

Как биоинформатика и системная биология появились на кафедре «Компьютерные технологии» Университета ИТМО

Часть 1. 2009-2011 гг.

<https://vk.com/@1077823-kak-bioinformatika-i-sistemnaya-biologiya-poyavilis-na-na-ka>

Поэт **Александр Кушнер** написал: «**Времена не выбирают, в них живут и умирают**». Если я бы не верил в классную поэзию, то у нас, в частности, не получилась бы история со сборкой генома, которая в дальнейшем переросла в суперуспешные исследования по биоинформатике и системной биологии под руководством **Алексея Сергушичева** (http://fppo.ifmo.ru/?page1=16&page2=52&page_d=1&page_d2=146244), а затем и **Владимира Ульяновцева** (http://fppo.ifmo.ru/?page1=16&page2=52&page_d=1&page_d2=146361).

2009 год

А дело было так. **21-25 сентября** проходил Санкт-Петербургский научный форум «Наука и общество» по теме: «Информационные технологии». Это была IV Петербургская встреча лауреатов Нобелевской премии и других выдающихся ученых (http://is.ifmo.ru/important/nobel_open/), которую ежегодно проводил **Жорес Иванович Алферов** (<http://is.ifmo.ru/photo/2010-03-15-Alferov-80/index.html>).

Там **21 сентября** я «пристал» к **Анатолию Борисовичу Чубайсу**, которого ничуть не заинтересовали наши талантливые молодые программисты. После этого я выбрал другую «жертву» – декана биофака МГУ (в то время еще и председателя Высшей аттестационной комиссии), академика РАН **Михаила Петровича Кирпичникова**, который немного послушав, подвел меня к весьма вальяжному господину и сказал: «**Костя, поговори с товарищем, возможно, тебе будет интересно**».

Костя оказался основателем и директором Центра «Биоинженерия» РАН, директором-координатором по биомедицинским технологиям «Научно-исследовательского центра «Курчатовский институт», заведующим кафедрой биотехнологии и профессором кафедры генетики биологического факультета МГУ, академиком РАН **Константином Георгиевичем Скрыбиным** (http://www.rusrep.ru/2010/20/interview_skyrbin/) и не просто академиком, а академиком в третьем поколении (по-моему, в академии наук такого еще не было).

Его дед – **Константин Иванович Скрыбин** был академиком АН СССР, Академии сельскохозяйственных наук и Академии медицинских наук, что, как мне кажется, беспрецедентно. Он был Героем Социалистического труда, лауреатом Ленинской и Сталинских премий, **кавалером шести орденов Ленина**. С 1939 г. его имя носит Всероссийский научно-исследовательский институт гельминтологии (https://ru.wikipedia.org/wiki/Скрыбин,_Константин_Иванович).

Его отец – **Георгий Константинович Скрыбин** – был академиком АН СССР и Академии сельскохозяйственных наук, с 1967 г. по 1974 г. – директор Научного центра биологических исследований, а 1971 г. по 1988 г. – Главный ученый секретарь Президиума Академии наук СССР. Герой Социалистического труда, лауреат Государственной премии. Кавалер двух орденов Ленина. Его имя носит Институт биохимии и физиологии микроорганизмов РАН (https://ru.wikipedia.org/wiki/Скрыбин,_Георгий_Константинович).

Теперь о самом **Констанине Георгиевиче**, который, к сожалению, скончался **05.11.2019 г.** Он был академиком АН РАН и Российской академии сельскохозяйственных наук, лауреатом Государственной премии (https://ru.wikipedia.org/wiki/Скрыбин,_Константин_Георгиевич).

Вот, кстати, что однажды я прочел у **Константина Георгиевича**: «**Мне было стыдно не прочесть Толстого, Достоевского или Гете, так как иначе за ужином нам бы не о чем было разговаривать**». Узнав это, многие, наверное, подумают – нам бы его заботы. У меня на это ответ простой: станьте академиком в третьем поколении, и у Вас такие же проблемы появятся. Они могут еще больше усугубиться, если одним из родственников будет кто-то того же масштаба, что и выдающийся композитор **Александр Николаевич Скрыбин** – родственник академиков!

Теперь, как говорится, ближе «к делу». Наш разговор с академиком длился долго – академик любил поговорить не меньше меня. Мне и до этого разговора было известно, что академик занимается геномикой – беседа только подтвердила это, и поэтому я предложил: «Давайте работать вместе в области сборки генома: я постараюсь привлечь к этой работе наших выдающихся молодых программистов с кафедры «Компьютерные технологии» (КТ) Университета ИТМО, а Вы – молодых биологов. «Давайте», – ответил академик. Так я «ввязался» в проведение на кафедре «Компьютерные технологии» Университета ИТМО исследований по биоинформатике, а в дальнейшем и по системной биологии (<http://is.ifmo.ru/belletristic/Shalyto-moi-shastlivye-gody-na-CT.pdf>).

Когда я рассказал об этом в университете, надо мной стали смеяться, и мне было сказано: «Какой Вы биолог, какие у нас биологи, да и какие биологи есть в нашей стране, ерунда какая-то».

Но такая хрень на меня действовала мало. Я подобные разговоры слышал всю свою «сознательную» жизнь. Например, мне постоянно говорили в НПО «Аврора», где я работаю уже более 50 лет: «Если ты написал несколько статей, то почему считаешь себя ученым – работаешь у нас инженером – вот и будь им». Многого бы я добился, если бы всех их слушал. Ведь известно, что «собака – лает, ветер – носит, а караван – идет». Конечно, если всех верблюдов отравить или перестрелять, то караван остановится, а потом и загнется, но «верблюды»-то обычно живучи... Интересно, что Хрущев в этом выражении вместо слова «верблюд» использовал слово «конь», которое ему было значительно ближе.

Сегодня я могу сказать, что «меня многие много лет поливали, и это помогало мне успешно расти!», а главное даже не мне, а нашим ребятам, которые выросли, да еще как. А еще я всегда помню высказывание: «Они пытались закопать нас, но они не знали, что мы – семена». В дальнейшем оказалось, что мое настойчивое желание, чтобы мы занялись сборкой генома, оказалось для нашего коллектива – судьбоносным, так как уже через несколько лет, когда Университет вошел в программу по повышению международной конкурентоспособности российских вузов «5 в 100», выяснилось, что в рамках этой программы для существования и нормального финансирования международной научной лаборатории (МНЛ) «Компьютерные технологии», сформированной нами, требовались высокорейтинговые публикации, которые по *Computer Sciences* практически недостижимы – в этой области главное не журнальные статьи, а материалы конференций в основном с низкими рейтингами, в то время как в биоинформатике, а тем более в системной биологии, все наоборот.

Через некоторое время, несмотря на указанную выше реакцию моего окружения, мы, все-таки, начали обсуждение сотрудничества нашей кафедры и центра «Биоинженерия», в результате которого было решено совместно создавать технологию сборки генома, превосходящую по качеству или производительности (в идеале – по обоим критериям сразу) мировой уровень. Эта технология должна была включать алгоритмы сборки генома, так и реализующее их программное обеспечение для кластеров или суперкомпьютеров. Потом мы со своей стороны решили, что технология должна быть применима, как для высокопроизводительных вычислительных средств, так и для значительно менее производительных, например, для персональных компьютеров. Последнее оказалось для нас главным.

Сначала мне надо было найти «работников» с нашей стороны. Для того, чтобы не зависеть от прихоти молодых людей можно либо выбирать интересующихся, либо принуждать талантливых, если они хотя бы некоторое время от Вас зависят! Последнее утверждение многим не понравится, но последовательное применение этих подходов (от второго к первому) привело нас к успеху.

Я заметил, что отбор для занятий наукой среди студентов в некотором смысле сравним либо с наливанием пива, либо, к сожалению, шампанского, так как когда в кружке с пивом осядет пена, то в ней обычно остается процентов семьдесят напитка, а когда осядет шампанское – процентов пятнадцать.

Месяца полтора мы договаривались с сотрудниками Скрыбина об их приезде для мотивации наших ребят на проведение исследований в этой области. На первую встречу мне удалось собрать около 100 наших студентов и выпускников, а от академика Скрыбина к нам приехали трое, включая

Егора Борисовича Прохорчука (<https://openmedcom.ru/lectors/258>), который сегодня возглавляет лабораторию геномики и эпигеномики позвоночных в Федеральном исследовательском центре «Фундаментальные основы биотехнологии» РАН, а также является деканом медико-биологического факультета Российского национального исследовательского медицинского университета им. Н.И. Пирогова. Биологи три часа рассказывали про геном, а наши молодые люди с интересом слушали, **но слушать – не делать**, и после лекции мы получили нечто среднее между пивом и шампанским – в зале осталось человек двадцать пять – после того, как лекция закончилась, человек семьдесят пять сразу «смылось».

Двадцать пять молодых людей еще полчаса подискутировали с биологами, и еще человек десять ушло. У нас, наконец, появилось «шампанское» – осталось человек пятнадцать.

Эти молодые люди ничего не обещали, но говорили, что им интересно **позаниматься** сборкой генома, и они попробуют что-нибудь сделать. Мне всегда «нравились» слова с приставкой «по»: **поработать, позаниматься, поделать**. Я, как и поэт Павел Коган, «с детства не любил овал – я с детства угол рисовал», и поэтому в моем лексиконе вместо этих звучат другие слова: **работать, заниматься, делать**.

Отступить было некуда (иначе я бы оказался в глазах Скрябина – треплом), и поэтому быстро разбил оставшихся наших молодых людей на три группы (по числу гостей из Москвы) и был рад этому, так как знал, что **в одном из ведущих московских вузов после такой же беседы ушли ... все!** Это к вопросу о привлечении в науку молодых людей, интересующихся на начальной стадии занятий принципиально новой и актуальной тематикой.

Когда через пару недель мы сами провели семинар, «пена еще больше осела» – пришло человек восемь, а еще через некоторое время «на дне» осталось три-четыре «капли». Двое из оставшихся молодых людей сказали, что им интересно заниматься указанной темой, и что они, якобы, будут продолжать это делать, но никаких обязательств, **естественно**, брать на себя не стали, что характерно для многих молодых людей. Более того, ни от кого из них я даже не услышал: «Дайте денег, и все будет хорошо», как иногда говорят сегодня.

Недели через две-три стало ясно, что **позаниматься наукой и заниматься ею** – две большие разницы :-), и что мы недалеко ушли от того московского вуза, в котором после беседы про геном ушли все!

Это окончательно показало, что мифом является то, что «дети» хотят заниматься наукой: **если раньше многих не устраивало предложенное мною автоматное программирование** (<http://is.ifmo.ru/works/2008/Vestnik/53/01-automata-based-programming.pdf>), **что можно было понять, то теперь молодежи была неинтересна уже сборка генома...** Что им интересно? Видимо, только то, что придумают сами и то неизвестно на долго ли, но ждать, что они придумают сами заняться тем, о чем мы договорились со Скрябиным – **век жизни не ведать...**

В общем, я понял, что «пахнет жареным», и **пора переходить от уговоров к принуждению**, иначе через пару встреч я останусь один на один ... с геномом, и это несмотря на то, что указанная тематика в то время практически всеми экспертами в мире признавалась одной из самых перспективных. Это, кстати, не оспаривал никто из «моих» молодых людей, но и что-то делать в этой области, тем более делать что-то серьезное, никто из них не торопился. Последнее, видимо, связано с тем, что многие из них являются буддистами – займется этим со всем энтузиазмом в следующей жизни.

