ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ ОБЩЕЙ ГЕНЕТИКИ ИМ. Н.И. ВАВИЛОВА
РОССИЙСКОЙ АКАДЕМИИ НАУК
(ИОГен РАН)

УТВЕРЖДАЮ
директор ИОГен РАН
_____ А.В. Мисюрин
«_7_» августа 2025 г.

ПРОГРАММА ПРОФЕССИОНАЛЬНОЙ ПЕРЕПОДГОТОВКИ «ПРАКТИЧЕСКАЯ БИОИНФОРМАТИКА»

Москва 2025

1. ОБЩАЯ ХАРАКТЕРИСТИКА ПРОГРАММЫ

1.1. Аннотация программы

Данная программа предназначена для подготовки и переподготовки специалистов-биологов в области практической биоинформатики. Программа включает в себя информацию по широкому кругу вопросов, связанных с применением вычислительных методов для сравнительно-геномного анализа, структурного анализа белков и нуклеиновых кислот, анализа данных высокопроизводительного секвенирования геномов, транскриптомов и метагеномов. Программа направлена на подготовку и переподготовку специалистов в области биоинформатики широкого профиля. Программа позволяет получить знания о современных генетических технологиях, прочих омискных технологиях, обработке данных высокопроизвеодительного секвенирования, работе с базами данных, сравнительной геномике, структурной биоинформатике. Программа сочетает в себе как лекционные, так и практические занятия, направленные на закрепление лекционного материала.

В случае успешного освоения программы слушателю выдается диплом о профессиональной переподготовке – «Практическая биоинформатика».

1.2. Цели и задачи реализации программы

Целью программы является обучение специалистов-биологов теоретическим и практическим навыкам современной биоинформатики для внедрения компьютерного анализа больших данных для сопровождения решения задач геномики, в том числе, для нужд медицины, экологии и сельского хозяйства.

В задачи программы входит совершенствование компетенций, необходимых для решения задач в области вычислительной геномики и структурной биоинформатики, а также формирование новых компетенций необходимых для профессиональной деятельности в области.

1.3. Планируемые результаты обучения

В результате освоения программы слушатель должен приобрести следующие знания, умения и опыт, необходимые для достижения планируемых результатов обучения:

Виды деятельности	**Профессиональные компетенции или трудовые функции** ***(формируются и (или) совершенствуются)***	**Знания**	**Умения**	**Практический опыт**
1. Основы работ с базами данных и программирования	Способен работать с основными базами данных и писать базовые коды	З1: знает основные базы данных белков и нуклеиновых кислот	У1: умеет экстрагировать необходимую информацию из публичных баз данных	В1: владеет навыками работы с основными базами данных белков и

		З2: знает основные принципы программирования на bash и python З3: знает принципы работы командной строки	У2: умеет использовать для экстракции данных коды, написанные на python У3: умеет пользоваться командной строкой	нуклеиновых кислот В2: владеет базовыми навыкам программирования на bash и python
2. Анализ омиксных данных	Способен анализировать различные типы данных секвенирования нового поколения	З1: знает принципы работы секвенаторов. З2: знает основные принципы организации работ по анализу геномов и транскриптомов для решения задач связанных с генетическим редактированием. З3: знает алгоритмы работы программ для анализа данных высокопроизводительного секвенирования.	У1: умеет оценивать качество имеющихся данных секвенирования. У2: умеет объяснить основные принципы лежащие в основе анализа данных высокопроизводительного секвенирования. У3: умеет использовать полученные знания для оценки корректности данных и их анализа в опубликованных работах и публичных базах данных, для планирования экспериментов с использованием омиксных технологий.	В1: владеет навыками интерпретации результатов контроля качества секвенирования. В2: владеет навыками использования биоинформатических программ для анализа омиксных данных. В3: владеет основными представлениями о достижениях в сфере методов обработки омиксных данных
3.Сравнительная геномика и эволюция	Способен к сравнительно-геномному анализу последовательностей	З1: знает принципы работы основных алгоритмов выравнивания. З2: знает принципы построения филогенетических деревьев. З3: знает принципы	У1: умеет проводить множественное выравнивание последовательностей и анализировать полученные результаты. У2: умеет строить и визуализировать различные типы	В1: владеет программами и алгоритмами, осуществляющим и множественное выравнивание последовательностей, включая структурное. В2: владеет построением и визуализацией

