

## ЗИМНЯЯ МОЛОДЕЖНАЯ НАУЧНАЯ ШКОЛА “EXPLORING A PERSONAL GENOME”, НОВОСИБИРСК, 2018

Стремительное развитие молекулярно-биологических технологий, в общем, и в частности, технологий секвенирования обусловило взрывной рост объемов геномной информации. К настоящему времени секвенированы геномы сотен тысяч людей. Существует распространенное мнение, что геном человека может рассказать о нем всё - его этническую принадлежность, рост, цвет волос и глаз, предрасположенность к полноте и различным заболеваниям, продолжительность жизни и то, как ему надо питаться, каким спортом заниматься и многое другое. Так ли это на самом деле? Что действительно может рассказать геном? И с какой точностью? Ответам на эти вопросы была посвящена зимняя молодежная научная школа “Exploring a personal genome”, которая прошла с 29 января по 6 февраля 2018 г. в Новосибирском государственном университете (НГУ) и Институте цитологии и генетики СО РАН (ИЦиГ СО РАН). Рабочий язык школы – английский. В работе школы приняли участие 6 бакалавров, 8 магистров, 4 аспиранта, а также четверо молодых ученых кандидатов наук из городов России (Иркутск, Новосибирск, Москва, Санкт-Петербург) и Казахстана. Преподавателями школы стали ученые из НГУ, ИЦиГ СО РАН и Института GIGA (Льежский университет, Бельгия): Юрий Аульченко, Мишель Жорж, Кароль Шарльи, Марианна Бегова и Ваутер Коппитерс.

Наряду с интенсивным лекционным курсом занятия включали исследовательскую работу над научными проектами с использованием геномных данных. Участие в школе предполагало знание основ генетики человека и геномики в объеме онлайн-курса “From disease to genes and back” (coursera.org), навыков программирования и анализа данных, и знание основ биоинформатики. Состав участников школы был представлен специалистами, связанными с анализом биологических данных в различных областях естественных наук: биологами, генетиками, биофизиками, математиками, медиками, биоинформатиками.

Занятия в школе были разделены на два блока: лекции и командная работа над научными проектами. На лекциях участников знакомили с методами изучения происхождения человека, моногенных и сложных полигенных признаков, которые зависят как от генетических факторов, так и от условий окружающей среды. Отдельное внимание уделялось принципам построения геномных моделей для предсказания антропометрических характеристик и предрасположенности к заболеваниям. Лекции вызвали большой интерес. Уточняющие вопросы и обсуждения продолжались во время перерывов.

Отличительной чертой школы стала творческая исследовательская работа над научными проектами. Здесь участники на практике познакомились с реальными задачами, которые решают ученые в научных лабораториях. Целью практической части школы было применение полученных на лекциях знаний при работе с генетическими данными реальных людей. Кроме того, участники получили неоценимый опыт работы в команде и опыт представления полученных в ходе выполнения проектов результатов.

### Проекты школы и работа над ними

Были сформулированы три научных проекта (рис. 1): определение этнического происхождения, менделевские признаки и сложные признаки. Для работы над научными проектами участникам школы предоставили геномные данные конкретных людей, а именно, трио (мать, отец и дочь). Участникам было предложено проанализировать эти данные и ответить на ряд вопросов о каждом индивидууме: Что вы можете сказать об этническом происхождении этого человека? Можете ли вы сказать, где он родился? Какая доля его генома была унаследована от неандертальцев? Каков пол этого человека, цвет глаз и рост? Насколько он предрасположен к определенным наследственным моногенным и полигенным заболеваниям?

В рамках четвертого, образовательного, проекта было необходимо подготовить материалы школы для онлайн

4 проекта	
<p><b>Этническое происхождение</b></p> <p>Место рождения, раса Доля генома от неандертальцев</p>	<p><b>Менделевские признаки</b></p> <p>Цвет кожи и волос Моногенные заболевания</p>
<p><b>Сложные признаки</b></p> <p>Риск ишемической болезни сердца и болезни Крона Рост</p>	<p><b>Образовательный курс</b></p> <p>Учебная программа Материалы для онлайн курса</p>

Рис. 1. Практические проекты школы.

курса. На основе видео- и аудиозаписей лекций школы участники проекта должны были разработать концепцию и программу курса. Для последующего монтажа, в видеоматериале нужно было найти границы отдельных смысловых блоков. Также, было необходимо собрать дополнительный материал, который помог бы в понимании идей, не рассмотренных детально в ходе лекций.