В общем, я, **как и Печорин, безуспешно пытался «влезть в жизнь честных контрабандистов»**, но всегда помнил о «тиграх», которые не хотят прыгать сквозь огненные обручи, но, все-таки, прыгают... Также я подумал, что если существуют операции «по принуждению к миру», то почему не осуществить **операцию «по принуждению к творчеству»**.

2010 год

У нас сложились условия для успешного проведения этой операции: на кафедре были очень талантливые молодые люди, которым в этом году **необходимо** было защитить две бакалаврские работы и магистерскую диссертацию!

Сначала я из немногочисленных оставшихся «героев» выбрал первую жертву – молодого человека, который считался очень умным. К тому же он был чемпионом мира по программированию и имел в этой области один из самых высоких рейтингов в мире. Кроме того, он говорил, что сборка генома его интересует, но главное было в том, что ему менее чем через полгода надо было защищать бакалаврскую работу. Потом к нему подтянулись еще двое героев (обратите внимание последнее слово я в кавычки не беру) – один из них тоже был чемпионом мира по программированию, а второй – призером чемпионата мира, причем первому из них тоже нужно было защитить бакалаврскую работу, а второму – магистерскую диссертацию.

Это были **Владислав Исенбаев, Евгений Капун и Михаил Дворкин**. Для того чтобы продемонстрировать молодым людям серьезность моих намерений, я впервые в жизни и специально для них написал два коротких, но очень проникновенных :-)) стихотворения. Первое – мотивирующее: «**Будет геном – будет диплом**», а второе – стимулирующее: «**Не будет генома – не будет диплома**» (http://is.ifmo.ru/belletristic/genom_belletristic).

Несмотря на то, что я не являюсь любителем Сталина, его подходы, видимо генетически :-)), действует и на меня, ведь, не зря сталинисты его, а также Берину, называют **эффективными менеджерами** (такими эффективными, что не дай Бог с ними встретиться даже в дурном сне). Вот один из примеров, который, якобы, имел место во время войны: «Сталин поручил заместителю народного комиссара нефтяной промышленности **Николаю Байбакову** открыть новые нефтяные месторождения. Когда Байбаков возразил, что это невозможно, Сталин ответил: «**Будет нефть, будет Байбаков, не будет нефти, не будет Байбакова!**». Вскоре были открыты месторождения в Татарии и Башкирии». Похоже на мои «стихи»?

Молодые люди поверили мне :-)), и, в конце концов, все для всех закончилось благополучно. При этом успешно были защищены следующие квалификационные работы: **Дворкин М.Э.** Методы минимизации необходимого числа цепей для секвенирования ДНК. Магистерская диссертация. СПбГУ ИТМО. 2010, 44 с. (http://is.ifmo.ru/papers/dvorkin_genom.pdf), **Исенбаев В.В.** Разработка системы секвенирования ДНК с использованием *paired-end* данных. Бакалаврская работа. СПбГУ ИТМО. 2010, 30 с. (http://is.ifmo.ru/genom/isenbaev_thesis.pdf) и **Капун Е.Д.** Разработка метода сравнения нуклеотидных последовательностей путем разбиения на фрагменты. Бакалаврская работа. СПбГУ ИТМО. 2010. 26 с. (http://is.ifmo.ru/genom/kapun_thesis.pdf). При этом ребята обещали продолжить работу с начала нового учебного года. Верилось с трудом, но, что мне оставалось делать – только надеяться, так как говоря по-сталински, «другого народа у меня не было».

Через некоторое время после защит дипломов к нам в Университет ИТМО приехали на предмет сотрудничества профессора одного из лучших в мире технических университетов – **Массачусетского технологического института (МТИ)**. В это время предполагалось, что российские университеты будут дружить с МТИ. Поэтому группа профессоров этого института **22 июня 2010 г.** посетила наш университет (<http://is.ifmo.ru/photo/2010-06-22-meeting-MIT/index.html>). Наше руководство уже было наслышано о дипломах наших «вундеркиндов», и один из докладов перед гостями наряду с докладами профессоров нашего университета было предложено сделать новоявленному магистру Мише Дворкину, который ко всему еще и свободно говорит по-английски, так как заканчивал школу в Нью-Йорке несмотря на то, что до этого учился в Санкт-Петербурге. Кстати, Миша мог сделать доклад и на эсперанто, но гости в нем были слабы...

У нас на кафедре ответственным за выполнение работ по этой тематике был **Георгий Корнеев**, совместно с которым мы с ребятами подготовили презентацию. Так как они осуществили очень быстрый «вход в геном», то мы также назвали и презентацию: **Dvorkin M., Isenbaev V., Kapun E. Genomics Quick Start. SPbSU ITMO. 21.06.2010. Scientific advisors: Acad. Konstantin Skryabin (Bioengineering RAS) & Prof. Anatoly Shalyto (SPbSU ITMO)** (<http://is.ifmo.ru/present/mit.ppt>), в которой объяснили как и почему они так быстро в новой для себя области добились хороших результатов.

Когда Миша закончил выступление, раздались аплодисменты. Видимо, для **большого укрепления дружбы :-))** один из приехавших профессоров (русскоязычный) сразу же стал сманивать Мишу к себе в аспирантуру. Я попытался объяснить ему, что сотрудничество между нашими университетами не сводится к сманиванию наших талантов, и предложил ему ставить задачи, которые ребята, возможно даже бесплатно, будут решать для нашего совместного блага... здесь. Мое

предложение профессору не понравилось, и он продолжил уговаривать Мишу, сказав мне, что тот после аспирантуры вернется в Россию. Я задал ему только один вопрос «А Вы вернетесь?». Это, почему-то, обидело гостя, и он свернул разговор, хотя визитные карточки все-таки дал: мне – просто так, а Мише – с надеждой. Вот и вся дружба! **Так закончилось наше сотрудничество с великим вузом**, правда, оказалось, что через несколько лет мы независимо от этого профессора стали сотрудничать *Broad Institute of MIT and Harvard*, что еще более круто, если учесть с кем мы там стали работать, но об этом позже.

Естественно, что тогда ни мы, ни сманивавший Мишу профессор, этого, естественно, не знали. Предложение американского профессора не заинтересовало Мишу (<http://www.school.ioffe.ru/school/news/article.html?692>, <http://www.school.ioffe.ru/staff/person.html?id=162>), так как у него были «свои счета» с МТИ: он был призером двух олимпиад Нью-Йорка и двух олимпиад – Америки и хотел, в частности, поступить в МТИ, но так как его тогда не приняли, то он, как Примаков (<https://ruspek.ru/events/evgeniy-primakov-razvernul-samolet-nad-atlantikoy>), развернул «самолет над океаном», вернулся в Россию и поступил к нам – в Университет ИТМО, благо, как один из победителей Всероссийской олимпиады школьников по информатике 2003 г. (<https://www.fontanka.ru/2003/04/14/47441/>), имел право на внеконкурсный прием. Второй раз входить в «одну и ту же воду» – снова поступать в МТИ, он не захотел. В том же году (2010 г.) он был принят в аспирантуру к нам с темой про геном и со мной в качестве научного руководителя.

Кстати, до этого отклонил предложение поступать в аспирантуру другого всемирно известного вуза – *ETH* (Цюрих) – Федя Царев. В общем, можно утверждать, что Федя и Миша поступили, как одна женщина, которая сказала: «**Зачем мне путешествовать, если я уже здесь?**».

В заключение истории с МТИ отмечу, что, как будет описано ниже несмотря на то, что сотрудничество с этим великим вузом у нас не получилось, этот вуз «обеспечил» возможность нашего знакомства с ... Максимом Артемовым!

Лето закончилось, закончилась осень и наступил декабрь, но про геном никто из молодых людей больше не вспоминал! Мотивирующая сила моих «стихов», как антител после короновируса, давно иссякла – ребята отбыли свой номер... Тем более, что они защитили свои дипломы. При этом Исенбаев не пошел в магистратуру и начал работать в промышленности, а Дворкин – окончил университет, и несмотря на то, что он поступил в аспирантуру, основным местом его работы уже тогда была физико-техническая школа при Академическом университете (АУ) – честь ему и хвала за это. Эту школу он свое время закончил.

Я вновь позвонил Скрыбину, и через несколько дней к нам вновь приехал Прохорчук, который провел встречу с небольшим числом, якобы, еще желающих собирать геном. При этом он сказал, что дальнейшая наша совместная деятельность будет интересовать их только в том случае, если к 15.02.2011 г. (потом этот срок перенесли на первое марта) **мы соберем тестовый искусственный геном в 1,8 млрд нуклеотидов**, причем подведение итогов этой сборки произойдет в апреле в Барселоне. Таким образом, **вопрос перед нами был поставлен ребром: или мы продолжаем сотрудничать, но по-настоящему, или на этом – все.**

Георгий Корнеев (он, как отмечено выше, в то время возглавлял нашу деятельность в области сборки геном) попросил меня и Прохорчука удалиться и провел с ребятами совещание. Решение было для меня неутешительным, так как они, и не без оснований, считали, что **уж больно неудобное сейчас время**. В подтверждение этого после моего возвращения было **озвучено несколько причин, которые не позволяли им эту работу продолжить**: во-первых, Новый год, во-вторых, сессия, в-третьих, студенческие сборы по программированию в Петрозаводске и наша Летняя Компьютерная Школа (в полном смысле последнего слова, так как она предназначена для повышения квалификации одаренных школьников), традиционно проводимая еще и ... зимой.

Особо хочу отметить позицию, которую занял **Миша Дворкин, принявший решение больше не заниматься геномом**, о чем через несколько дней сказал **Андрею Станкевичу**. Интересно, что до этого **Миша произнес фразу, повергшую меня в шок**: он не понимал, зачем писать статьи. Интересно, что это высказывание последовало почти сразу после опубликования им и с еще одним призером чемпионата мира по программированию – **Искандером Акишевым**, статьи в весьма солидном математическом журнале, выпускаемом Томским государственным университетом, которая была представлена туда «с моей подачи»: *Акишев И.Р., Дворкин М.Э.* О построении

минимальных детерминированных конечных автоматов, распознающих префиксный код заданной мощности // Прикладная дискретная математика. 2010. № 2, с. 104-116 (<http://www.lib.tsu.ru/mminfo/000349342/08/image/08-104.pdf>). Этот научный инфантилизм меня удивил лишь частично, так как в свое время Георгий Корнеев сомневался в необходимости ссылок на предшественников – придумал нечто и ладно. Надо отметить, что эта «детская болезнь» у Гоши давно прошла.

