорию Usearch_analysis_euk. Затем из директории /home/isu/DNA_Barcoding_School/fastq_data_euk скопировать в свою папку Usearch_analysis_euk все fastq файлы и файл silva_18s_v123.fa – база данных для таксономической идентификации OTU.

Затем необходимо провести полный цикл анализа данных, используя набор команд script_usearch_bac.sh (скачать на образовательном портале – Тема 7) и скрипт для языка программирования R - R_Usearch_data_preparation.r (скачать на образовательном портале – Тема 7). Команды файла script_usearch_bac.sh запускаются на вычислительном сервере. Команды скрипта R_Usearch_data_preparation.r выполняются на локальном компьютере. Для выполнения команд скрипта R_Usearch_data_preparation.r после выполнения команд скрипта script_usearch_bac.sh с вычислительного сервера необходимо скачать результирующие файлы M.otutab.tsv – таблица представленности OTU в пробах, M.sintax.tsv – таблица таксономии OTU, M.otus.fa – файл формата fasta с нуклеотидными последовательностями OTU.

При анализе необходимо учитывать длину праймерных областей: прямой праймер - 20 пар оснований, обратный праймер - 15 пар оснований.

Результат выполнения анализа в виде итогового файла 18S_count_tax_table_silva.tsv необходимо загрузить на образовательный портал.

Ответ:

OTU_id	Apr12	Jul10	Jul19	Jun6	Mar29	May27	Sep14	Sep3
OTU1	10628	2190	1714	956	1023	1783	2863	3619
OTU2	5589	1890	319	2538	1895	6954	0	0
OTU3	2501	930	2119	922	7788	905	262	675
OTU4	584	3283	2773	1876	1667	2299	1225	844
OTU5	4131	1115	482	1684	801	2440	877	790
OTU6	775	2858	1799	141	67	259	1206	4853
OTU7	1379	87	46	2573	171	6475	0	0
OTU8	852	1407	227	878	241	454	3596	2716
OTU9	277	467	429	335	3079	71	4841	231
OTU10	83	0	0	118	186	8012	281	0
OTU11	2262	1391	500	503	1064	1008	468	1077
OTU12	376	1239	2822	54	2170	83	383	588
OTU13	800	731	59	425	1123	579	2240	869
OTU14	979	1598	1985	54	826	135	235	492
OTU15	0	38	2966	25	582	27	1811	681
OTU16	0	1935	600	0	0	0	2105	639
OTU17	0	0	0	97	0	216	529	4083
OTU18	115	532	399	105	76	99	2316	1081
OTU19	83	106	25	250	66	49	3171	922
OTU20	2195	0	0	0	2097	30	0	0
OTU21	0	249	1127	0	0	0	1331	1479
OTU22	0	58	251	0	0	0	1708	2105
OTU23	0	209	583	0	70	0	1656	1326

Задание по теме 5:

На вычислительном сервере в директории `/home/isu/DNA_Barcoding_School/fastq_data_euk` представлен набор данных в виде файлов `fastq` формата, представляющих результаты исследования сезонной динамики сообщества эукариотических микроорганизмов из фотического слоя воды из озера Байкал на основе ампликона фрагмента V4 гена 18S rRNA.

Необходимо в собственной паке «Фамилия_Имя» расположенной в рабочей директории `/home/isu/DNA_Barcoding_School/` школы создать директорию `DADA2_analysis_euk`. Затем из директории `/home/isu/DNA_Barcoding_School/fastq_data_euk` скопировать в свою папку `DADA2_analysis_euk` все `fastq` файлы и файл `silva_132.18s.99_rep_set.dada2.fa.gz` – база данных для таксономической идентификации OTU.

Затем необходимо провести полный цикл анализа данных, используя набор команд скрипта `R_dada2_analiz.r` (скачать на образовательном портале – Тема 8) для языка программирования R. Все стадии анализа проводятся на вычислительном сервере. Дополнительную информацию по запуску консоли языка программирования R на вычислительном сервере смотрите в руководстве Использование_пакета_DADA2_практика.pdf (скачать на образовательном портале – Тема 8).

При анализе необходимо учитывать длину праймерных областей: прямой праймер - 20 пар оснований, обратный праймер - 15 пар оснований.

После выполнения анализа с вычислительного сервера необходимо скачать файл `18S_count_tax_table_Silva.tsv` с полным результатом анализа и загрузить в качестве итога выполнения задания на образовательный портал.

Ответ:

ASV_num	Apr12	Jul10	Jul19	Jun6	Mar29	May27	Sep14	Sep3
ASV1	932	1417	9002	150	190	225	8509	14390
ASV2	1203	605	1444	244	617	383	16657	13604
ASV3	3376	3988	3321	4816	4310	3302	2814	1355
ASV4	4159	2129	1571	5119	4361	4236	1363	962
ASV5	1670	2086	2397	2890	1857	2018	1547	944
ASV6	787	2216	3100	1809	881	1368	1206	1156
ASV7	148	226	1110	140	65	91	3836	3266
ASV8	1219	1029	874	1820	1126	1275	480	363
ASV9	1553	225	319	293	1444	244	2314	1630
ASV10	733	1358	1373	884	989	706	804	768
ASV11	1434	358	47	2292	572	2723	0	0
ASV12	273	186	758	124	118	67	3026	2664
ASV13	344	1798	1962	575	347	393	784	390
ASV14	844	724	510	1412	711	1058	414	748
ASV15	464	2063	1020	1149	574	803	39	28
ASV16	997	663	637	906	1102	741	686	232
ASV17	102	0	77	21	78	37	4029	1324
ASV18	779	373	360	1168	802	845	485	315
ASV19	475	724	962	635	539	598	642	383

Задание по теме 6:

По ссылке в Теме 6 - "Файл с данными для самостоятельного анализа" находится таблица 16S_count_tax_table_rdp.tsv - результат анализа ДНК-метабаркодинга бактериальных сообществ в программе USEARCH.