В первый день школы участники выбрали один из четырех проектов. Затем, организаторами были сформированы команды участников. Команды работали над проектами самостоятельно, при консультационной поддержке преподавателей. В каждом проекте был получен конкретный теоретический и практический результат.

Например, одна из команд изучала этническое происхождение исследуемого трио. При сопоставлении исследуемых образцов ДНК с общедоступными генетическими базами данных было выявлено, что образцы ДНК получены от людей европеоидной расы.

Работа другой команды была посвящена менделевским признакам, которые определяются мутациями только в одном гене. Своим опытом работы над этим проектом поделилась выпускница биологического факультета СПбГУ, специалист по анализу геномных данных, Parseq Lab **Анастасия Коростелева**:

*- Я выбрала проект по поиску и изучению менделевских признаков человека по данным полногеномного и полноэкзомного секвенирования. Руководитель проекта дала участникам большую свободу в выборе направления работы, наша команда разделилась на тех, кому ближе подход “от данных – к вариантам”, и тех, кто предпочитает анализировать уже известные варианты в геномных данных. В результате работы наша команда определила такие фенотипы, как, например, цвет глаз, цвет кожи и во-*

*лос, толерантность к лактозе и резистентность к некоторым типам лекарств.*

Задачей третьей команды было построение предсказательных моделей для сложных полигенных признаков. На основании данных из литературы о полногеномных исследованиях ассоциаций была построена модель, связывающая рост взрослого человека с набором генетических вариантов и других факторов. Кроме определения роста исследуемых индивидов, также предлагалось оценить риск ишемической болезни сердца и болезни Крона.

Отдельно стоит сказать об исходных экспериментальных данных для проектов. Они предоставлялись участникам «как есть», без дополнительных пояснений. В процессе практической работы над проектами у участников возникало множество вопросов. Часто требовалось конвертировать данные из одного формата в другой, анализировать качество этих данных, фильтровать в зависимости от условий задачи, визуализировать. Стало очевидно, что без базовых знаний программирования обойтись невозможно или, по крайней мере, нужно иметь в команде программиста широкого профиля.

Также важным “откровением” для участников, большинство из которых имели базовое биологическое или медицинское образование, стала необходимость понимания теории вероятностей, статистики и линейной алгебры. И дело даже не в знании конкретных формул, а в понимании и применении основ: для чего нужен корреляционный анализ, что такое генеральная совокупность, что такое собственные числа и как найти главные компоненты в многомерных данных. Математические методы невозможно освоить за несколько дней. Поэтому, наряду с программистом в команде должен быть математик и/или биоинформатик.

В то время, пока три команды работали с генетическими данными, четвертая команда составила учебную программу образовательного онлайн курса, разбила длинные лекции школы на короткие уроки до 15 минут каждый, подготовила дополнительные материалы к занятиям и вопросы после каждого занятия для проверки усвоенного материала.

В результате исследования геномов конкретных людей выяснилось, что некоторые признаки, например, цвет глаз и этническое происхождение - можно предсказывать с высокой точностью. Однако при предсказании сложных признаков – таких как риск ишемической болезни сердца и рост – добавочная ценность геномной информации зачастую минимальна.

Работа над проектами завершилась презентацией полученных результатов в последний день школы, 6 февраля, на мини-конференции, куда в качестве слушателей кроме участников и преподавателей школы также были приглашены сотрудники научно-исследовательских институтов г. Новосибирска.

## Итоги школы

Подведение итогов школы завершилось торжественным вручением сертификатов об участии, коллективной фотосъемкой и, конечно же, опросом. Отзывы участников свидетельствовали о том, что школа соответствовала их ожиданиям, а лекции были интересными и полезными, что позволило лучше разобраться в освещаемых темах. Работа в командах над практическими проектами принесла пользу и удовольствие большинству слушателей, а некоторые признаются, что преподаватели и члены команды за время школы стали для них хорошими друзьями и почти семьей. Участникам хотелось бы больше практической работы непосредственно с преподавателями на примере решения простых задач еще до погружения в работу над проектом. Поскольку уровень подготовки и опыт участников был разным, полезным было бы проведение семинаров с практическими советами и ознакомлением с существующими методами анализа данных и особенностями применения каждого из подходов.