Мне в пору было сдать, и если бы я поступил так, то нашей международной научной лаборатории «Компьютерные технологии», созданной через несколько лет после этого в рамках программы «5 в 100», давно бы уже не было, так как по числу публикаций требования этой программы мы бы выполнили с помощью работ по *Computer Science*, а вот по показателям качества источников, где эти работы публикуются – нет, так как в этой области они существенно ниже, чем в биоинформатике, а тем более, чем в системной биологии. Естественно, что из-за этого не было бы и большинства ребят, работающих в лаборатории сейчас.

Однако я сдаваться и на этот раз не стал, так как вспомнил, как я и мой ученик, с которым мы вместе работали в НПО «Аврора», опубликовали четыре классные статьи по одной и той же тематике в одном из лучших российских журналов – «Автоматика и телемеханика», который переводится на английский язык, и начали писать пятую. Но в это время молодому человеку предложили высокооплачиваемую работу, он долго не раздумывал и уволился с нашего предприятия. На этом наша совместная деятельность закончилась.

Примерно через год он позвонил мне и сказал, что попал в больницу, но чувствует себя хорошо и готов продолжить писать незавершенную статью. Последнее меня, естественно, обрадовало, и я подключился к работе. Затратив много усилий, мы завершили статью, и стали ждать из журнала вердикт. Он, естественно, последовал, но был отрицательным, так как за это время те же результаты были получены и опубликованы другими людьми.

Оказывается, пока мы меняем работу, женимся, отдыхаем, болеем **мир не останавливается и не ждет**, когда мы соблаговолим вернуться к тому, чем занимались до этого. Причина, из-за которой мы остановились, была объективной, но, как оказалось, только для нас...

Приведу по этому поводу одну из «Заметок о мотивации» (http://is.ifmo.ru/belletristic/zametki_o_motivacii.pdf): «Не сделаешь ты, сделает кто-то другой. Не напишешь ты, напишет кто-то другой. Не будешь рядом ты, будет кто-то другой. Не возьмешь все в свои руки, возьмет кто-то другой. **Всегда будет кто-то другой**». Никогда не говорите, что «это не зависит от Вас». Если не от Вас, так от кого же? И помните слова **Феликса Кривина**: «**Там, где каждый считает, что лично он не делает погоды, погода бывает самая отвратительная**». Если Вам кто-то с пафосом сказал, «**если не мы, то кто же**», но ничего не сделал, знайте, что пафос легко сбивается словами: «**Если не Вы, то другие**». И помните, что **свято место пусто не бывает**».

Как следует из изложенного, похожая история чуть не случилась у нас и со сборкой генома, переросшей в дальнейшем в суперуспешные исследования по биоинформатике и системной биологии под руководством Леши Сергушичева.

А тогда отказ работать в рассматриваемой области (**все свалили!**) привел меня в бешенство, кровь ударила в голову, и я на повышенных тонах прочел присутствующим стихотворную строку – на этот раз не свою, а поэта **Александра Кушнера**, которая была приведена в самом начале этого текста.

На этом все должно было бы закончиться, но неожиданно ко мне подошел **Андрей Станкевич** и, волнуясь, сказал, что возглавит работу по сборке генома и начнет руководить в указанном направлении **двумя Антонами (Ахи и Банных) и Сергеем (Поромовым)** на сборах в Петрозаводске. Я удивился, но стало немного легче. Однако уверенности, что дело будет сделано, у меня не появилось.

2011 год
Миша Дворкин

В 2009 г. на кафедре **на постоянной основе остался работать** чемпион мира по программированию **Федор Царев**, который у нас работал с 2007 г. Он стал вторым чемпионом мира (после Павла

Маврина), работающим на кафедре, и помогал мне в организации научных исследований. В середине новогодних каникул он позвонил и сказал, что **Миша Дворкин передумал...**

Оказалось, что Михаил, все-таки, решил продолжить заниматься сборкой генома, но ... не с нами, а в лаборатории «Алгоритмическая биология» Академического университета (АУ), которой по мегагранту руководил выдающийся в этой области ученый **Павел Аркадьевич Певзнер** из Калифорнийского университета в Сан-Диего (Компо Ф., Певзнер П. Алгоритмы биоинформатики. М.: ДМК-Пресс. 2023 г., 682 с., <https://www.labirint.ru/books/931230/>). Однако Миша не поставил нас в известность о своем уходе, что не делает ему чести. При этом я продолжал оставаться научным руководителем этого умного и очень вежливого мальчика, у которого эти свойства, похоже, подменили какие-то другие человеческие качества – всегда помню фразу: **«По сути, вежливость и приличия – это частный случай лицемерия».**

Прошло полгода, а Миша, как страус, прятал «голову» и, видимо, думал, что его «не видно», и за это время ни разу даже не позвонил. Я понимаю, что это не телефонный разговор, но мужества поговорить «с глазу на глаз» у него, видимо, не хватало. Интересно, что, **общаясь со мной несколько лет до этого, он так и не осознал, что я могу понять его:** Миша преподает информатику в лицее, который располагается в помещении АУ, и ему очень удобно там же заниматься наукой, за которую ему платили значительно больше, чем в то время могли платить мы.

Однако, возможно, он вообще не считал, что со мной следует об этом поговорить, хотя ничего плохого я ему не сделал, а, наоборот, старался, как мог, помочь ему, например, «святая» его при всех в аспирантуру к **Дэвиду Хопкрофту** (http://is.ifmo.ru/important/hopcroft_itmo.pdf) или предлагая ему создать в Университете ИТМО совместно с **Сергеем Евгеньевичем Рукшиным** центр подготовки одаренных детей в области информатики и математики. А еще я помню, как в сентябре 2011 г. при моем появлении на обеде на соревнованиях по программированию *Russian Code Cup* в Москве, Миша, как ни в чем не бывало, поздоровался со мной на весь зал.

Через какое-то время Миша перестал заниматься наукой и там. Сегодня – он преподаватель программирования в лицее «Физико-техническая школа» при АУ.

Формирование новой команды

После того, как молодые люди, занимавшиеся сборкой генома, которых к тому времени практически не осталось, сообщили мне, что время для сборки генома неудачное :-), я с помощью **Феди Царева** решил **подключить к решению этой задачи более надежную «публику»** – команду в составе: **Алексей Сергушичев, Сергей Казаков и Антон Александров**, которая достаточно хорошо выступала на полуфинальных соревнованиях чемпионата мира по программированию.

Кроме того, эти ребята прекрасно зарекомендовали себя при выполнении под моим и Феиным руководством курсовой работы по автоматному программированию, в которой они на основе тестов с помощью генетического программирования строили автоматы, управляющие режимами модели самолета, учитывая при этом не только дискретные, но и непрерывные параметры.

Важно отметить, что я мучил этих ребят почти четыре года (!), пока мы писали, все-таки дописали и опубликовали классную статью: **Aleksandrov A.V., Kazakov S.V., Sergushichev A.A., Tsarev F.N., Shalyto A.A.** The Use of Evolutionary Programming Based on Training Examples for the Generation of Finite State Machines for Controlling Object with Complex Behavior // Journal of Computer and Systems Sciences International. 2013. № 3, pp. 410-425 (<https://link.springer.com/article/10.1134/S1064230713020020>). (**Александров А.В., Казаков С.В., Сергушичев А.А., Царев Ф.Н., Шальто А.А.** Применение эволюционного программирования на основе обучающих примеров для генерации конечных автоматов, управляющих объектами со сложным поведением // Известия РАН. Теория и системы управления. 2013. № 3, с. 85-100 (http://is.ifmo.ru/works/2013/alexandrov_samolet.pdf).

При этом отмечу, что ребята **начали работать над статьей в рамках моего курсовика на третьем курсе в 2009 г., а статья вышла в свет после защиты ими магистерских диссертаций в 2013 г., и все это время они что-то делали в указанном направлении!**

Однако пойти на «бой» с геномом их надо было еще уговорить... При этом надо отметить, что в начале четвертого курса в 2010 г. Алексей Сергушичев сказал мне, что собирается писать бакалаврскую работу по когнитивным компиляторам у русского профессора из Финляндии. Я

спросил Лешу, что этот профессор обещает ему, и с удивлением узнал: «Пятьдесят долларов за решенную задачу». «А аспирантуру или работу он тебе обещает?», – не унимался я. Оказалось – ничего не обещал, «только тихо ботами качал». После этого я посоветовал Леше забыть об этом «блестящем предложении» и подумать о сборке генома, которой в течение 2010 г. В. Исенбаев, Е. Капун и М. Дворкин занимались в рамках выполнения своих квалификационных работ.

Мне потребовалось достаточно много времени для того, чтобы уговорить Алексея заняться сборкой генома. Во время «уговоров» я привел ему слова из «Заметок о мотивации»: «Станьте как лазер: копите энергию и выстреливайте ее узким лучом в цель. Не растрачивайте себя на многие дела сразу – **не будьте похожими на душлаг, из которого через скользкие макароны в разные стороны без напора вытекает теплая и мутная вода.** И помните, что лазеру тоже хочется расслабиться, но он не может – ему надо торопиться к цели :-)). В конце концов, Алексей согласился заняться сборкой генома, и, **как говорит сегодня, ни разу не пожалел об этом.**

После этого уговорить Антона с Сергеем заняться этой тематикой было уже проще, при этом главный мой довод состоял в том, что ими снова будет руководить Федор Царев, с которым у ребят при выполнении курсовика сложились прекрасные отношения. Самого Федю руководить этими ребятами и возглавить работу упрощать не пришлось!

Снова вспоминаю одну из «Заметок»: «**Никогда, никогда не сдавайтесь** (угрозу жизни и пытки я не обсуждаю). Если же Вам, все-таки, придет мысль сдаться, вспомните, ради чего Вы все начинали... **Идите до конца.** Вы можете **потерять все**, но, если выиграете... **Развивайте в себе бойцовские качества и самообладание.** Помните, что «устают все, но люди с бойцовским характером, сначала делают работу и только потом устают».

О сборке генома – сначала искусственного

После того, как новый коллектив был сформирован началась разработка алгоритмов сборки генома, состоящей из четырех этапов: **1.** Исправление ошибок **в ридх** – данных, поступивших с секвенирующей машины; **2.** Восстановление фрагментов геномной последовательности длиной примерно по 500 нуклеотидов (**сборка квазиконтигов**) на основе исправленных ридов; **3. Сборка контигов** – длинных непрерывных фрагментов геномной последовательности (для этого этапа тогда использовался сборщик *Newbler*, предоставленный центром «Биоинженерия»); **4.** Дополнительный, оптимизирующий этап: определение взаимного расположения контигов друг относительно друга (**построение скэффолдов**) – для этой цели применялся один из модулей *Open-Source* сборщика *AbySS* (<https://www.bcgsc.ca/resources/software/abyss>).