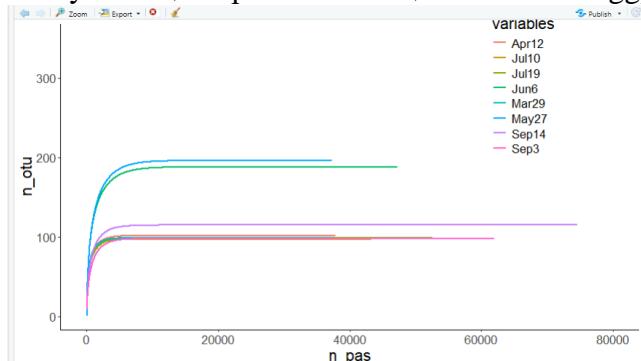
На основе этих данных необходимо провести следующие виды анализа:

- 1) По построить кривые насыщения образцов видами.
- 2) Построить кривые видового обилия.
- 3) Рассчитать индексы разнообразия, видового богатства, Chao1, ASE, Щенноа, Симпсона.

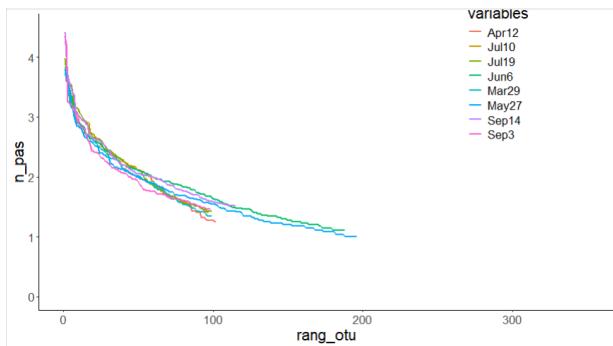
Представить результат виде таблицы.

Ответ:

#визуализация кривых насыщения в пакете ggplot2



визуализация кривых видового обилия с помощью пакета ggplot2



индексы разнообразия, видового богатства, Chao1, ASE, Щенноа, Симпсона. Представить результат виде таблицы.

sample_id	reads_number	S.obs	S.chao1	se.chao1	S.ACE	se.ACE	shannon	simpson
Apr12	37802	102	102	0	NaN	NaN	3.48	0.93
Jul10	43269	97	97	0	NaN	NaN	3.54	0.94
Jul19	52532	99	99	0	NaN	NaN	3.35	0.93
Jun6	47196	188	188	0	NaN	NaN	3.669	0.93
Mar29	37321	99	99	0	NaN	NaN	3.42	0.93
May27	37224	196	196	0	196	2.930	3.68	0.94
Sep14	74607	115	115	0	NaN	NaN	2.85	0.84
Sep3	61945	98	98	0	NaN	NaN	2.44	0.78

Критерий оценивания самостоятельной работы – результаты по каждой работе оформляются по указанным требованиям (смотрите в описании задания) и загружаются на образовательный портал ИГУ (<https://educa.isu.ru/>). Преподаватель оценивает задания, если все решено верно, студент получает зачет по заданию, если имеются недочеты или ошибки, задание отправляется на доработку с указанием допущенных ошибок. Отчёт по переработанному заданию загружается на образовательный портал для повторного оценивания.

3.Оценочные средства для промежуточной аттестации

Промежуточная аттестация проходит в форме зачета (9 семестр), к которому допускаются студенты, выполнившие в полном объеме аудиторную нагрузку, самостоятельную работу. Студенты, имеющие задолженность, должны выполнить все обязательные виды деятельности.

Фонд оценочных средств для промежуточной аттестации включает:

- тестовые задания для зачета.

Назначение оценочных средств: выявить сформированность компетенций ПК-1 (см. п. III). Тестовое задание включает два варианта по 20 вопросов по всем темам курса. К тесу допускаются студенты, задавшие все домашние заданий и получившие по каждому заданию зачет.

Критерий оценивания тестового задания для зачета

№	Тип задания	Критерии оценки	Результат оценивания
1	Задание закрытого типа на установление соответствия	Считается верным, если правильно установлены	Полное совпадение с верным ответом – 1

		все соответствия (позиции одного столбца верно соотнесены с позициями другого столбца)	балл Все остальные случаи – 0 баллов
2	Задание закрытого типа на установление последовательности	Считается верным, если правильно указана вся последовательность цифр	Полное совпадение с верным ответом – 1 балл Все остальные случаи – 0 баллов
3	Задание комбинированного типа с выбором одного верного ответа из четырех предложенных и обоснованием выбора	Считается верным, если правильно указана цифра (буква) правильного ответа и приведены корректные аргументы, используемые при выборе ответа	Полное совпадение с верным ответом – 1 балл Все остальные случаи – 0 баллов
4	Задание комбинированного типа с выбором нескольких верных ответов из четырех предложенных и обоснованием выбора	Считается верным, если правильно указаны цифры (буквы) правильного ответа и приведены корректные аргументы, используемые при выборе ответа	Полное совпадение с верным ответом – 1 балл Все остальные случаи – 0 баллов
5	Задание открытого типа с развернутым ответом	Считается верным, если ответ совпадает с эталонным ответом по содержанию и полноте	Полное соответствие эталонному ответу – 1 балл Все остальные случаи – 0 баллов

Система получения баллов за тестирование

Оценка	критерий
зачтено	15 и более баллов
незачтено	14 баллов и менее

3.1 Оценочные материалы для промежуточной аттестации (зачета)

Тестирование (Вариант 1).

