

# Путевые заметки — второй месяц стажировки студентов СФУ в Барселоне

Подошёл к концу второй месяц стажировки двух студентов ИФБиБТ СФУ Артура Туманяна и Ивана Денисова в Центре Геномных Регуляций (Барселона, Испания). У ребят остался всего месяц на завершение своих проектов. Приобретённые в течении стажировки знания и навыки по биоинформатике должны помочь доделать работу на достойном уровне. Познакомимся немного с биоинформатикой на примере текущих проектов Артура и Ивана, но для начала небольшое введение в проблематику этой новой области.

Биоинформатика начала формироваться, когда использование компьютерных технологий в биологии стало необходимостью. При масштабных работах по генетике пришлось обрабатывать объёмный наследственный материал животных — ДНК, а позже анализировать РНК и белки. Результативную обработку последовательности ДНК, которая у человека состоит из 3 миллиардов нуклеотидных оснований возможно производить только на компьютере. Сегодня активно применяются вычислительные кластеры, много компьютеров, объединённых в единую систему для вычислений. Биоинформатика также активно использует достижения математики, химии, физики и играет значительную роль в теоретической биологии.

Биоинформатика помогает молекулярным биологам сужая область поиска лекарств, помогает им ориентироваться в молекулярном мире, создаёт новые компьютерные приложения для работы с биологической и медицинской информацией. В руках биологов биоинформатика расширяет представление о живых организмах и популяциях, например, помогая отвечать на вопросы о механизмах эволюции, или, восстанавливая данные о едином предке всех живых организмов.



Артур: По приезду в Барселону я попал в лабораторию эволюционной геномики Федора Алексеевича Кондрашова, и работал там над очень интересной темой, связанной с мутациями в альфа-синуклеине (далее синукленине), которые вызывают болезнь Паркинсона. Эта болезнь связана с образованием большого количества ненормальных белковых образований (телец Леви) внутри нейронов. Тельца Леви состоят из «неправильно» свернувшегося синукленина. Если в гене, кодирующем синукленин возникает мутация, то меняется аминокислотная последовательность белка, что может привести к изменению третичной структуры, а следовательно повлиять на функцию белка и его взаимодействие с другими белками.

Известно, что некоторые виды обезьян не болеют болезнью Паркинсона, хотя имеют Тимин в 53-ем сайте, который патогенен для человека. Было высказано предположение что болезнь у обезьян не возникает благодаря полезной мутации в аминокислотной последовательности белка, который взаимодействует с синуклеином. Такая замена называется «компенсаторной», поскольку она стабилизирует процесс взаимодействия белка с «неправильным» синуклеином, и, тем самым, улучшает приспособленность организма к среде.

Поиск аминокислотных замен осуществлялся согласно филогенетическим данным синуклеина. Необходимым условием являлось то, что если синуклеин претерпел замену в последовательности, то и белок, в котором произошла компенсация должен изменить последовательность в тот же эволюционный промежуток. Важно, что нам нужны не все замены, а только те, которые происходили одновременно с заменой Тимина на Аденин в 53-ем сайте в синуклеине человека, поскольку именно эта замена предположительно влияет на появление болезни.

Исходя из этих условий был найден белок, в котором находятся 2 возможных сайта компенсации. Данные в ближайшее время будут отправлены в экспериментальную лабораторию для подтверждения. В ходе стажировки я научился работать со специализированными приложениями для биоинформатики, профессиональными базами данных, познакомился с компенсаторной и коэволюционной теориями.

В настоящее время занимаюсь изучением скорости эволюции белков, которые первоначально дублировались, а затем, одна из копий была потеряна. Известно, что после дубликации скорость накопления и фиксирования точечных мутаций (скорость эволюции) возрастает. Вопрос, на который должна ответить моя работа звучит так: как изменится скорость эволюции гена, если после дубликации будет потеряна одна из его копий? В этом вопросе многое остаётся неясным, но до отъезда надеюсь успеть разобраться в явлении дубликаций более основательно.



*Иван Денисов:* Моё исследование в лаборатории сравнительной геномики Тони Габалдона направлено на разгадку истории генов. Генами называют участки ДНК, в которых содержится информация о всех биомолекулах в живых организмах. Информация эта меняется со временем согласно некоторым правилам и иногда копируется, таким путём под контролем окружающей среды образуются новые виды животных, растений и грибов.

Проблема состоит в том, что изучая сегодня гены, мы с большим трудом можем восстановить их генеалогическое дерево, сказать кто от кого произошёл и когда это случилось. А чтобы построить филогенетическое дерево животных, например, важно опираться только на те гены, которые не претерпели большого числа дубликаций (копирований), иначе будет большая погрешность. Также считается, что если два гена являются копиями друг друга в прошлом, то один очень вероятно изменил свою функцию или локализацию, иначе он не сохранился бы до наших дней и стёрся за ненадобностью. И наоборот, если гены являются прямыми потомками, то и роль свою они по наследству передадут. На основе этого предположения и филогенетических данных делается много предсказаний о наследовании функций генами. А поскольку исследовать функции всех генов всех организмов нельзя из-за невероятного объёма геномов, то такое предсказание играет важную роль. Но наиболее полезным в медицине является предсказание взаимодействия двух белков на основе их родословных. Если истории двух генов очень похожи, то велика вероятность, что они как-то связаны в переплетённой сети метаболизма, и, что если меняется один ген, то и второй должен измениться, чтобы белки смогли продолжать работать в как и прежде в тандеме.

Моя задача стояла в том, чтобы показать, что классическая теория, которая описана в предыдущем абзаце, на практике не всегда выполняется, поскольку историю многих дубликаций мы уже никогда не сможем восстановить. Так что, те белки, которые выглядят прямыми наследниками, могут оказаться копиями, чей истинный предок пропал. Поэтому вероятность иметь разные функции у эволюционно отдаленных генов одинакова, не зависимо от того какое мы приписали им происхождение.

Доказательство описанного этого опровержения построено на вероятностной модели возникновения дубликаций, потери генов и приобретения ими новой функции. Параметры модели планировалось оценить из анализа экспериментальных данных. А именно, большой базы данных филогенетических деревьев (<http://phylomedb.org/>), которую подготовили ребята в лаборатории Тони Габалдона. Но по ходу анализа данных было обнаружено необычное статистическое разделение свойств генов которое заслуживает отдельного внимания. Это разделение свойств стало предметом моего изучения и дополнило модель дубликаций генов новым пониманием проблемы. На сегодняшний день удалось доказать, что распределение не является результатом бинарной топологии деревьев, и, что

вероятность дубликации не является постоянной, как и скорость эволюции гена. В ходе работы разработаны алгоритмы генерации случайных деревьев, описана модель дубликаций. Планирую завершить проверку функций и локализацию работы генов основываясь на базе данных их онтологии, объединить все результаты в единый отчёт который совместно с работами Испанских коллег, надеюсь, станет достоин полноценной публикации.

Иван и Артур благодарны руководству СФУ и банку Сантандер за предоставленную стажировку, которая стала серьёзным шансом реализовать исследовательский потенциал, получить уникальный опыт и вдохновиться на покорение новых вершин в науке.



Иван Денисов — справа

*20 апреля 2010 г.*

© Сибирский федеральный университет. Редакция сайта: +7 (391) 246-98-60, [info@sfu-kras.ru](mailto:info@sfu-kras.ru).

Адрес страницы: <https://news.sfu-kras.ru/node/5814>