При этом отмечу, что алгоритмы сами по себе никого не интересовали – необходимо было срочно (дедлайн в соревнованиях в Барселоне был близок) не только разработать алгоритмы для каждого из этапов, но и программно реализовать каждый из них. Это осложнялось тем, что в момент начала работ ребята были на каникулах – каждый на своей малой Родине, а наличие у них ноутбуков было не тем оружием, чтобы собрать искусственный геном в 1,8 млрд нуклеотидов, однако Егор Борисович Прохорчук предоставил им доступ к суперкомпьютеру (не знаю был ли у них до этого опыт работы на таком вычислителе).

Федина команда, к которой подключились **Сергей Мельников, Паша Федотов** и частично Слава Исенбаев, в январе закончила **первый этап** создания сборщика генома – **разработала и реализовала алгоритм исправления ошибок в ридх**. Запуск этой программы проводился на суперкомпьютере центра «Биоинженерия» (24 гигабайта оперативной памяти, два четырехъядерных процессора). При этом объем входных данных составлял порядка 250 гигабайт. В результате трех запусков программы, занявших суммарно примерно 60 часов, было исправлено около 18,5% исходных ридов, и появилась возможность перейти ко второму этапу.

В феврале команда Царева завершила второй этап – **разработала и реализовала алгоритм восстановления фрагментов геномной последовательности (осуществила сборку квазиконтигов)**. Это было выполнено на основе построения и обхода графа де Брейна (https://en.wikipedia.org/wiki/De_Bruijn_graph). На этот раз запуск разработанной программы проводился на суперкомпьютере, предоставленном компанией «Т-Платформы» (64 гигабайта оперативной памяти, два двенадцатиядерных процессора). В результате работы программы в течение 40 часов были сформированы данные, которые были поданы на вход сборщика *Newbler* –

начался третий этап сборки (сборка контигов), продолжавшийся около 60 часов. Четвертый этап мы тогда не проводили.

Полученные результаты показали работоспособность и перспективность разработанных нами алгоритмов и программ для отдельных этапов сборки генома.

Итак, **в нашем первом биоинформатическом проекте** ответственным исполнителем от кафедры КТ был Ф. Царев, а от центра «Биоинженерия» – Е. Прохорчук. Исполнителями были А. Сергушичев, А. Александров, С. Казаков, П. Федотов и С. Мельников.

В апреле 2011 г. наша команда из пяти человек (трое из них ездили за счет нашего университета) приняла участие в семинаре, организованном национальным центром геномного анализа в Барселоне, на котором подводились предварительные итоги выполнения проекта *De novo genome assembly project (dnGASP)*. В работе семинара **участвовали девять исследовательских центров**, таких как, например, *Beijing Genomics Institute* (Пекинский геномный институт, Китай), *European Bioinformatics Institute* (Европейский институт биоинформатики, Великобритания) и *Canada's Michel Smith Genome Science Centre* (Канадский институт геномных исследований, Канада). **Сравнение полученных результатов показало, что наши ребята оказались на уровне с «сильными мира сего».**

По результатам семинара всеми его участниками (32 исследователя) была написана статья, которая была опубликована в 2011 г. на сайте *Centre National de Rechercheen Genomique Humaine* (Франция): *Alexandrov A., Fedotov P., Kazakov S., Melnikov S., Sergushichev A., Tsarev F. et al.* The de novo Genome Assembly Assessment Project – *dnGASP* (https://www.cnrg.fr/READNA/meetings/downloads/READNA_Deliverable_6.8_Manuscript.pdf).

Среди соавторов – Е.Б. Прохорчук.

Сборка генома рыбы

В то время сборкой генома, кроме организаций, принявших участие в указанном выше проекте, занимались еще несколько американских университетов, которые, как и мы, **в июне 2011 г.** приняли участие в соревнованиях *Assemblathon-2* (<https://assemblathon.org/assemblathon2>) по сборке генома рыбы (в соревновании *Assemblathon-1* мы не участвовали). По результатам этих соревнований в 2013 г. (см. ниже) была опубликована статья, авторами которой стали все участники проекта, в том числе и наши ребята.

В этих соревнованиях мы участвовали без представителей центра «Биоинженерия». В то время у нас уже был свой «не первой свежести» суперкомпьютер (32 гигабайта оперативной памяти, восемь процессорных ядер), который я «выклянчил» у технического директора компании *Mail.ru* Владимира Александровича Габриеляна (http://is.ifmo.ru/photo/2011-09-18-Russian-code-cup/IMG_0421.jpg) в ходе подготовки к проведению соревнования по программированию *Russian Code Cup 2011*.

Наше успешное участие в этом и предыдущем проектах уже в то время **позволило говорить, что Университет ИТМО стал одним из немногих университетов и исследовательских центров мира, обладающих технологией сборки геномных последовательностей** на основе данных о чтении на секвенаторах второго поколения.

Мы выиграли два крупных Всероссийских конкурса, что было абсолютной сенсацией Первый грант (27.04.2011 г.)

2011-2013. Федеральная целевая программа «Научные и научно-педагогические кадры инновационной России» на 2009–2013 годы. «Проведение научных исследований научными группами под руководством докторов наук». Открытый конкурс на выполнение НИР в следующих областях: биокаталитические, биосинтетические и биосенсорные технологии; биомедицинские и ветеринарные технологии жизнеобеспечения и защиты человека и животных; геномные и постгеномные технологии создания лекарственных средств; клеточные технологии; биоинженерия; биоинформационные технологии. Тема: **«Разработка метода сборки геномных последовательностей на основе восстановления фрагментов по парным чтениям».** Руководитель – А.А. Шалыто.

Когда мы подавали документы на этот грант, наш ректор **Владимир Николаевич Васильев**, которому предстояло подписать около 300 страниц пояснительной записки, как ему, видимо, тогда казалось на абсолютно безнадежное начинание, спросил меня: «**А кто у Вас является доктором биологических наук? Не Вы ли, Анатолий Абрамович?**». На это я ответил, что в условиях этого конкурса нет ограничения на специализацию доктора наук, который руководит группой исследователей – руководителем должен быть доктор наук и не обязательно биологических. Мой ответ не внушил ректору уверенности в нашей победе :-), но он, все-таки, подписал заявку: по условиям конкурса – каждую страницу.

Но, оказалось, что ректор не зря мучился (известно же, что, если долго мучиться, то что-нибудь получится :-)) – **мы заняли ... первое место в стране и выиграли этот грант** в условиях огромной конкуренции (были поданы заявки от более, чем ста организаций). Потом мы подавали заявки на разные гранты, в том числе и по биологии, но меня больше никто не спрашивал доктором каких наук я являюсь. Сегодня я мог быть сказать – «**психологических**», так как **понимаю на каких ребят надо «ставить», и с кем и что мы можем выиграть!**

Здесь приведены ссылки на отчеты по каждому этапу этой темы:
<http://is.ifmo.ru/bioinformatica/fragments/report-310276-1.pdf>,
<http://is.ifmo.ru/bioinformatica/fragments/report-310276-2.pdf>,
<http://is.ifmo.ru/bioinformatica/fragments/report-310276-3.pdf>,
<http://is.ifmo.ru/bioinformatica/fragments/report-310276-4.pdf>.

Второй грант (01.08.2011 г.)

Очень скоро мы вновь **удивили всех, в том числе и себя**, войдя в число победителей по второму гранту, причем за нами остался МГУ со своими биологами и суперкомпьютерами, который не вошел в число победителей – побеждали заявки, занявшие только первое и второе места.

2011-2013. Федеральная целевая программа «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2013 гг.» «Проведение проблемно-ориентированных поисковых исследований и создание научно-технического задела по перспективным технологиям в области информационно-телекоммуникационных систем». Тема: «**Разработка алгоритмов сборки геномных последовательностей для систем эксафлопсного уровня производительности**». Руководитель – А.А. Шалыто.

Статья

Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Царев Ф.Н., Шалыто А.А. Метод исправления ошибок в наборе чтений нуклеотидной последовательности // Научно-технический вестник информационных технологий, механики и оптики. 2011. Вып. 5, с. 81-84 (<https://ctlab.itmo.ru/~svkazakov/papers/alexandrov-vestnik-itmo-2011-5.pdf>).

Конференции

1. Сергушичев А.А., Исенбаев В.В., Царев Ф.Н., Шалыто А.А., Прохорчук Е.Б. Разработка метода восстановления фрагментов нуклеотидных последовательностей по парным чтениям / Сборник «СПИСОК-2011». Материалы второй межвузовской научной конференции по проблемам информатики». СПб.: ВВМ. 2011, с. 320-325 (<http://is.ifmo.ru/works/2011/SPISOK/01-Sergushichev-Isenbaev-Tsarev-Prohorchuk-Shalyto.pdf>).

2. Александров А.В., Казаков С.В., Мельников С.В., Царев Ф.Н., Шалыто А.А., Прохорчук Е.Б. Разработка метода удаления ошибок из набора чтений нуклеотидной последовательности / Сборник «СПИСОК-2011». Материалы второй межвузовской научной конференции по проблемам информатики». СПб.: ВВМ. 2011, с. 326-329 (<http://is.ifmo.ru/works/2011/SPISOK/02-Aleksandrov-Kazakov-Melnikov-Prohorchuk-Tsarev-Shalyto.pdf>).

3. Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Царев Ф.Н. Метод сборки геномных последовательностей на основе совместного применения графов де Брейна и графов перекрытий / Тезисы II Международной научно-практической конференции «Постгеномные методы анализа в биологии, лабораторной и клинической медицине: геномика, протеомика, биоинформатика». Новосибирск. 2011. Т. 2, с. 188.

4. Александров А.В., Казаков С.В., Сергушичев А.А., Мельников С.В., Исенбаев В.В., Царев Ф.Н. Метод сборки генома с помощью восстановления его аргументов по парным чтениям / VIII

Всероссийская межвузовская конференция молодых ученых. СПбГУ ИТМО. 12-15 апреля 2011. III сессия научной школы «Технологии программирования и искусственный интеллект».

Свидетельство о регистрации программы

Александров А.В., Исенбаев В.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Царев Ф.Н.
Программное средство для удаления ошибок из набора чтений нуклеотидной последовательности
// Свидетельство о регистрации программы для ЭВМ. № 2011 614454. Дата регистрации – 06.06.2011. (http://is.ifmo.ru/certificates/2011_06_06_genom_certificate.pdf).

Бакалаврские работы

1. *Александров А.В.* Разработка метода удаления ошибок из набора чтений нуклеотидной последовательности. Бакалаврская работа. Университет ИТМО. 2011.
2. *Сергушичев А.А.* Разработка метода восстановления фрагментов нуклеотидной последовательности по парным чтениям. Бакалаврская работа. Университет ИТМО. 2011 (<http://is.ifmo.ru/diploma-theses/2011-bachelor-sergushichev/>).