Индекс и содержание формируемой компетенции	Индикаторы компетенций	Тестовые задания для промежуточной аттестации
<p>ПК-1 Способен осуществлять научно-исследовательскую деятельность по решению фундаментальных и прикладных задач в области физико-химической биологии, биотехнологии, биоинформатики и смежных дисциплин</p>	<p><i>ИДК ПК-1.1</i> Знает перспективы междисциплинарных исследований, основные понятия, идеи, достижения и современные направления развития физико-химической биологии, биотехнологии, биоинформатики и смежных дисциплин, основные методологические подходы и методы решения задач по тематике научных исследований</p>	<p>Задание комбинированного типа с выбором одного или нескольких верных ответов из четырех предложенных с аргументацией выбора</p> <p>Вопрос 1. Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какой маркер чаще всего используется для идентификации бактерий в ДНК-баркодинге? A) COI B) ITS C) 16S rRNA D) 18S rRNA Ответ _____ Правильный ответ: С Аргументация: Ген 16S рpНК (16S rRNA) является стандартным маркером для идентификации бактерий и архей, так как содержит как консервативные, так и переменные участки, что позволяет различать даже близкородственные виды</p> <p>Вопрос 2. Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какой подход используется для оценки родства организмов по маркерным последовательностям? A) Филогенетическое дерево B) Генетическая дистанция C) Секвенирование по Сенгеру D) Биохимический анализ Ответ _____ Правильный ответ: В Аргументация: Генетическая дистанция — количественная мера различий между последовательностями. Она применяется при выравнивании баркодовых фрагментов, чтобы оценить степень родства между организмами.</p>
	<p><i>ИДК ПК-1.2</i> Умеет использовать в профессиональной деятельности современные</p>	

<p>представления о процессах жизнедеятельности на всех уровнях организации биологических систем, правильно ставить задачи исследования, обосновывать актуальность, новизну, теоретическую и практическую значимость исследования, выбирать и применять классические и современные методы, прогнозировать перспективы дальнейших исследований</p>	<p>Вопрос 3. <i>Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i> Что представляет собой функция pairwiseAlignment в контексте ДНК-баркодинга? А) Визуализация данных Б) Массовое секвенирование В) Сравнение последовательностей Г) Расчет биологических индексов Ответ _____ Правильный ответ: С Аргументация: pairwiseAlignment используется в биоинформатике для парного выравнивания нуклеотидных или аминокислотных последовательностей с целью оценки их сходства.</p> <p>Вопрос 4. <i>Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i> Какая база данных предназначена для хранения баркод-последовательностей животных? А) BOLD Б) GenBank В) SILVA Г) DADA2 Ответ _____ Правильный ответ: А Аргументация: BOLD (Barcode of Life Data Systems) — специализированная база данных, содержащая баркодные последовательности (в основном COI) для животных.</p>
--	---

	<p><i>ИДК ПК-1.3</i></p> <p>Владеет логикой и терминологическим аппаратом научного исследования, приемами организации работы по сбору, анализу, проведению научных исследований биосистем с использованием соответствующих методов, прикладного ПО и баз данных</p>	<p>Вопрос 5. <i>Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Что такое E-value в BLAST-сравнении?</p> <p>A) Количество ошибок в чтении B) Скорость секвенирования C) Статистическая значимость совпадения D) Мера мутаций</p> <p>Ответ _____</p> <p>Правильный ответ: C</p> <p>Аргументация:</p> <p>E-value показывает вероятность случайного совпадения между двумя последовательностями. Чем меньше значение, тем значимее результат выравнивания.</p>
		<p>Вопрос 6. <i>Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Какой метод применяется при массовом секвенировании ДНК в метабаркодинге?</p> <p>A) ПЦР-амплификация B) Электрофорез C) High-throughput sequencing (HTS) D) Нуклеарная окраска</p> <p>Ответ _____</p> <p>Правильный ответ: C</p> <p>Аргументация:</p> <p>Методы HTS (высокопроизводительное секвенирование) позволяют параллельно расшифровывать тысячи маркерных последовательностей, что особенно важно при анализе сложных микробных сообществ.</p> <p>Вопрос 7. <i>Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Что означает термин «фильтрация по качеству» в анализе HTS-данных?</p> <p>A) Удаление дубликатов B) Удаление коротких последовательностей C) Исключение некорректных чтений D) Расчет таксономического индекса</p> <p>Ответ _____</p> <p>Правильный ответ: C</p>

	<p>Аргументация: Фильтрация по качеству — это удаление низкокачественных, сомнительных прочтений, которые могут повлиять на точность анализа таксономического состава.</p> <p>Вопрос 8. <i>Прочитайте вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Какой тип кластеризации часто применяется в метабаркодинге для группировки похожих последовательностей?</p> <p>A) UPGMA B) OTU-кластеризация C) Neighbor-Joining D) PCA</p> <p>Ответ _____</p> <p>Правильный ответ: B</p> <p>Аргументация: OTU-кластеризация (Operational Taxonomic Units) — стандартный метод группировки прочтений по сходству (обычно 97%), применяемый в метабаркодинге.</p> <p>Вопрос 9. <i>Прочитайте вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Что такое ASV (Amplicon Sequence Variant)?</p> <p>A) Клон организма B) Видовой уровень идентификации C) Уникальная нуклеотидная последовательность D) Метод сбора образцов</p> <p>Ответ _____</p> <p>Правильный ответ: C</p> <p>Аргументация: ASV — это уникальная последовательность ампликона, определённая с точностью до одного нуклеотида, что делает анализ более точным по сравнению с OTU.</p> <p>Вопрос 10. <i>Прочитайте вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Какая мера таксономического разнообразия учитывает как богатство, так и равномерность видов?</p> <p>A) Общее количество видов</p>
--	--

Б) Индекс Шеннона

С) Число OTU

Д) Длина последовательности

Ответ _____

Правильный ответ: В

Аргументация:

Индекс Шеннона учитывает как общее количество таксонов (богатство), так и равномерность их распределения в сообществе.

Вопрос 11.

Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.

Для чего используется rarefaction-кривая?