		эволюционного анализа.	филогенетических деревьев. У3: умеет проводить эволюционный анализ последовательностей	филогенетических деревьев. В3: владеет основными методами эволюционного анализа, включая построение ландшафта приспособленности
4.Структурная биоинформатика	Способен к моделированию структур белков и нуклеиновых кислот с использованием современных алгоритмов	З1: знает принципы работы основных алгоритмов структурного моделирования белков. З2: знает принципы работы основных алгоритмов моделирования структур нуклеиновых кислот. З3: знает принципы работы молекулярного докинга и молекулярной динамики.	У1: умеет моделировать структуры белков с использованием современных алгоритмов. У2: умеет моделировать структуры ДНК и РНК с использованием современных алгоритмов. У3: умеет подбирать оптимальные структурные конформации белков и лигандов с использованием молекулярного докинга и молекулярной динамики.	В1: владеет основными современными алгоритмами reference-guided и de novo структурного моделирования белков. В2: владеет основными современными алгоритмами моделирования структуры ДНК и РНК В3: владеет методами молекулярного докинга и молекулярной динамики для подбора оптимальных лигандов и эффекторных молекул, а также оптимизации конформации белков-мишеней.

1.4. Нормативную правовую основу разработки программы составляют:

– Федеральный закон от 29 декабря 2012 г. № 273-ФЗ «Об образовании в Российской Федерации»;
– Приказ Минобрнауки России от 1 июля 2013 г. № 499 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным профессиональным программам»;
– Постановление Правительства Российской Федерации от 22 января 2013 г. № 23 «О Правилах разработки, утверждения и применения профессиональных стандартов»;

– Приказ Минтруда России от 12 апреля 2013 г. № 148н «О утверждении уровней квалификаций в целях разработки проектов профессиональных стандартов»;

Программа разработана на основе:

Профессионального стандарта «Специалист по научно-исследовательским и опытно-конструкторским разработкам» (утв. приказом Министерства труда и социальной защиты РФ от 4 марта 2014 г. № 121н)

1.5. Область применения программы (целевая аудитория)

Настоящая программа предназначена для специалистов, работающих в сфере сфере анализа живых систем.

Требования к слушателям (категории слушателей):
- высшее или средне-специальное образование в области биологии/биотехнологии/биоорганической химии/фармации;
- понимание основ процессов передачи генетической информации, базовые знания молекулярной биологии, микробиологии и биохимии

1.6. Реализация программы

Программа реализуется Федеральным государственным бюджетным учреждением науки Институтом общей генетики им. Н.И. Вавилова Российской академии наук (ИОГен РАН).

Форма обучения слушателей программы: очная.
Трудоемкость программы: 300 часов.
Язык обучения: русский.

2. СОДЕРЖАНИЕ ПРОГРАММЫ

2.1. Учебный план.

№	Наименование разделов/тем программы	Трудоемкость (час.)	Всего контактной работы (час.)	в том числе (час.)			СРС (час.)	Форма аттестации
				ЛК	ПР	ЛБ		
1	Работа с базами данных	10	10	4	6	0	0	
2	Методы и алгоритмы сравнительно-геномного анализа	36	30	12	18	0	6	Зачет по итогам выполнения практического задания

3	Основы работы в командной строке, работы на сервере.	16	16	2	14	0	0	
4	Анализ данных высокопроизводительного секвенирования геномов. Картирование прочтений	16	12	2	10	0	4	Зачет по итогам выполнения практического задания
5	Структурная биоинформатика	48	40	16	24	0	8	Зачет по итогам выполнения практического задания
6	Основы метагеномного анализа	48	40	16	24	0	8	Зачет по итогам выполнения практического задания
7	Функциональная аннотация генов и белков	6	6	2	4	0	0	
8	Сравнительная геномика и эволюция	12	12	6	6	0	0	
9	Системная биология и работа с омиксными данными	60	48	22	26	0	12	Зачет по итогам выполнения практического задания
14	Итоговая аттестация	8	-	-			8	Итоговый проект
	Итого (час.)	260	214	82	132	0	46	

2.2. Календарный учебный график

Календарный учебный график формируется непосредственно при реализации программы.

2.3. Рабочая программа разделов

№	Наименование компонентов программы	Содержание учебного материала и формы организации деятельности слушателей	Всего (час.)
1	Работа с базами данных.	Лекции, совмещенные с практическими занятиями -GenBank и ассоциированные базы -UniProt -PDB -PFAM, RFAM -Ensembl -UCSC Genome Browser -STRING -GEO -KEGG -OMIM -специализированные базы данных.	10
2	Методы и алгоритмы сравнительно-геномного анализа	Лекции, совмещенные с практическими занятиями, выполнение самостоятельных заданий -глобальные и локальные выравнивания; -парные выравнивания, динамическое программирование (основные принципы алгоритмов Ниделмана-Вунша и Смита-Уотермана); -множественные выравнивания, иерархический подход (MUSCLE); -проблема выбора параметров и структурные выравнивания (TM-align и FATCAT); -множественное локальное выравнивание и определение мотивов (MEME); -поиск сходных последовательностей в больших массивах (BLAST); -картирование на геномы (bowtie); -полногеномные выравнивания (Mauve).	36
3	Основы работы в командной строке, работы на сервере.	Лекции, совмещенные с практическими занятиями -основы программирования на bash и python -основы работы с командной строкой -основы работы на удаленном сервере	16
4	Анализ данных высокопроизводительного секвенирования геномов. Картирование прочтений	Лекции, совмещенные с практическими занятиями, выполнение самостоятельных заданий -Первичная оценка качества прочтений (FastQC) -тримминг адапторов (Trimmomatic) -картирование на геном (bowtie) -reference-guided и de novo сборка	16
5	Структурная биоинформатика	Лекции, совмещенные с практическим занятиями; выполнение самостоятельных заданий Структура белка	48