Несмотря на сложности, с которыми пришлось столкнуться во время выполнения практических проектов, участники готовы рекомендовать школу своим друзьям и коллегам.

## Что говорят участники о прошедшей школе

Участники школы поделились своими впечатлениями через два месяца после возвращения из Новосибирска и рассказали, как полученные знания и опыт повлияли на их учебу и работу.

**Иван Петрушин** (доцент Иркутского государственного университета, к.т.н., г. Иркутск):

*– Важнейшим результатом школы для себя вижу знакомство с актуальными результатами статистической геномики. Теперь лучше понимаю, что современная наука и математические методы могут или пока не могут рассказать о конкретном человеке и популяции. Пускай практическая часть занятий охватила лишь небольшую часть тем, но разбудила немалый интерес к направлению. Я после школы специально разбирался с методом главных компонент, так как он применим и в других областях с многомерными данными.*

**Сергей Славский** (магистр Сколтеха, г. Москва):

*– Когда у нас в Сколтехе начался курс по эволюционной биологии, мне пригодились знания из популяционной генетики, о происхождении человека и о методах проверки статистических гипотез. Пожалуй, самое важное, что я вынес из работы на школе – нужно учиться работать с реальными данными. На лекции всё кажется понятным, но, когда пытаешься повторить описанное уже самостоятельно и на реальных примерах, то обнаруживается много подводных камней.*

**Анастасия Коростелева** (биологический факультет СПбГУ; специалист по анализу геномных данных, Parseq Lab, г. Санкт-Петербург):

*– О возможности поехать на научную школу мне сообщила коллега. Я только начала осваиваться в области работы с данными NGS и геномики, поэтому участие в подобных школах и семинарах мне необходимо для интенсивного обучения, а также для общения с единомышленниками. Одной из моих обязанностей в Parseq Lab является работа с редкими генетическими вариантами, поэтому на школе мне было интересно узнать, какие существуют подходы в этом вопросе. Статистике в биологии на естественно-научных факультетах традиционно уделяется недостаточное внимание. Курс лекций от организаторов отлично подходит для введения в тему и рождает интерес для самостоятельного обучения. Также школа стала источником новых знакомств с сильными и начинающими специалистами.*

**Андрей Курицын** (студент магистратуры по биоинформатике, Санкт-Петербургский политехнический университет им. Петра Великого, Институт прикладной математики и механики, г. Санкт-Петербург):





**Рис. 2.** Участники зимней молодежной научной школы “Exploring a personal genome” вместе с преподавателями после представления проектов. Новосибирск, февраль 2018 г.

Слева направо, нижний ряд: В. Коппитерс, К. Шарльи, А.Ж. Газизова, Л.В. Рожненко, П.С. Белокопытова, Э.Ю. Хлебус, А.А. Курицын, С.А. Славский, М. Жорж, Е.Е. Елгаева, К.А. Юзаю, Л.И. Шагам, А.А. Торгашева, Д.Д. Романовская, Е.В. Путинцева; верхний ряд: Ю.С. Аульченко, И.С. Петрушин, С.Ж. Шарапов, М.Р. Бевова, Т.И. Шашкова, А.В. Северинов, Д.Д. Горев, Д.С. Вибе, А.И. Ташкеев, А.Е. Барановский.

– Вспоминаю неделю, проведенную в школе, как самое яркое впечатление за последние несколько лет. Так как в настоящее время я сейчас работаю не по специальности биоинформатика, то на работу школы не повлияла. Зато на учебе, когда мы проходили методы секвенирования и GWAS, мне было гораздо легче “схватывать” суть. Больше всего я рад тому, что смог проехать половину нашей необъятной страны и познакомиться с большим количеством интересных людей и замечательных преподавателей, каждый из которых смог значительно расширить мой кругозор и показать, какая же на самом деле интересная наука.

**Полина Белокопытова** (магистр 1 курса ФЕН НГУ, г. Новосибирск):

– Полезно было поработать с Linux. Оказывается, большинство инструментов биоинформатики разработано для этой операционной системы. Познакомилась с разными базами данных по геномам, онлайн-инструментами для работы с генетическими вариантами. Когда общалась с лекторами, то они с удовольствием рассказывали и о своей

стране, например, что благодаря развитой медицине, в Бельгии сделать анализ наследственной информации не так и дорого.