А) Определение точности чтения

Б) Оценка полноты отбора видов

С) Идентификация патогенов

Д) Построение филогенетического дерева

Ответ _____

Правильный ответ: В

Аргументация:

Rarefaction-кривые позволяют определить, насколько глубоко охвачено видовое разнообразие и хватает ли числа прочтений для оценки сообщества.

Вопрос 12.

Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.

Какая R-функция используется для визуализации таксономического состава с помощью пузырьковой диаграммы?

А) bubblePlot()

Б) barplot()

С) ggplot()

Д) heatmap()

Ответ _____

Правильный ответ: С

Аргументация:

`ggplot()` — мощный инструмент визуализации в R, позволяющий создавать сложные графики, включая пузырьковые диаграммы через `geom_point(size = ...)`.

Вопрос 13.

Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.

Какая из метрик расстояния используется в кластеризации микробиомных данных?

- A) Hamming
- B) Jaccard
- C) Bray–Curtis
- D) Euclidean

Ответ _____

Правильный ответ: C

Аргументация:

Bray–Curtis — наиболее часто применяемая метрика для оценки различий между сообществами по относительному числу таксонов, подходит для экологических данных.

Задание закрытого типа на установление соответствия

Вопрос 14.

Прочтите вопрос и установите соответствие.

Установите соответствие между функциями языка программирования R и их назначением:

- | Функция | Назначение |
|-----------|---|
| 1. mean | A. Вычисление дисперсии выборки |
| 2. var | B. Построение линейной модели |
| 3. lm | C. Расчет среднего значения |
| 4. ggplot | D. Визуализация данных с помощью графиков |

Ответ _____

Правильный ответ:

- 1. mean — C
- 2. var — A
- 3. lm — B
- 4. ggplot — D

Вопрос 15.

Прочтите вопрос и установите соответствие.

Установите соответствие между функциями языка программирования R и их назначением:

- | Функция | Назначение |
|-----------|---|
| 1. length | A. Определение количества элементов в векторе |

- | | |
|--|---|
| | <p>2. seq В. Создание числовой последовательности
 3. table С. Построение частотного распределения
 4. summary D. Получение основных статистических характеристик</p> |
|--|---|

Ответ _____

Правильный ответ:

1. length — A
2. seq — B
3. table — C
4. summary — D

Задание закрытого типа на установление последовательности

Вопрос 16.

Прочтите вопрос и установите последовательность.

Расположите перечисленные ниже этапы кластерного анализа состава сообщества в правильной последовательности. Предполагается использование иерархической кластеризации на основе дистанций Брея-Кертиса.

Этапы:

A. Визуализация результатов: Построение дендрограммы и ее интерпретация. B. Выбор алгоритма кластеризации: Определение метода агглюмерации (например, метод одиночной связи, метод полной связи, метод средней связи). C. Определение числа кластеров: Выбор оптимального числа кластеров на основе дендрограммы и/или других критериев (например, коэффициент силуэта, метод локтя). D. Сбор и подготовка данных: Формирование матрицы "сайты x виды" (или "образцы x таксоны"). E. Расчет матрицы дистанций: Вычисление матрицы дистанций Брея-Кертиса между всеми парами сайтов (или образцов).

Правильный ответ: D -> E -> B -> A -> C

Вопрос 17.

Прочтите вопрос и установите последовательность.

Расположите перечисленные ниже этапы расчета индекса разнообразия Шеннона в правильной последовательности.

Этапы:

A. Расчет доли каждого вида (pi): Вычисление доли каждого вида в общей численности/биомассе всех видов в выборке. B. Суммирование: Суммирование значений ($- pi * \log_2(pi)$) для всех видов в выборке. C. Определение общей численности/биомассы (N): Определение общего числа особей/значения биомассы всех видов в выборке. D. Определение индекса Шеннона (H): Полученная сумма является индексом Шеннона (H). E. Определение численности/биомассы каждого вида (ni): Определение числа особей/значения биомассы каждого вида в выборке.

	<p>Правильный ответ: Е -> С -> А -> В -> Д</p> <p>Задание открытого типа с развернутым ответом</p> <p>Вопрос 18. <i>Прочтите вопрос и запишите развернутый обоснованный ответ.</i> Индексы для оценки таксономического разнообразия состава сообществ их характеристики? Ответ: Индексы для оценки таксономического разнообразия состава сообществ и их характеристики играют ключевую роль в понимании структуры и функционирования экосистем. Эти индексы позволяют количественно оценить богатство видов, их относительную численность и выровненность распределения, предоставляя важные сведения для мониторинга изменений в биоразнообразии, выявления угроз и разработки стратегий сохранения. Различные индексы обладают своими преимуществами и ограничениями, и выбор конкретного индекса зависит от цели исследования и характеристик изучаемого сообщества. Одним из наиболее распространенных индексов является индекс Шеннона (Shannon diversity index), который учитывает как число видов, так и их относительную численность. Этот индекс чувствителен к изменениям в численности редких видов и полезен для сравнения разнообразия в различных сообществах или в одном сообществе в разные периоды времени. Индекс Симпсона (Simpson's diversity index) фокусируется на доминирующих видах и показывает вероятность того, что две случайно выбранные особи принадлежат к разным видам. Он менее чувствителен к редким видам и больше отражает структуру сообщества, определяемую наиболее обильными видами. Помимо индексов Шеннона и Симпсона, существуют и другие показатели, такие как индекс Маргалефа (Margalef's index), который оценивает разнообразие на основе общего числа видов и общего числа особей, и индекс Бергера-Паркера (Berger-Parker index), который отражает долю наиболее обильного вида в сообществе. Каждый из этих индексов предоставляет уникальную перспективу на структуру сообщества и позволяет оценить различные аспекты биоразнообразия. Важно отметить, что использование индексов разнообразия требует осторожности и понимания их ограничений. Например, индексы могут быть чувствительны к размеру выборки и могут не учитывать функциональные различия между видами. Поэтому, при интерпретации результатов необходимо учитывать контекст исследования и использовать несколько индексов для получения более полной картины. Кроме того, рекомендуется комбинировать индексы разнообразия с другими методами анализа, такими как построение кривых накопления видов и анализ функционального разнообразия.</p> <p>Вопрос 19. <i>Прочтите вопрос и запишите развернутый обоснованный ответ.</i> Индексы наблюдаемого и ожидаемого видового разнообразия и их использование в метабаркодинге? Ответ: Индексы наблюдаемого и ожидаемого видового разнообразия играют ключевую роль в анализе данных</p>
--	---