		-банк данных структур (PDB) -предсказание вторичной структуры, трансмембранных сегментов (TMHMM), сигнальных пептидов (SignalP), неупорядоченных сегментов (SPOT-Disorder) -предсказание по гомологии (SWISS-MODEL и I-TASSER); -предсказание de novo (AlphaFold and RosettaFold); -структурные (SCOPE), функциональные и эволюционные (PFAM) домены; -докинг и предсказание белок-белковых взаимодействий; -молекулярная динамика. Структура РНК: -предсказание вторичной структуры; -динамическое программирование (основные принципы алгоритмов Нуссинов и Цукера); -поиск паттерна в последовательности (RFAM); -множественное выравнивание с учетом структуры (LocARNA), ковариации позиций.	
6	Основы метагеномного анализа	Лекции, совмещенные с практическим занятиями; выполнение самостоятельных заданий Направленная метагеномика: -подбор праймеров, особенности планирования метагеномных экспериментов -анализ таксономии (SILVA, GreenGenes, специализированные базы данных грибов) -визуализация (QUIIME II) Shotgun метагеномика: -особенности работы с рвзличными образцами; -сборка метагеномов, включая фрагментированные -таксономия -функциональная аннотация -реконструкция метаболизма	48
7	Функциональная аннотация генов и белков	Лекции, совмещенные с практическим занятиями. -аннотация по сходству; -гомология, ортологи и паралоги, ксенологи и гомеологи; -изоферменты, нео- и субфункционализация, радикальное изменение функций в эволюции; -автоматическая аннотация (RAST, Prokka); -сравнительно-геномные методы – ко-локализация, ко-регуляция и ко-экспрессия (STRING); -филогенетический футпринтинг.	6
8	Сравнительная геномика и эволюция	Лекции, совмещенные с практическим занятиями Филогенетические деревья и эволюция: -методы, основанные на расстояниях (объединение соседей); -метод максимального правдоподобия (MEGA и PHYLIP); bootstrapping; -согласование деревьев; -признаки отрицательного и положительного отбора; -построение ландшафтов приспособленности.	12

9	Системная биология и работа с омиксными данными	Лекции, совмещенные с практическим занятиями; выполнение самостоятельных заданий -Первичная оценка качества прочтений (FastQC) -тримминг адапторов (Trimmomatic) -картирование на геном (bowtie) Транскриптомика -анализ дифференциальной экспрессии генов -анализ генных онтологий (GO) -картирование генов с использованием транскриптомов и de novo (Glimmer, GenMark, AUGUSTUS); -тРНК (tRNA-Scan) и другие некодирующие РНК; -регуляторные сайты ChIP-Seq -гистоны и факторы транскрипции; -подбор контролей для элиминации ложных пиков -поиск пиков (MACS2) -поиск мотивов (MEME, ChIP-Munk) Структура хроматина -ATAC-Seq и HiC; -Ribo-Seq; -проблема стандартизации экспериментальных и вычислительных протоколов, batch-эффект.	60
10	**Итоговая аттестация**		8
	Итого		260

2.4. Оценка качества освоения программы

2.4.1. Форма аттестации и оценочные материалы

Итоговая аттестация осуществляется по результатам выполнения индивидуальных практических заданий. Для успешного прохождения программы слушатель должен успешно выполнить не менее 50% предложенных заданий.

2.4.2. Анализ удовлетворенности требований слушателей.

По итогам обучения проводится анкетирование слушателей для анализа удовлетворенности их требований к организации и качеству обучения.