Эксперимент в области экстремального программирования

В ходе проведения исследований по сборке генома нами осуществлялся эксперимент по применению предложенной мною новой разновидности экстремального программирования – **программирование олимпиадными командами**, при котором одна или несколько команд, успешно выступающих на соревнованиях по спортивному программированию, на «тренировках» и после них совместно с тренерами и представителями предметной области решают и обсуждают, завершив многочасовой «мозговой штурм», решение не многих олимпиадных задач, как это принято на соревнованиях по спортивному программированию, а **только одной задачи, которая является научной**.

Есть основания надеяться, что многолетние тренировки, командный дух и наличие у участников «одной крови» могут позволить решать задачи, основанные на знании алгоритмов дискретной математики более эффективно, чем при традиционной форме работы. Об этом нами **первого ноября** был сделан доклад на конференции «Разработка ПО 2011» (*CEE-SECR 2011*) в Москве: *Царев Ф.Н., Шалыто А.А.* Программирование олимпиадными командами как разновидность экстремального программирования для решения задач дискретной математики на примере сборки генома (<http://is.ifmo.ru/present/2011/tsarev-shalyto-secr-2011.pdf>). Результаты участия в соревнованиях по сборке генома в течение первых двух месяцев проведения эксперимента, в которые входили новогодние праздники, каникулы и сессия, свидетельствуют об эффективности предложенного подхода к решению сложных задач указанного класса.

Федор Царев рассказал о сборке генома Президенту РФ

29.10.2011 г. Федор Царев рассказал, в частности, о сборке генома (<http://is.ifmo.ru/aboutus/2011/2011-10-2011-Tsarev-Medvedev.pdf>) Президенту РФ в Московской школе экономики и управления «Сколково», где происходила его встреча с молодыми учеными, инноваторами и предпринимателями (<https://www.youtube.com/watch?v=4zgVcQuzq9E>).

Решение Ученого совета

27.12.2011 г. решением Ученого совета Университета ИТМО было утверждено создание лаборатории «**Алгоритмы сборки геномных последовательностей**» при кафедре КТ. Сайт лаборатории: <http://genome.ifmo.ru/>. Этот адрес сейчас автоматически преобразуется в такой: <http://ctlab.ifmo.ru/>.

Часть 2. 2012-2016 гг.

<https://vk.com/@1077823-chast-2-2012-2016-kak-bioinformatika-i-sistemnaya-biologiya>

2012 год

Выигрыш еще двух грантов

Первый грант

2012, 2013. Государственное задание Министерства образования и науки РФ. Тема: «**Разработка алгоритмов генетического программирования и сборки генома**». Руководитель – А.А. Шалыто.

Второй грант

2012, 2013. Федеральная целевая программы «Научные и научно-педагогические кадры инновационной России» на 2009–2013 годы. «Поддержка исследований, проводимых коллективами **научно-образовательных центров**» по научному направлению «Науки о жизни. Живые системы» в области «Геномные, протеомные и постгеномные технологии». Тема: «**Разработка методов сборки генома, транскриптома и динамического анализа протеома**». Руководитель – А.А. Шалыто.

Статьи

1. **Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А. Царев Ф.Н.** Метод сборки контигов геномных последовательностей на основе совместного применения графов де Брейна и графов перекрытий // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2012. № 6, с. 93-98. (<https://ctlab.itmo.ru/~svkazakov/papers/alexandrov-vestnik-itmo-2012.pdf>).

В статье предложен **метод сборки контигов геномных последовательностей**. Его особенностью является **разбиение процесса сборки контигов на два этапа – сборка квазиконтигов из чтений и сборка контигов из квазиконтигов**. На первом этапе используется **граф де Брейна, а на втором – граф перекрытий**. Экспериментальное исследование метода проводилось на чтениях генома рыбы *Maylandia zebra*, размер генома которой примерно миллиард нуклеотидов. Преимущество разработанного метода состоит в том, что для его работы **требуется существенно меньше оперативной памяти** по сравнению с существующими программными средствами для сборки генома.

2. **Сергушичев А.А., Царев Ф.Н.** Сборка генома и технология MapReduce // Суперкомпьютеры. 2012. № 4, с. 40-43. (<https://is.ifmo.ru/works/2012/supercomputers-genome-mapreduce.pdf>).

Конференции

1. **Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Царев Ф.Н., Шалыто А.А.** Метод сборки контигов геномных последовательностей на основе совместного применения графов де Брейна и графов перекрытий / I Всероссийский конгресс молодых ученых. III сессия «Технологии программирования и искусственный интеллект». НИУ ИТМО. 2012.

2. **Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Федотов П.В., Царев Ф.Н., Шалыто А.А.** Метод сборки генома с использованием технологии MapReduce / I Всероссийский конгресс молодых ученых. III сессия «Технологии программирования и искусственный интеллект». НИУ ИТМО. 2012, с. 679-683 (<https://docplayer.ru/121335765-Parallelnyy-algoritm-de-novo-sborki-genoma-s-ispolzovaniem-tehnologii-mapreduce.html>).

3. **Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Царев Ф.Н., Шалыто А.А.** Метод сборки контигов геномных последовательностей на основе совместного применения графов де Брейна и графов перекрытий / Материалы Всероссийской научной конференции по проблемам информатики СПИСОК-2012. СПб.: ВВМ. СПбГУ. 2012, с. 415-418 (<http://is.ifmo.ru/works/2012/SPISOK/alexandrov-de-bruijn.pdf>).

4. **Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Царев Ф.Н., Шалыто А.А.** Метод сборки генома с использованием технологии MapReduce / Материалы Всероссийской научной конференции по проблемам информатики СПИСОК-2012. СПб.: ВВМ. СПбГУ. 2012, с. 419-422 (<http://is.ifmo.ru/works/2012/SPISOK/alexandrov-mapreduce.pdf>).

5. **Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Царев Ф.Н., Шалыто А.А.** Совместное применение графов де Брейна, графов перекрытий и микросборки для *de novo* сборки генома / Сборник тезисов III международной конференции «Постгеномные методы анализа в биологии, лабораторной и клинической медицине». Казань. 2012, с. 45, 46.

6. **Ахи А., Нигматуллин Н., Сергушичев А., Царев Ф.** Метод оценки расстояния между контигами на основе принципа максимального правдоподобия / Стендовый доклад на III международной научно-практической конференции «Постгеномные методы анализа в биологии, лабораторной и клинической медицине». Казань, 2012.

7. **Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Царев Ф.Н., Шалыто А.А.** Метод сборки контигов геномных последовательностей на основе совместного применения графов де Брейна и графа перекрытий / Сборник трудов Международного конгресса по интеллектуальным

системам и информационным технологиям IS&IT'2012. М.: Физматлит, 2012. Т. 3, с. 283-288 (<http://is.ifmo.ru/works/2012/genome-isit.pdf>).

8. Царев Ф.Н., Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Федотов П.В. Параллельный алгоритм *de novo* сборки генома с использованием технологии *MapReduce* / Международная суперкомпьютерная конференция «Научный сервис в сети Интернет 2012. Поиск новых решений»

(<http://agora.guru.ru/display.php?conf=abrau2012&page=item015&PHPSESSID=pueoj7ptig8373fbagh29dlc64>).

9. Alexandrov A., Kazakov S., Melnikov S., Sergushichev A., Shalyto A., Tsarev F. Combining de Bruijn graph, overlap graph and microassembly for *de novo* genome assembly / Proceedings of the 12th Annual Conference in Bioinformatics (Bioinformatics 2012). Stockholm, Sweden. 11.06.2012, p. 72.

<http://ctlab.itmo.ru/~svkazakov/papers/alexandrov-bioinf-2012-abstract.pdf>.

Первое наше выступление и публикация по этой тематике за рубежом.

Свидетельство о регистрации программы

Александров А.В., Казаков С.В., Мельников С.В., Сергушичев А.А., Федотов П.В., Царев Ф.Н.

Программное средство для сборки квазиконтингов из парных чтений // Свидетельство о государственной регистрации программы для ЭВМ. № 2012 616774. Дата регистрации – 27.07.2012 (http://is.ifmo.ru/certificates/2012_07_27_alexandrov_kazakov_melnikov_sergushichev_fedotov_tsarev/).

Знакомство с Павлом Певзнером

Теперь расскажу, как мы познакомились с Павлом Певзнером. Почти через два года после начала работ по сборке генома мы попросились выступить на семинаре его лаборатории. Доклад Феди Царева и Леша Сергушичева на тему «Метод сборки генома на основе совместного применения графов де Брейна и графов перекрытий» (слайды к докладу по запросу в Google «16 ноября 2012 года семинар по алгоритмической сборке генома») продолжался часа полтора. Он проходил в присутствии не только почти всех сотрудников лаборатории Певзнера, но и многих других представителей университета, включая Михаила Владимировича Дубину и Александра Владимировича Омельченко.

Нашим ребятам было задано множество вопросов, на каждый из которых они обстоятельно ответили. Последний вопрос задал Певзнер: «А где вы раньше были?» Я посчитал, что это признание, и ответил: «Готовились к этому семинару». В своем выступлении Павел Аркадьевич отметил актуальность решаемой нами задачи (сборка генома при ограничении на используемую память), высокий уровень исследований и качество доклада. После этого мы обсуждали возможность дальнейшего сотрудничества, которое, правда, не осуществилось. Однако наши достижения в этой области, полученные своими силами без помощи специалистов по биоинформатике, тем более зарубежных, приглашенных в Россию за огромные деньги, видимо, произвели на слушателей такое впечатление, что через четыре года, когда защищали диссертации Сергушичев и Казаков, отзывы о их диссертациях легко дали и Дубина, и представители Певзнера!

Знакомство с Максимом Артемовым

Девятого, десятого июля 2012 г. Сколтех провел совместно с МТИ II Международную конференцию «Создание Центров науки, инноваций и образования» (<https://www.golos-ameriki.ru/a/skolkovo-conference/1381361.html>). В состав делегации МТИ входил Максим Артемов (<https://scholar.google.com/citations?user=ReYYZ8oAAAAJ&hl=en>), закончивший там аспирантуру. После конференции делегация посетила Санкт-Петербург, где на одном из мероприятий «пересеклись» Максим и Федор Царев.

2013 год

Этот год стал ключевым в развитии кафедры КТ в этой области. В конце года Леше Сергушичеву удалось съездить на длительную стажировку в Университет Вашингтона в Сент-Луисе к молодому исследователю, а теперь Professor Максиму Артемову, который там работал и работает в настоящее время. Тогда же, Лешин одноклассник – Владимир Ульянов, а сейчас руководитель международной научной группы «Компьютерные технологии», начал обучение в аспирантуре также с длительной стажировки в группе Дмитрия Алексеева в Федеральном научно-клиническом центре физико-химической медицины (ФНКЦ ФХМ) в Москве. С тех пор Володя наряду с основной

тематикой (синтезом дискретных моделей) занимается и биоинформатикой: разработкой и реализацией алгоритмов для сравнительной метагеномики. В процессе коллаборации с этим центром были реализованы следующие программные средства:

1. **MetaFast** (*METAgenome FAST analysis toolkit*) is a toolkit for calculating a number of statistics of metagenome sequences and building the distance matrix between them (<https://github.com/ctlab/metafast>);

2. **MetaCherchant** is a tool for analysing genomic environment of a nucleotide sequence within a metagenome. The implementation is based on **MetaFast** source code (<https://github.com/ctlab/metacherchant>). Большой вклад в эту работу внес Артем Васильев.