	<p>метабаркодинга, предоставляя ценную информацию о структуре и составе биологических сообществ. Наблюдаемое видовое разнообразие, как правило, оценивается с помощью различных индексов, таких как индекс Шеннона, индекс Симпсона или индекс Чao1, которые учитывают как количество обнаруженных таксонов (видов или операционных таксономических единиц, OTU), так и их относительную численность. Эти индексы позволяют количественно оценить разнообразие, представленное в образце, и сравнить его между различными образцами или условиями среды.</p> <p>Однако, наблюдаемое видовое разнообразие часто недооценивает истинное разнообразие, существующее в изучаемой среде. Это связано с ограничениями методов секвенирования, такими как неполный охват библиотеки, ошибки секвенирования, а также возможность обнаружения только наиболее распространенных таксонов. Для учета этих факторов используют индексы ожидаемого видового разнообразия, которые стремятся оценить общее видовое богатство сообщества, включая те таксоны, которые не были обнаружены в процессе секвенирования.</p> <p>Существуют различные подходы к оценке ожидаемого видового разнообразия, включая экстраполяционные методы, основанные на построении кривых разрежения и экстраполяции на основе модели распределения видов. Эти методы позволяют предсказать количество таксонов, которые могли бы быть обнаружены при увеличении глубины секвенирования. Другие методы, такие как моделирование на основе непараметрических оценок (например, Chao2 или ACE), используют информацию о распределении редких таксонов для оценки общего видового богатства.</p> <p>Использование как наблюдаемого, так и ожидаемого видового разнообразия в анализе метабаркодинга позволяет получить более полную картину структуры сообщества. Сравнение наблюдаемого и ожидаемого разнообразия может выявить, например, степень недооценки разнообразия в конкретном образце, что, в свою очередь, влияет на интерпретацию результатов и на выбор стратегий для дальнейших исследований. Кроме того, изменение разницы между наблюдаемым и ожидаемым разнообразием может указывать на воздействие определенных факторов среды или антропогенного влияния на структуру сообщества.</p> <p>Вопрос 20. <i>Прочтите вопрос и запишите развернутый обоснованный ответ.</i> Охарактеризуйте маркеры, которые применяются для метабаркодинга для исследования таксономического состава сообществ растений?</p> <p>Ответ:</p> <p>В метабаркодинге, мощном инструменте для исследования таксономического состава сообществ растений, применение специфичных и надежных маркеров играет решающую роль в обеспечении точности и полноты получаемых данных. Выбор оптимальных маркеров определяется рядом факторов, включая цель исследования, таксономический уровень анализа и возможности секвенирования. Среди наиболее распространенных и изученных маркеров, используемых в метабаркодинге растений, выделяются регионы генов рибосомальной РНК (рРНК), такие как ITS2 и trnL (UAA) инtron хлоропластов.</p> <p>ITS2, или интернализированный транскрибуемый спейсер 2, является наиболее часто используемым универсальным маркером для идентификации грибов и растений. Этот регион обладает высокой вариабельностью между видами, что позволяет проводить точную идентификацию на видовом уровне. Однако, в некоторых случаях,</p>
--	--

		<p>ITS2 может проявлять внутривидовую изменчивость и затруднять разделение близкородственных видов. В то же время, <i>trnL</i> (UAA) инtron хлоропластов является альтернативным маркером, который часто используется в сочетании с ITS2 для повышения точности идентификации растений. Этот регион обладает меньшей вариабельностью по сравнению с ITS2, но он позволяет проводить идентификацию на более высоких таксономических уровнях, таких как род или семейство.</p> <p>Помимо ITS2 и <i>trnL</i> (UAA), в метабаркодинге растений также используются другие маркеры, такие как <i>rbcL</i> (рибулозобисфосфаткарбоксилаза большая субъединица) и <i>matK</i> (созревание тРНК лизина). Эти маркеры обладают меньшей вариабельностью по сравнению с ITS2, но они могут быть полезны для анализа филогенетических взаимоотношений между растениями и для идентификации растений на самых высоких таксономических уровнях. В последнее время все большее внимание уделяется разработке и применению видоспецифических маркеров для решения конкретных задач в метабаркодинге растений. Такие маркеры позволяют проводить более точную идентификацию растений в сложных сообществах и изучать их экологические взаимодействия.</p> <p>Выбор оптимального маркера или комбинации маркеров для метабаркодинга растений зависит от целей исследования и особенностей изучаемого сообщества. Для исследования таксономического состава сообществ растений на видовом уровне, ITS2 является предпочтительным маркером. Однако, для повышения точности идентификации и для анализа филогенетических взаимоотношений, рекомендуется использовать комбинацию ITS2 и <i>trnL</i> (UAA). Для анализа сообществ растений на более высоких таксономических уровнях, <i>rbcL</i> и <i>matK</i> могут быть полезны.</p>
--	--	--

Тестирование (Вариант 2).