3. ОРГАНИЗАЦИОННО-ПЕДАГОГИЧЕСКИЕ УСЛОВИЯ РЕАЛИЗАЦИИ ПРОГРАММЫ

3.1. Материально-техническое обеспечение

Наименование учебных помещений	Вид занятий	Наименование оборудования, программного обеспечения
Аудитория для проведения лекционных занятий, аудитория для проведения практических занятий.	Лекции, вычислительные практикумы.	– Стационарный ПК или ноутбук, оснащенный выходом в сеть Интернет, и имеющий характеристики, достаточные для функционирования x64 ОС и прикладного программного обеспечения, с объемом оперативной памяти не менее 8 ГБ, свободным местом на жестком диске не менее 80 ГБ. – сервер с объемом оперативной памяти не менее 2 ТБ, не менее 256 вычислительных потоков, не менее 40 ТБ жестких дисков. – Программное обеспечение: Linux, (опционально) Microsoft Windows версии 10 или более поздней, браузер.

3.2. Учебно-методическое обеспечение и информационное обеспечение

Перечень программного обеспечения (подлежит обновлению при необходимости).

NCBI SRA Toolkit https://github.com/ncbi/sra-tools/wiki/01.-Downloading-SRA-Toolkit#header-global
Fastp https://github.com/OpenGene/fastp
Seqtk https://github.com/lh3/seqtk
NCBI BLAST: https://blast.ncbi.nlm.nih.gov/Blast.cgi
Bowtie2: http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
BWA MEM 2 https://github.com/bwa-mem2/bwa-mem2
Minimap2: https://github.com/lh3/minimap2
STAR: https://github.com/alexdobin/STAR
SAMtools, BCFtools: http://www.htslib.org/
Spades https://cab.spbu.ru/software/spades/
Tablet: https://ics.hutton.ac.uk/tablet/
IGV: https://software.broadinstitute.org/software/igv/home
Braker: https://github.com/Gaius-Augustus/BRAKER
I-Tasser: https://zhanggroup.org/I-TASSER/
AlphaFold: https://alphafold.ebi.ac.uk/
MEME: https://meme-suite.org/meme/
MUSCLE: https://www.ebi.ac.uk/jdispatcher/msa/muscle
T-Coffee: https://tcoffee.crg.eu/

Перечень профессиональных баз данных и информационных справочных систем (подлежит обновлению при необходимости)

Базы данных национального центра биотехнологической информации (NCBI):
https://www.ncbi.nlm.nih.gov/
Ensembl https://www.ensembl.org/index.html

cBioPortal for cancer genomics: http://www.cbioportal.org/
UCSC Genome Browser: https://genome.ucsc.edu/
UniProt: https://www.uniprot.org/
PDB: https://www.rcsb.org/
PFAM (InterPro): https://www.ebi.ac.uk/interpro/
KEGG: https://www.genome.jp/kegg/
GEO: https://www.ncbi.nlm.nih.gov/geo/

3.3. Организация образовательного процесса

Основой программы являются лекционные и практические вычислительные занятия. Часть учебного времени отводится на выполнение индивидуальных зачётных заданий.

3.4. Кадровое обеспечение образовательного процесса

Кадровое обеспечение программы осуществляют сотрудники ИОГен РАН и Сколтеха.

4. РУКОВОДИТЕЛЬ И РАЗРАБОТЧИКИ ПРОГРАММЫ

Руководитель программы:
Миронов Андрей Александрович, д.б.н., профессор МГУ, главный научный сотрудник лаборатории сравнительной и функциональной геномики регуляторных систем ИОГен РАН. Имеет более, чем 40-летний опыт разработки курсов и преподавания вычислительных методов в биологии и практической биоинформатики для биологов (Биофак МГУ, ФББ МГУ, ВШЭ). Имеет большой опыт успешного руководства студентами и аспирантами в области биоинформатики, а также научными проектами, поддержанными РФФИ и РНФ, посвященными биоинформатическому анализу больших массивов данных.

Разработчики программы:
Гельфанд Михаил Сергеевич, д.б.н., профессор, вице-президент Сколковского института науки и технологий. Имеет около 40 лет опыта разработки и чтения курсов по вычислительной молекулярной биологии (Биофак МГУ, ФББ МГУ, Сколтех, ВШЭ). Имеет большой опыт руководства студентами и аспирантами, а также работы со старшими школьниками. Имеет опыт руководства научными проектами, поддержанными РФФИ и РНФ, посвященными сравнительно-геномному анализу.

Казнадзей Анна Денисовна, к.б.н., старший научный сотрудник лаборатории сравнительной и функциональной геномики регуляторных систем ИОГен РАН. Имеет более 10 лет опыта чтения лекций и проведения практических занятий по биоинформатике студентам и старшим школьникам.

Тутукина Мария Николаевна, к.б.н., зав. лабораторией сравнительной и функциональной геномики регуляторных систем ИОГен РАН. Имеет более 10 лет опыта чтения лекций и проведения практических занятий по сравнительной и функциональной геномике бактерий и метагеномике студентам (МФТИ, ВШЭ) и старшим школьникам.