В этом же году наш университет вошел в состав участников «Программы 5-100», которая позволила многим ребятам оставаться в университете и занимать наукой.

Наконец-то, по результатам соревнований *Assemblathon-2* вышла статья, в которой 91 соавтор: **Alexandrov A., Fedotov P., Kazakov S., Melnikov S., Sergushichev A., Tsarev F. et al.** *Assemblathon 2: Evaluating de novo Methods of Genome Assembly in Three Vertebrate Species* // **Giga Science**. 2013. V. 2. № 10. Open Access (<https://pubmed.ncbi.nlm.nih.gov/23870653/>).

В этом же году мы опубликовали статью: **Сергушичев А. А., Александров А. В., Казаков С. В., Царев Ф. Н., Шалыто А. А.** Совместное применение графа де Брейна, графа перекрытий и микросборки для *de novo* сборки генома // Известия Саратовского университета. Серия «Математика. Механика. Информатика». 2013. Вып. 2. Ч. 2, с. 51-57 (<https://ctlab.itmo.ru/~svkazakov/papers/sergushichev-isu-2013.pdf>).

В ней описан разработанный нами метод сборки контигов геномных последовательностей из парных чтений, который был **ориентирован на сокращение оперативной памяти** для *de novo* сборки генома, что достигалось совместным применением графов де Брейна и перекрытий, а также микросборки. Особенность метода состоит в том, что **процесс сборки контигов состоит уже не из двух, а из трех этапов: сборка квазиконтигов из чтений, сборка контигов из квазиконтигов и микросборка**. На первом этапе использовался граф де Брейна, на втором – граф перекрытий. Описаны результаты экспериментального исследования разработанного метода на чтениях геномов бактерии *E. Coli* (размер генома – 4.5 миллиона нуклеотидов) и рыбы *Maylandia zebra* (размер генома – миллиард нуклеотидов). Преимущество разработанного метода состоит в том, что для его работы требуется существенно меньше оперативной памяти по сравнению с существующими программными средствами для сборки генома.

Оценки сложности в биоинформатике – выдающиеся результаты

Статья для **RECOMB-seq 2013**, который проходил в рамках конференции **RECOMB 2013** (<http://bioinfo.au.tsinghua.edu.cn/recomb2013/recomeseqaccepted.html>)

Kapun E., Tsarev F. De Bruijn Super walk with Multiplicities Problem is NP-hard // BMC Bioinformatic. 2013, 14 (Suppl 5): S7. Proceeding of the Third Annual RECOMB Satellite Workshop on Massively Parallel Sequencing (RECOMB-seq 2013). Tsinghua University. Beijing, China (<http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-14-S5-S7>, http://is.ifmo.ru/articles_en/2013/kapun-de-bruijn-superwalk.pdf). Из рецензии на этот доклад: «This is an exciting result and an important contribution to our understanding of the theoretical limitations of genome assembly» / «**Это захватывающий результат и важный вклад в наше понимание теоретических ограничений сборки генома**».

13th Workshop on Algorithms in Bioinformatics (WABI 2013), который проходил во Франции **Kapun E., Tsarev F.** On NP-Hardness of the Paired de Bruijn Sound Cycle Problem // Lecture Notes in Bioinformatics. 2013. Vol. 8128, pp. 59-69. 13th Workshop on Algorithms in Bioinformatics. Sophia Antipolis, France. 2013 (<https://arxiv.org/pdf/1307.7806.pdf>).

Конференции

1. **Akhi A., Sergushichev A., Tsarev F.** Maximum Likelihood Scaffold Assembly / 17th Annual International Conference on Research in Computational Molecular Biology (RECOMB). 2013. V. 4, p. 156. Book of Abstracts. Tsinghua University. Beijing, China (<http://is.ifmo.ru/posters/2013/recomb-2013-poster.pdf>, https://www.researchgate.net/publication/266158164_Maximum_likelihood_scaffold_assembly).

2. **Александров А.В., Казаков С.В., Сергушичев А.А., Царев Ф.Н.** Метод сборки контигов геномных последовательностей с ошибками вставки и удаления на основе совместного применения графов де Брейна и графа перекрытий / Материалы научной конференции по проблемам информатики СПИСОК-2013. СПбГУ с. 578-582 (<http://is.ifmo.ru/works/2013/SPISOK/alexandrov-kazakov-sergushichev-tsarev.pdf>).
3. **Александров А.В., Казаков С.В., Сергушичев А.А.** Метод сборки контигов геномных последовательностей из парных чтений с ошибками вставки и удаления на основе совместного применения графов де Брейна и графа перекрытий / Сборник тезисов докладов II конгресса молодых ученых. Вып. 1. СПб.: Университет ИТМО. 2013, с. 145, 146 (https://research.itmo.ru/file/stat/118/sbornik_tezisov_dokladov_kongressa_molodyh_uchenyh_vypusk_1..pdf).
4. **Ахи А.А., Сергушичев А.А.** Метод оценки расстояния между контигами на основе принципа максимального правдоподобия / Сборник тезисов докладов II конгресса молодых ученых. Вып. 1. СПб.: Университет ИТМО. 2013, с. 146, 147 (https://research.itmo.ru/file/stat/118/sbornik_tezisov_dokladov_kongressa_molodyh_uchenyh_vypusk_1..pdf).
5. **Долганов В.О.** Восстановление фрагментов парных чтений при сборке транскриптома / Сборник тезисов докладов конгресса молодых ученых. Вып. 1. СПб.: Университет ИТМО. 2013, с. 154, 155 (https://research.itmo.ru/file/stat/118/sbornik_tezisov_dokladov_kongressa_molodyh_uchenyh_vypusk_1..pdf).
6. **Сергушичев А.А., Федотов П.В., Царев Ф.Н., Шалыто А.А.** Разработка методов сборки генома, сборки транскриптома и динамического анализа протеома / XLII научная и учебно-методическая конференция НИУ ИТМО. Секция «Автоматное программирование, эволюционные вычисления и биоинформатика». СПбГУ ИТМО. 2013.
7. **Сергушичев А.А., Казаков С.В., Федотов П.В., Царев Ф.Н., Шалыто А.А.** Разработка метода сборки геномных последовательностей на основе восстановления фрагментов по парным чтениям / XLII научная и учебно-методическая конференция НИУ ИТМО. Секция «Автоматное программирование, эволюционные вычисления и биоинформатика». СПбГУ ИТМО. 2013.
8. **Сергушичев А.А., Федотов П.В., Царев Ф.Н., Шалыто А.А.** Разработка алгоритмов сборки геномных последовательностей для вычислительных систем экзафлопсного уровня производительности / XLII научная и учебно-методическая конференция НИУ ИТМО. Секция «Автоматное программирование, эволюционные вычисления и биоинформатика». СПбГУ ИТМО. 2013.

Свидетельства о регистрации программ

1. **Александров А.В., Казаков С.В., Царев Ф.Н., Сергушичев А.А., Федотов П.В.** Программное средство, реализующее алгоритм поиска перекрытий между квазиконтигами // Свидетельство о регистрации программы для ЭВМ. № 2013 616471. Дата регистрации – 09.07.2013 (http://is.ifmo.ru/certificates/2013_07_09_alexandrov_et_al.pdf).
2. **Александров А.В., Казаков С.В., Царев Ф.Н., Мельников С.В., Сергушичев А.А., Федотов П.В.** Программное средство, реализующее запуск этапов сборки генома через графический интерфейс пользователя // Свидетельство о регистрации программы для ЭВМ. № 2013 619155. Дата регистрации – 26.09.2013 (http://is.ifmo.ru/certificates/2013_09_26_alexandrov_et_al.pdf).
3. **Александров А.В., Казаков С.В., Сергушичев А.А.** Программное средство, реализующее алгоритм упрощения графа перекрытий при сборке геномных последовательностей // Свидетельство о регистрации программы для ЭВМ. № 2013 660881. Дата регистрации – 21.11.2013 (http://is.ifmo.ru/certificates/2013_11_21_alexandrov_et_al_simplification.pdf).
4. **Александров А.В., Казаков С.В., Сергушичев А.А.** Программное средство, реализующее алгоритм исправления ошибок вставки и удаления в наборе чтений нуклеотидной последовательности // Свидетельство о регистрации программы для ЭВМ. № 2013 660882. Дата регистрации – 21.11.2013 (http://is.ifmo.ru/certificates/2013_11_21_alexandrov_et_al_indel.pdf).

Магистерские диссертации

1. **Александров А.В.** Разработка метода исправления ошибок вставки и удаления в наборе чтений нуклеотидной последовательности. Магистерская диссертация. Университет ИТМО. 2013 (<http://is.ifmo.ru/diploma-theses/2013/master/alexandrov/alexandrov.pdf>).
2. **Казаков С.В.** Разработка алгоритма упрощения графа перекрытий при сборке геномных последовательностей. Магистерская диссертация. Университет ИТМО. 2013 (<http://is.ifmo.ru/diploma-theses/2013/master/kazakov/kazakov.pdf>).

3. **Сергушичев А.А.** Разработка метода восстановления фрагментов генома по парным чтениям с ошибками вставки и удаления. Магистерская диссертация. Университет ИТМО. 2013 (<http://is.ifmo.ru/diploma-theses/2013/master/sergushichev/sergushichev.pdf>).

4. **Ахи А.А.** Разработка метода сборки скэффолдов геномных последовательностей на основе принципа максимального правдоподобия. Магистерская диссертация. Университет ИТМО. 2013 (<http://docplayer.ru/54733363-A-a-ahi-razrabotka-metoda-sborki-skeffoldov-genomnyh-posledovatelnostey-na-osnove-principa-maksimalnogo-pravdopodobiya-magisterskaya-rabota.html>).

У Александрова и Ахи научным руководителем был Федя Царев, а у Казакова и Сергушичева – я.

Победа в конкурсе «Лучший инновационный регион России»

В 2013 г. проект нашей лаборатории по сборке генома *Easy Genomics* принес победу Санкт-Петербургу как лучшему инновационному региону России. От Санкт-Петербурга приз получила ... **Инна Шалыто** (<http://is.ifmo.ru/photo/2013-11-01-Challenge-Cup/index.html>), которая в то время работала заместителем Председателя комитета по промышленной политике и инновациям Администрации Санкт-Петербурга (<http://news.ifmo.ru/ru/archive/archive2/news/3578/>).

Научно-популярные выступления и статьи Федора Царева

Пятого августа Федя на Летней школе «Биотехнологии будущего» сделал доклад на тему «Сборка генома *de novo*: мифы и реальность» (<http://docplayer.ru/48839450-Sborka-genoma-de-novo-mify-i-realnost.html>).

15 октября он выступил на *Третьем Science Slam* в Санкт-Петербурге на тему «Что биоинформатики не знают о геноме» (<https://www.youtube.com/watch?v=5ufX3EgJbuA>).