Индекс и содержание формируемой компетенции	Индикаторы компетенций	Тестовые задания для промежуточной аттестации
<p>ПК-1 Способен осуществлять научно-исследовательскую деятельность по решению фундаментальных</p>	<p><i>ИДК ПК-1.1</i> Знает перспективы междисциплинарных исследований, основные понятия, идеи, достижения и современные направления</p>	<p>Задание комбинированного типа с выбором одного или нескольких верных ответов из четырех предложенных и аргументацией выбора</p> <p>Вопрос 1. Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</p> <p>Какой маркер используют для баркодинга растений?</p> <p>A) 16S rRNA B) COI</p>

<p>и прикладных задач в области физико-химической биологии, биотехнологии, биоинформатики и смежных дисциплин</p>	<p>развития физико-химической биологии, биотехнологии, биоинформатики и смежных дисциплин, основные методологические подходы и методы решения задач по тематике научных исследований</p>	<p>C) ITS D) rbcL Ответ _____ Правильный ответ: D Аргументация: Ген rbcL кодирует большую субъединицу рубиско — фермента фотосинтеза, широко используется как стандартный маркер для ДНК-баркодинга растений. Остальные маркеры (16S — бактерии, COI — животные, ITS — грибы и растения, но менее стабилен) применяются в других таксонах.</p> <p>Вопрос 2. Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какой метод позволяет оценить степень эволюционного различия между последовательностями?</p> <p>A) Фильтрация B) Расчет генетических дистанций C) Редактирование чтений D) Сборка генома Ответ _____ Правильный ответ: B Аргументация: Генетические дистанции количественно отражают различия между нуклеотидными или аминокислотными последовательностями, что позволяет оценивать эволюционную удаленность организмов. Остальные методы не связаны с эволюционной оценкой.</p> <p>Вопрос 3. Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Что делает программа Clustal Omega? A) Строит гистограмму частот B) Строит множественное выравнивание последовательностей C) Идентифицирует белки D) Сравнивает уровни экспрессии Ответ _____ Правильный ответ: B Аргументация: Clustal Omega — популярный инструмент для множественного выравнивания последовательностей, который позволяет выявить консервативные участки и подготовить данные для построения филогенетических деревьев.</p>
	<p>ИДК ПК-1.2 Умеет использовать в профессиональной деятельности современные представления о процессах жизнедеятельности на всех уровнях организаций биологических систем, правильно ставить задачи исследования, обосновывать актуальность, новизну, теоретическую и практическую</p>	

<p>значимость исследования, выбирать и применять классические и современные методы, прогнозировать перспективы дальнейших исследований</p>	<p>Вопрос 4. Прочтайте вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</p> <p>Какая из баз данных содержит данные по грибам и микроэкариотам?</p> <p>A) UNITE B) SILVA C) KEGG D) PDB</p> <p>Ответ _____</p> <p>Правильный ответ: A</p> <p>Аргументация:</p> <p>UNITE — специализированная база данных для грибов и микроэкариот, содержит ITS-последовательности, используемые в ДНК-баркодинге грибов. SILVA — бактерии и археи, KEGG — метаболические пути, PDB — структуры белков.</p>
<p>ИДК ПК-1.3 Владеет логикой и терминологическим аппаратом научного исследования, приемами организации работы по сбору, анализу, проведению научных исследований биосистем с использованием соответствующих методов, прикладного ПО и баз данных</p>	<p>Вопрос 5. Прочтайте вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</p> <p>Какой алгоритм используется в BLAST для выравнивания последовательностей?</p> <p>A) Needleman-Wunsch B) Smith-Waterman C) Heuristic search D) Maximum likelihood</p> <p>Ответ _____</p> <p>Правильный ответ: C</p> <p>Аргументация:</p> <p>BLAST использует эвристический (heuristic) подход, что делает поиск выравниваний намного быстрее, чем точные алгоритмы вроде Smith-Waterman. Это основной метод в базах данных при поиске похожих последовательностей.</p> <p>Вопрос 6. Прочтайте вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</p> <p>Что означает термин "sequencing depth"?</p> <p>A) Количество видов в сообществе B) Количество последовательностей, полученных для каждого образца C) Длина ампликона D) Скорость анализа</p>

	<p>Ответ _____ Правильный ответ: В Аргументация: Глубина секвенирования (sequencing depth) — это число прочтений одной и той же области ДНК, характеризует полноту охвата генома и влияет на достоверность результатов анализа.</p> <p>Вопрос 7. <i>Прочтайте вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Для чего используется программа FastQC?</p> <p>A) Сбор метаданных B) Построение деревьев C) Анализ качества чтений D) Анализ OTU</p> <p>Ответ _____ Правильный ответ: C Аргументация: FastQC — инструмент для оценки качества чтений после секвенирования. Он показывает длину чтений, распределение качества по базам, количество дупликатов и GC-состав.</p> <p>Вопрос 8. <i>Прочтайте вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Что такое «котсев по длине ампликона» в анализе HTS-данных?</p> <p>A) Удаление длинных последовательностей B) Исключение последовательностей вне заданного диапазона C) Фильтрация по виду D) Редактирование ошибок</p> <p>Ответ _____ Правильный ответ: B Аргументация: Фильтрация по длине ампликона — важный шаг в препроцессинге, который исключает слишком короткие или слишком длинные последовательности, не соответствующие ожидаемому размеру гена-маркера.</p> <p>Вопрос 9. <i>Прочтайте вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Что такое кластер OTU (Operational Taxonomic Unit)?</p>
--	---

- A) Вид
B) Группа схожих последовательностей
C) Род
D) Индекс

Ответ _____

Правильный ответ: B

Аргументация:

OTU — это группировка близких последовательностей (обычно с 97% сходства), используемая как прокси для таксонов в анализе биоразнообразия. Это единица учета в микробиологических исследованиях.

Вопрос 10.

Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.

Какой индекс учитывает только разнообразие видов, но не равномерность?

- A) Индекс Шеннона
B) Индекс Симпсона
C) Число OTU
D) Индекс Чao1

Ответ _____

Правильный ответ: C

Аргументация:

Число OTU отражает видовое богатство (количество таксонов), но не говорит о распределении или равномерности, в отличие от индексов Шеннона или Симпсона.

Вопрос 11.

Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.

Что показывает rank-abundance кривая?