**Элеонора Хлебус** (аспирантка МФТИ (ГУ); м.н.с. лаборатории молекулярной генетики ФГБУ «Национальный медицинский исследовательский центр профилактической медицины» Минздрава России, г. Москва):

– Одним из наиболее интересных проектов мне показался образовательный проект, задачей которого было создание онлайн курса на основе лекций, прочитанных на школе. Работа над этой задачей помогла мне основательно разобраться в учебном материале и применить некоторые методы при работе над моей кандидатской диссертацией, например, для анализа исследуемой выборки я использовала метод главных компонент. В ходе обсуждения некоторых результатов моей диссертации с Мишелем Жоржем и благодаря его “правильным” вопросам, я посмотрела на свое исследование с новой стороны, и поняла, что и как нужно доделать, и с большей частью уже справилась после школы. Посещение школы и общение как с преподавателями, так и с другими

*участниками зарядило меня энергией и мотивировало на более активную и продуктивную работу у себя в лаборатории. А новые идеи, которые я почерпнула на лекциях, дали толчок самообразованию и профессиональному росту.*

## Заключение

На наш взгляд, ценность подобных школ в том, что молодые специалисты работают над научно и практически актуальными задачами, параллельно обучаясь и применяя полученные знания на практике. В отличие от лабораторных работ, на школе не было заранее заготовленных планов работ над проектами, а результат не был заранее известен. Более того, сведения о нужных методах, алгоритмах и описание программ приходилось искать самостоятельно в первоисточниках -- научных статьях и сайтах разработчиков программного обеспечения. Лекции чередовались с практической работой и регулярным обменом опытом между участниками школы и преподавателями, а также, что немаловажно, между членами одной команды -- специалистами в разных областях знаний. Как

говорили участники: "Видишь проблему - скажи товарищам и ищи решение". Такой интенсивный режим работы способствует проактивному поведению и отличному усвоению теоретического материала.

## Благодарности

Зимняя школа "Exploring a personal genome" была организована лабораторией теоретической и прикладной функциональной геномики (САЕ «Синтетическая биология» ФЕН НГУ) в сотрудничестве с Институтом GIGA (Льежский университет, Бельгия) и Институтом цитологии и генетики СО РАН. Школа прошла при поддержке Министерства образования и науки РФ по Программе повышения конкурентоспособности ведущих российских университетов среди ведущих мировых научно-образовательных центров (Проект 5-100).

Авторы выражают особую благодарность всем тем, кто участвовал в интервью после окончания школы и поделился своими впечатлениями от школы и личным опытом использования полученных знаний.

Э.Ю. Хлебус<sup>1,2\*</sup>, И.С. Петрушин<sup>3</sup>, М.Р. Бегова<sup>4</sup>, Ю.С. Аульченко<sup>5,6</sup>

<sup>1</sup> Федеральное государственное бюджетное учреждение "Национальный медицинский исследовательский центр профилактической медицины" Минздрава России, Москва, Россия

<sup>2</sup> Московский физико-технический институт (государственный университет), Москва, Россия

<sup>3</sup> Иркутский государственный университет, Иркутск, Россия

<sup>4</sup> GIGA научно-исследовательский центр, Льежский университет, Льеж, Бельгия

<sup>5</sup> Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук», Новосибирск, Россия

<sup>6</sup> Федеральное государственное автономное образовательное учреждение высшего образования «Новосибирский государственный университет», Новосибирск, Россия

\*elkhlebus@gmail.com

Публикуется в авторской редакции.  
Опубликовано онлайн 20.06.2018 г.

## КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Э.Ю. Хлебус, И.С. Петрушин, М.Р. Бегова, Ю.С. Аульченко. Зимняя молодежная научная школа "Exploring a personal genome", Новосибирск, 2018. Письма в Вавиловский журнал. 2018. [http://www.bionet.nsc.ru/vogis/download/conference/appx\\_3.pdf](http://www.bionet.nsc.ru/vogis/download/conference/appx_3.pdf)

## HOW TO CITE THIS ARTICLE:

E.Yu. Khlebus, I.S. Petrushin, M.R. Bevova, Y.S. Aulchenko. Winter school for young scientists "Exploring a personal genome", Novosibirsk, 2018. Pisma v Vavilovskii Zhurnal. 2018. [http://www.bionet.nsc.ru/vogis/download/conference/appx\\_3.pdf](http://www.bionet.nsc.ru/vogis/download/conference/appx_3.pdf)