Перед этим выступлением – 26 сентября – на портале «Бумага» был опубликован текст: «Биоинформатик Федор Царев – о притягательности живой природы и сложности бытовых алгоритмов» (<https://paperpaper.ru/science-bioinformatics/>), а после выступления – 19 октября – на портале «Первая линия» был опубликован еще один текст: «Федор Царев: еще пять лет назад я не планировал заниматься биоинформатикой» (<https://1-line.spbu.ru/ljudi/234-1105.html>).

Участие Максима Буздалова в моделировании конформтно-зависимых свойств белков

В феврале 2013 г. были подведены итоги конкурса в рамках Федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2013 гг.». Университет ИТМО выиграл грант на исследования по теме «**Суперкомпьютерное моделирование конформационно-зависимых свойств белков в задачах рационального дизайна лекарственных препаратов**». Руководитель – **Александр Валерьевич Бухановский** (<https://4science.ru/project/14-514-11-4068>). Наша кафедра в лице Максима Буздалова была соисполнителем этой работы. У него по этой тематике две публикации: 1. *Спельников Д.М., Князев С. Н., Балахонцева М. А., Буздалов М.В., Порозов Ю.Б., Маслов В. Г., Бухановский А.В.* Высокопроизводительный программный комплекс моделирования конформационно-зависимых свойств белков в задачах рационального дизайна лекарственных препаратов // Динамика сложных систем. 2013. Т. 7. № 3, с. 12-16 (<https://pureportal.spbu.ru/en/publications/высокопроизводительный-программный-комплекс-моделирования-конформ>); 2. *Buzdalov M., Knyazev S., Porozov Yu.* Protein Conformation Motion Modeling Using Sep-CMA-ES / Proceedings of 13th International Conference on Machine Learning and Applications (ICMLA 2014). USA. 2014, pp. 35-40 (http://is.ifmo.ru/articles_en/2014/2014-icmla-cmaes-proteins.pdf).

Майн кайф

Можно подумать, что если молодой талантливый человек остался после уговоров учиться и работать в Университете, то и все хорошо. Это только в военных позывных бывает, что «над всей Испанией безоблачное небо». В реальной жизни за «небом» все время надо следить.

На примере Леши Сергушичева расскажу историю про мобильность, которая высоко ценится при определении рейтингов вуза. По этой причине мы не возражали против четырехмесячной стажировки Леши, который к тому времени уже четыре года занимался научной работой, в Университете Вашингтона в Сент-Луисе у Максима Артемова. За это время он принял участие в написании статьи для журнала с очень высоким импакт-фактором, публикация в котором также

поднимает рейтинг нашего университета. Но за все хорошее в жизни надо платить, и когда время стажировки приближалось к концу, Алексей сообщил мне, что ему «ненавязчиво» предлагают поступить в аспирантуру МТИ.

Молодой человек не собирался никуда уезжать, по крайней мере, в ближайшие несколько лет, но там нашлись «сладкоголосые сирены» и случилось то, что и должно было случиться. **Я ничего не смог возразить на это, так как об аспирантуре в МТИ мечтают многие, а некоторые, например, я, даже и не мечтали.** Осознав, что произошло ожидаемое (во-первых, кого им брать в аспирантуру, как не таких талантливых мальчиков, как Леша, а, во-вторых, он дома не живет уже семь лет, то почему бы ему не пожить еще где-то), я ничего не смог возразить.

В этой истории, как мне первоначально казалось, все шло, **как в сказке**: сердце молодого человека пронзила «стрела Амура», и у него не было сил **поступить иначе**, чем влюбиться в «ту», кто направила эту стрелу. Однако уже через месяц выяснилось, что это больше похоже не на сказку, а на **бытовой роман**. Тут, оказывается, была не внезапная страсть, а «любовь по расчету», которая подтолкнула молодого человека «жениться не обязательно на МТИ, а на Америке», чем сознательно он и занялся, подав документы в аспирантуру нескольких университетов, из которых ожидал ответов. Получается, как говорится, в одном неприличном анекдоте: **вроде бы одно и то же, но есть нюанс.**

Хочется спросить (только не знаю кого): «Потеря талантов – не слишком ли высокая цена за повышение рейтингов университетов?». И не надо в этот момент мне начинать читать мантру, что это не потеря талантов, а дальнейшее их развитие... Изложенное **очень напоминает жертвоприношение или отопление за счет сжигания ценных пород дерева.** Если так дело пойдет и дальше, то *рейтинг наших университетов будет целесообразно оценивать по минимуму суммы рейтингов*, в которых остались наши ребята. Итак, я свой вклад может быть уже внес: МТИ – в то время был первым в мире по рейтингу QS. А Вы можете столь же «эффективно» повысить рейтинг своего вуза? И еще. Несмотря на то, что некоторые, все-таки, возвращаются на Родину, это настолько редко, что является исключением, подтверждающим «правило».

Как Вы думаете, когда молодой человек, если его примут там в аспирантуру, вернется на Родину, особенно учитывая то, что он родом из провинции, а в аспирантуре даже МТИ, в отличие, например, от работы в *Microsoft* или *Google*, на квартиру в Санкт-Петербурге не заработать? Кроме того, ему всегда нравились *Hard Rock Cafe*... И еще. **Мы, конечно, не МТИ, но если все будут «сваливать», то никогда и не будем...** «Стимулируя обучение за границей, мы обрекаем на деградацию собственную систему образования. Опыт Петра I неуместен – тогда в России не было системы образования. Здесь показателен опыт Индии. Поскольку ее элита могла учиться в британских вузах, в Индии так и не появились хорошие университеты» (В. May).

Тем временем указанную выше статью, где Леша был соавтором, в журнал с очень-очень высоким импакт-фактором не приняли, а его, как я и ожидал, пригласили в аспирантуру, правда, не в МТИ или Гарвард, а в университет Вашингтона в Сент-Луисе, в котором он стажировался. Это тоже неплохой :-) университет: в нем учились и работали 22 (!) Нобелевских лауреата (http://en.wikipedia.org/wiki/List_of_Nobel_laureates_by_university_affiliation). В общем, большая польза для нас могла получиться...

Однако **американцы оказались жмотами, а мы сегодня не нищие**, и поэтому стипендия на все про все в 28 тыс. долларов в год (при курсе 30 руб. за \$), вряд ли это те деньги, которые привлекут молодого человека к активному занятию наукой в этом университете в течение четырех-шести лет. При этом следует учесть, что университет Вашингтона при всем уважении к нему не МТИ или Гарвард, а область исследований молодого человека – биоинформатика. Это не медицина, в которой такую зарплату можно терпеть долго, так как после окончания учебы в этой области в США обычно открываются прекрасные финансовые возможности.

В общем, мы сохранили еще одного талантливого молодого человека. При этом его совместные работы с университетом Вашингтона нами только приветствуются, и он может ездить туда в командировки и не на один месяц. **А тем временем статью с его участием и аффилиацией «ITMO University» опубликовал журнал *Immunity* с импакт-фактором 20!** В следующем году им в соавторстве была опубликована статья в журнале *Molecular Cell* с импакт-фактором 15, а в январе

2016 г. – в журнале *Cell Host & Microbe* с импакт-фактором более 12. Такое решение проблемы сохранения талантов в российском университете считаю наиболее целесообразным.

При этом я всегда помню, что все в мире меняется, и то, что молодому человеку было не интересно в 2014 г. в связи с изменившимися обстоятельствами может стать интересным, например, в 2015 г. (за это время курс рубля к доллару девальвировал в два раза). Это сильно мотивирует меня на продолжение борьбы за таланты. И еще. «**Концепцию эмиграции я считаю устаревшей.** У меня здесь «Гоголь-центр», в России живут родители, друзья. Я могу ставить за границей. Зачем же уезжать?» (К. Серебренников).

А мы жмотами не оказались – в 2013 г. в нашу аспирантуру поступило сразу пять (!) однокурсников: Леша Сергушичев, Володя Ульянов, Сережа Казаков, Антон Александров и Даня Чивилихин, у которых научным руководителем стал ... я, так как с каждым из них с третьего курса так или иначе занимался наукой.

2014 год

«Мой первый грант»

2014, 2015. Александров А.В. (руководитель), **Капун Е.Д.** **Разработка метода сборки генома из чтений, содержащих ошибки вставки и удаления.** Конкурс инициативных научных проектов Российского фонда фундаментальных исследований, выполняемых молодыми учеными «Мой первый грант».

Статья

Mardanov A., Toymentseva A., Gilyazeva A., **Kazakov S.**, Shagimardanov E., Khaitlina S., Sharipo M. Draft Genome Sequence of *Serratia Grimesii* Strain A2 // Genome Announcements. V. 2. 2014. № 5 (<https://mra.asm.org/content/ga/2/5/e00937-14.full.pdf>).

Конференции

1. **Kazakov S., Shalyto A.** Overlap Graph Simplification Using Edge Reliability Calculation / 8th International Conference Intelligent Systems and Agents. 2014. Lisbon, Portugal, pp. 220-226 (<https://pdfs.semanticscholar.org/3aaf/7558481899419a581b31ce8d39c2ed608652.pdf>).
2. **Ульянцев В.И., Казаков С.В., Дубинкина В.Б., Тяхт А.В., Алексеев Д.Г.** MetaFast – программное средство для высокопроизводительного сравнительного анализа метагеномов / Сборник трудов IV международной научно-практической конференции «Постгеномные методы анализа в биологии, лабораторной и клинической медицине». Казань, 2014, с. 103 (<http://is.ifmo.ru/works/2014/postgenome2014-ulyantsev.pdf>).
3. **Сергушичев А.А., Пирс Э., Артемов М.Н.** GAM: конвейер для совместного анализа данных транскрипционного и метаболического профайлинга / Сборник трудов IV международной научно-практической конференции «Постгеномные методы анализа в биологии, лабораторной и клинической медицины». Казань, 2014, с. 104 (<http://is.ifmo.ru/works/2014/postgenome2014-sergushichev.pdf>).
4. **Долганов В.О., Царев Ф.Н.** Алгоритм сборки транскриптома на основе анализа компонент связности графа де Брейна / Сборник трудов IV международной научно-практической конференции «Постгеномные методы анализа в биологии, лабораторной и клинической медицины». Казань, 2014, с. 123 (<http://is.ifmo.ru/works/2014/postgenome2014-dolganov.pdf>).
5. **Долганов В.О.** Разработка метода сборки транскриптома на основе анализа компонент связности графа де Брейна / Всероссийская научная конференция по проблемам информатики СПИСОК-2014. СПб.: ВВМ. СПбГУ. 2014 (<http://spisok.math.spbu.ru/2014/s15.asp>).
6. **Buzdalov M., Knyazev S., Porozov Yu.** Protein Conformation Motion Modeling Using Sep-CMA-ES / Proceedings of 13th International Conference on Machine Learning and Applications (ICMLA 2014). USA. 2014, pp. 35-40 (http://is.ifmo.ru/articles_en/2014/2014-icmla-cmaes-proteins.pdf).