- A) Глубину секвенирования
B) Равномерность распределения видов
C) Скорость чтения
D) Количество образцов

Ответ _____

Правильный ответ: B

Аргументация:

Rank-abundance кривая показывает распределение относительного обилия видов, и позволяет визуально оценить доминирование и равномерность биоразнообразия в сообществе.

	<p>Вопрос 12. <i>Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Какая функция в R может использоваться для построения тепловой карты?</p> <p>A) heatmap() B) line() C) dotchart() D) matrix()</p> <p>Ответ _____</p> <p>Правильный ответ: A</p> <p>Аргументация: heatmap() — стандартная функция в R для визуализации матриц данных в виде цветной карты, часто используется в микробиомике для отображения обилия таксонов по образцам.</p> <p>Вопрос 13. <i>Прочтите вопрос, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа.</i></p> <p>Какая метрика расстояния подходит для бинарных данных при кластеризации?</p> <p>A) Jaccard B) Euclidean C) Mahalanobis D) Manhattan</p> <p>Ответ _____</p> <p>Правильный ответ: A</p> <p>Аргументация: Коэффициент Жаккара (Jaccard) широко применяется для сравнения бинарных данных, таких как наличие/отсутствие OTU или ASV в микробиологических исследованиях.</p> <p>Задание закрытого типа на установление соответствие</p> <p>Вопрос 14. <i>Прочтите вопрос и установите соответствие.</i></p> <p>Установите соответствие между функциями языка R и их назначением:</p> <table border="0" data-bbox="781 1235 1275 1267"> <thead> <tr> <th>Функция</th> <th>Назначение</th> </tr> </thead> </table> <p>1. heatmap A. Построение тепловой карты на основе матрицы 2. dist B. Расчет матрицы расстояний между объектами 3. prcomp C. Проведение анализа главных компонент</p>	Функция	Назначение
Функция	Назначение		

4. hclust D. Иерархическая кластеризация объектов

Ответ _____

Правильный ответ:

1. heatmap — A
2. dist — B
3. prcomp — C
4. hclust — D

Вопрос 15.

Прочтите вопрос и установите соответствие.

Установите соответствие между функциями и их назначением в анализе данных в R:

Функция

Назначение

- | | |
|----------------------------|--|
| 1. phyloseq | A. Объединение данных микробиомных исследований |
| 2. tax_grom | B. Агрегация таксонов до заданного таксономического уровня |
| 3. plot_bar | C. Визуализация относительной абундансности таксонов |
| 4. transform_sample_counts | D. Нормализация данных по образцам |

Ответ _____

Правильный ответ:

1. phyloseq — A
2. tax_grom — B
3. plot_bar — C
4. transform_sample_counts — D

Задание закрытого типа на установление последовательности

Вопрос 16.

Прочтите вопрос и установите последовательность.

Расположите перечисленные ниже этапы кластерного анализа состава сообщества в правильной последовательности. Предполагается использование иерархической кластеризации на основе дистанций Жаккара. Этапы:

- A. Визуализация результатов: Построение дендрограммы и ее интерпретация.
- B. Выбор алгоритма кластеризации: Определение метода аггломерации (например, метод одиночной связи, метод полной связи, метод средней связи).
- C. Определение числа кластеров: Выбор оптимального числа кластеров на основе дендрограммы и/или других критериев (например, коэффициент силуэта, метод локтя).
- D. Преобразование данных в бинарный формат (присутствие/отсутствие): Если данные представлены обилием, необходимо преобразовать их в формат 0/1 (отсутствует/присутствует).
- E. Сбор данных: Формирование матрицы "сайты x виды" (или "образцы x таксоны").
- F.

	<p>Расчет матрицы дистанций: Вычисление матрицы дистанций Жаккара между всеми парами сайтов (или образцов). Правильный ответ: E -> D -> F -> B -> A -> C</p> <p>Вопрос 17. <i>Прочтите вопрос и установите последовательность.</i></p> <p>Расположите перечисленные ниже этапы расчета индекса разнообразия Симпсона в правильной последовательности.</p> <p>Этапы:</p> <p>А. Возвведение в квадрат: Возвведение в квадрат доли каждого вида (pi^2). В. Вычитание из единицы: Вычитание суммы квадратов долей видов из единицы (1 - D). Результат - индекс разнообразия Симпсона. С. Расчет доли каждого вида (pi): Вычисление доли каждого вида в общей численности/биомассе всех видов в выборке. Д. Суммирование квадратов долей: Суммирование квадратов долей всех видов. Это значение обозначается как D (доминирование). Е. Определение общей численности/биомассы (N): Определение общего числа особей/значения биомассы всех видов в выборке. F. Определение численности/биомассы каждого вида (ni): Определение числа особей/значения биомассы каждого вида в выборке.</p> <p>Правильный ответ: F -> E -> C -> A -> D -> B</p> <p>Задание открытого типа с развернутым ответом</p> <p>Вопрос 18. <i>Прочтите вопрос и запишите развернутый обоснованный ответ.</i></p> <p>Различные виды дистанций и их характеристики при сравнении таксономического состава нескольких сообществ?</p> <p>Ответ:</p> <p>Для сравнения таксономического состава сообществ в экологических исследованиях широко используются различные меры дистанции (или меры сходства/несходства). Выбор конкретной меры зависит от типа данных (качественные или количественные), целей исследования и свойств самих сообществ. Важно понимать, что каждая мера дистанции акцентирует разные аспекты различий между сообществами, и неправильный выбор может привести к искаженным результатам и неверным выводам.</p> <p>Одним из наиболее распространенных типов мер дистанции являются дистанции на основе присутствия-отсутствия видов (бинарные данные). Классическим примером является дистанция Жаккара (Jaccard distance), которая учитывает только виды, встречающиеся в обоих сообществах, и игнорирует ситуации, когда вид отсутствует в обоих. Другой распространенной мерой является дистанция Серенсена (Sørensen distance), которая придает большее значение общим видам, что делает её более устойчивой к различиям в размерах выборки. Эти дистанции особенно полезны при анализе видов, которые трудно количественно оценить, например, редких видов или видов, обнаружение которых зависит от интенсивности поиска.</p> <p>Для анализа количественных данных, таких как численность или биомасса видов, часто используются метрические дистанции, основанные на разностях в обилиях. Дистанция Евклида (Euclidean distance) является одной из простейших и наиболее интуитивно понятных мер, но она чувствительна к различиям в общих масштабах обилия и</p>
--	---

	<p>может быть искажена доминирующими видами. Дистанция Брея-Кертиса (Bray-Curtis distance) является более робастной мерой, поскольку она нормирует различия в обилиях на общую сумму обилий, что делает её менее чувствительной к различиям в размерах выборки и более устойчивой к влиянию доминирующих видов. Кроме того, существуют дистанции, учитывающие филогенетические отношения между видами. Например, дистанция Юнифрак (UniFrac distance) оценивает разницу в филогенетическом разнообразии между сообществами, принимая во внимание структуру филогенетического дерева. Эта мера полезна для выявления различий в функциональной роли сообществ, даже если их таксономический состав схож. Выбор между весовой и невесовой версиями Юнифрак позволяет акцентировать либо присутствие-отсутствие видов, либо их относительное обилие в филогенетическом контексте.</p> <p>Вопрос 19. <i>Прочтите вопрос и запишите развернутый обоснованный ответ.</i> Многомерное шкалирование для сравнения нескольких сообществ по результатам метабаркодинга?</p> <p>Ответ:</p> <p>Многомерное шкалирование (MDS) представляет собой мощный инструмент для визуализации и анализа сложных данных, основанных на матрицах расстояний или сходства объектов. В контексте метабаркодинга, где исследователи стремятся понять различия в микробных сообществах между различными образцами или группами, MDS предоставляет возможность для представления этих различий в виде наглядной карты. Эта карта отображает взаимосвязи между сообществами, позволяя исследователям выявлять кластеры, тренды и другие важные структурные особенности данных.</p> <p>В задачах сравнения нескольких сообществ по результатам метабаркодинга, MDS используется для представления каждого сообщества в виде точки в многомерном пространстве. Расстояния между точками отражают степень различия между соответствующими сообществами. Различия могут быть основаны на самых разных метриках, таких как Bray-Curtis, Jaccard, Unifrac, и т.д., каждая из которых акцентирует разные аспекты различий между сообществами. Выбор оптимальной метрики зависит от конкретной задачи и целей исследования.</p> <p>Применение MDS в метабаркодинге позволяет исследователям выявить факторы, влияющие на структуру микробных сообществ. Например, можно сравнить микробные сообщества в разных типах почв, или в различных участках тела человека. Визуализация результатов с помощью MDS позволяет увидеть, насколько сильно различаются эти сообщества, и какие факторы могут быть ответственны за эти различия. Это может быть pH почвы, влажность, температура, диета, использование антибиотиков и т.д.</p> <p>Более того, MDS может использоваться для выявления временных изменений в микробных сообществах. Сравнивая сообщества, взятые в разные моменты времени, можно проследить динамику их изменения и выявить факторы, влияющие на эти изменения. Например, можно изучать изменения в микробиоме кишечника после изменения диеты, или после приема антибиотиков.</p> <p>Вопрос 20. <i>Прочтите вопрос и запишите развернутый обоснованный ответ.</i> Охарактеризуйте маркеры, которые применяются для в метабаркодинге для исследования таксономического состава</p>
--	--

	<p>сообществ животных?</p> <p>Ответ:</p> <p>В метабаркодинге, как мощном инструменте для исследования таксономического состава сообществ животных, выбор маркеров имеет решающее значение для получения точных и полных результатов. Идеальный маркер должен обладать рядом характеристик, включая консервативные фланкирующие регионы для универсального амплифицирования и гипервариабельные участки для разрешения таксономических групп. Среди наиболее часто используемых и проверенных маркеров можно выделить ген цитохром с оксидазы I (COI) для идентификации животных, рибосомальные гены 16S и 12S рРНК для бактерий и архей, а также ITS регион для грибов. Однако, для анализа сообществ животных COI является золотым стандартом благодаря его универсальности и развитой базе данных.</p> <p>Эффективность маркера в определении таксономического разнообразия зависит от его способности различать близкородственные виды. COI, например, обладает высокой скоростью мутирования, что позволяет дифференцировать даже виды с небольшими генетическими различиями. В то же время, важно учитывать возможность амплификации нецелевых последовательностей, таких как псевдогены или последовательности других организмов, и принимать меры для их исключения из анализа. Разработка блокирующих праймеров и тщательная проверка результатов выравнивания позволяют минимизировать ошибки и получить более точные данные о составе сообщества.</p> <p>При выборе маркера необходимо учитывать целевые группы организмов. Универсальные маркеры, такие как COI, могут обеспечивать широкое охватывающее множество видов, но их эффективность может варьироваться для разных таксономических групп. В некоторых случаях, использование видоспецифических или группоспецифических маркеров может быть более предпочтительным для детального анализа определенных групп животных. Например, для изучения сообществ насекомых могут использоваться маркеры, оптимизированные для этой группы, которые обеспечивают более высокое разрешение на видовом уровне.</p> <p>Качество и полнота баз данных являются критически важными для точной таксономической идентификации последовательностей, полученных в результате метабаркодинга. Базы данных, такие как GenBank и BOLD содержат огромное количество последовательностей и таксономической информации по маркеру COI, но их качество и полнота могут варьироваться. Регулярное обновление и проверка баз данных, а также использование нескольких баз данных для перекрестной проверки результатов, позволяют повысить точность идентификации и минимизировать ошибки.</p>
--	--

Разработчик:

Букин доцент Букин Ю.С.
(подпись)