Практическая школа по обработке данных секвенирования

18-20 февраля лабораторией «Компьютерные технологии» Университета ИТМО и компаниями «Life Technologies», ООО «Компания Хеликон» была проведена «Практическая школа по обработке данных секвенирования нового поколения». В школе приняло участие 14 человек. На этой школе Алексей Сергушичев и Федор Царев сделали доклад «Сборка генома *de novo*: мифы и реальность» (<http://genome.ifmo.ru/files/materials/bioinf2014/>).

Доклад Царева

Девятого апреля. Федя Царев сделал доклад «Сборка генома: мифы и реальность» на пленарном заседании III Всероссийского конгресса молодых ученых. СПб.: Университет ИТМО (<https://pdfslide.net/science/-iii--558ddc8acfa20.html>).

Магистерская диссертация

Долганов В.О. Разработка метода сборки транскриптома на основе анализа компонент связности графа де Брейна. Магистерская диссертация. Университет ИТМО. 2014 (<http://is.ifmo.ru/diploma-theses/2014/master/dolganov/dolganov.pdf>).

Сборка генома редкой бактерии

В этом году Сергей Казаков принял участие в проекте сборки генома редкой бактерии *Serratia grimesii*. Образцы клеток бактерии были получены в Институте цитологии РАН в Санкт-Петербурге, а ее секвенирование было выполнено в Институте фундаментальной медицины и биологии Казанского федерального университета. Геном бактерии включен в международную базу данных биотехнологической информации NCBI (<http://www.ncbi.nlm.nih.gov/nucleotide/JGVP000000000>).

Семинар по системной биологии

15-18 мая Федор Царев, Алексей Сергушичев и Павел Федотов совместно с представителями Института биоинформатики (<https://bioinf.me/>) провели **первый семинар** по системной биологии, в которой спикерами были **Максим Артемов** и **Александр Предеус** из Университета Вашингтона в Сент-Луисе, а также брат Максима – **Никита Артемов** из Гарвардского университета. Каждая тема включала практические занятия, в проведении которых участвовали Алексей Сергушичев и Павел Федотов (<http://bioinformaticsinstitute.ru/sbw2014>). В семинаре приняло участие 25 человек.

Роль Максима Артемова в наших успехах

Знакомство с Максимом Артемовым **стало для нас знаковым** – он вдохнул новую жизнь в наши исследования по биоинформатике и стал профессором-исследователем в Университете ИТМО. Важно отметить, что, когда рядом с нами появился Максим Артемов, ребята за счет предшествующих работ уже были готовы к работам по биоинформатике широком смысле слова.

И сам Максим становится все известнее: «На конференции *Keystone Dublin* Люк О'Нилл (один из «крутых парней» в области иммунитета и один из организаторов конференции) в своем докладе про макрофаги рассказывал про вещество «итаконат» и раз двадцать упоминал некоего Макса и его статьи. Макс Артемов в своем докладе про итаконат объяснил для тех, кто не в курсе, что тот Макс, про которого говорил Люк, – это он».

2015 год

Начало прорыва – статьи в классных журналах

Статьи

1. Jha A., Ching-Cheng Huang S., **Sergushichev A.**, Lampropoulou V., Ivanova Y., Loginicheva E., Chmielewski K., Stewart K., Ashall J., Everts B., Pearce E., Driggers E., Artyomov M. Parallel Metabolic and Transcriptional Data Reveals Metabolic Modules that Regulate Macrophage Polarization Distinct Metabolic Modules Promote Macrophage Polarization // **Immunity**. 2015. V. 42, № 3, pp. 419-430. **IF: 24.082, SJR: 16.215** (уже к июлю 2017 г. статья цитировалась 156 раз) (<https://www.sciencedirect.com/science/article/pii/S1074761315000801?via%3Dihub>).
2. Vincent E., **Sergushichev A.**, Griss T., Gingras M., Samborska B., Ntimbane T., Coelho P., Blagih J., Raissi T., Choinière L., Bridon G., Loginicheva E., Flynn B., Thomas E., Tavaré J., Avizonis D., Pause A., Elder D., Artyomov M., Jones R. Mitochondrial Phosphoenolpyruvate Carboxykinase Regulates Metabolic Adaptation and Enables Glucose-Independent Tumor Growth // **Molecular Cell**. 2015. Vol. 60, Issue 2, pp. 195-207. **5-Year IF: 15.052. IF: 14.018. SJR: 10.968** (<https://pubmed.ncbi.nlm.nih.gov/26474064/>).
3. Glotov A., **Kazakov S.**, Zhukova E., **Alexandrov A.**, Glotov O., Pakin V., Danilova M., Tarkovskaya I., Niyazova S., Chakova N., Komissarova S., Kurnikova E., Sarana A., Sherbak S., **Sergushichev A.**, **Shalyto A.**, Baranov V. Targeted Next-Generation Sequencing (NGS) of Nine Candidate Genes with Custom AmpliSeq in Patients and a Cardiomyopathy Risk Group // **Clinica Chimica Acta**. 2015. V. 446, pp.132-140. **IF: 2.7** (<https://pubmed.ncbi.nlm.nih.gov/25892673/>).

4. Toymentseva A, ..., **Kazakov S.**, ..., Sharipova M. Draft Genome Sequence of *Bacillus ginsengihumi* Strain M 2.11 with Phytase Activity // Genome Announcement. V.3. 2015 № 4. (<https://mra.asm.org/content/ga/3/4/e00851-15.full.pdf>).
5. **Suleimanova A.**, ..., **Kazakov S.**, ..., Sharipova M. High-quality draft genome sequence of a new phytase-producing microorganism *Pantoea* sp. 3.5.1 // **Standarts in Genomic Sciences**. Vol. 10. 2015. Article number: 95 (<https://environmentalmicrobiome.biomedcentral.com/articles/10.1186/s40793-015-0093-y>).

Конференции

1. **Kazakov S., Ulyantsev V., Dubinkina V., Tyakht A., Alexeev D.** MetaFast: fast reference-free graph-based comparison of shotgun metagenomic data / Proceedings of the International Moscow Conference on Computational Molecular Biology 2015 (MCCMB'15). M.: 2015. <http://ctlab.itmo.ru/~svkazakov/papers/2015-mccmb-kazakov.pdf>.
2. **Казаков С.В., Ульянов В.И., Дубинкина В.Б., Тяхт А.В., Алексеев Д.Г.** MetaFast: высокопроизводительный сравнительный анализ метагеномов на основе графа де Брейна / Сборник тезисов I международной школы-конференции «Биомедицина, материалы и технологии XXI века». Казань, 2015, с. 98. <http://ctlab.itmo.ru/~svkazakov/papers/2015-mt21-kazakov.pdf>.

Семинар по системной биологии

21-24 мая Институт биоинформатики и Университет ИТМО провели **второй семинар** по системной биологии с участием приглашенных лекторов из *Washington University in St. Louis* и *Harvard University*. Каждая тема включала практические занятия, в проведении которых участвовали Алексей Сергушичев, Павел Федотов и Антон Александров из международной лаборатории «Компьютерные технологии» Университета ИТМО (<http://bioinformaticsinstitute.ru/sbw2015>).

Выступление

Третьего июня на семинаре «Биоинформатика и системная биология» в Политехническом университете Санкт-Петербурга в качестве подготовки к защите кандидатской диссертации был сделан доклад **Казаков С., Шалыто А.** ITMO Genome Assembler: *de novo* сборка генома в условиях ограниченных системных ресурсов (<https://compbio.spbstu.ru/lab/seminary/page/4/>).

Публикация в известной газете

Публикация в журнале **Immunity** привлекла внимание известной российской газеты «Комсомольская правда» (<http://www.izh.kp.ru/daily/26449.4/3318658/>).

2016 год

Прорыв – публикации в классных журналах

Статьи

1. Campbell J., **Alexandrov A.**, Kim J., Wala J., Berger A., Pedamallu C., Shukla S., Guo G., Brooks A., Murray B., Imielinski M., Hu X., Ling S., Akbani R., Rosenberg M., Sougnez C., Ramachandran A., Collisson E., Kwiakowski D., Lawrence M., Weinstein J., Verhaak R., Wu C., Hammerman P., Cherniack A., Getz G., **Artyomov M.**, Schreiber R., Govindan R. Distinct Patterns of Somatic Genome Alterations in Lung Adenocarcinomas and Squamous Cell Carcinomas // **Nature Genetics**. 2016. V. 48. No 6, pp. 607-616. IF: 31.616. SJR: 23.762!!! (<https://pubmed.ncbi.nlm.nih.gov/27158780/>).
2. Lampropoulou V., **Sergushichev A.**, Bambouskova M., Nair S., Vincent E., Loginicheva E., Cervantes-Barragan L., Ma X., Huang S., Griss T., Weinheimer C., Khader S., Randolph G., Pearce E., Jones R., Diwan A., Diamond M., **Artyomov M.** Itaconate Links Inhibition of Succinate Dehydrogenase with Macrophage Metabolic Remodeling and Regulation of Inflammation // **Cell Metabolism**. [http://www.cell.com/cell-metabolism/fulltext/S1550-4131\(16\)30253-4](http://www.cell.com/cell-metabolism/fulltext/S1550-4131(16)30253-4) 2016. V. 24. No.1, pp. 158-166. IF: 17.565, SJR: 9.487 (<https://pubmed.ncbi.nlm.nih.gov/27374498/>).
3. Lu Q., Yokoyama C., Williams J., Baldrige M., Jin X., DesRochers D., Bricker T., Wilen C., Bagaitkar J., Loginicheva E., **Sergushichev A.**, Kreamalmeyer D., Keller B., Zhao Y., Kambal A., Green D., Martinez J., Dinauer M., Holtzman M., Crouch E., Beatty W., Boon A., Zhang H., Randolph G., **Artyomov M.**, Virgin H. Homeostatic Control of Innate Lung Inflammation by Vici Syndrome Gene *Epg5* and Additional Autophagy Genes Promotes Influenza Pathogenesis // **Cell Host&Microbe**. 2016. V.19, pp. 102-113. IF: 12.328. SJR: 6.8 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4714358/>).

4. **Sergushichev A., Loboda A., Jha A., Vincent E., Driggers E., Jones R., Pearce E. Artyomov M.** GAM: a Web-Service for Integrated Transcriptional and Metabolic Network Analysis // **Nucleic Acids Research**. 2016. V. 44. No. 1, pp.194-200. **IF: 9.112, SJR: 6.16** (<https://pubmed.ncbi.nlm.nih.gov/27098040/>).
5. Derr A., ..., **Sergushichev A.**, ..., Artyomov M., Garber M. End Sequence Analysis Toolkit (ESAT) expands the extractable information from single-cell RNA-seq data // **Genome Research**. 2016. 26: 1397-1410. (<https://genome.cshlp.org/content/26/10/1397>).
6. **Ulyantsev V., Kazakov S., Dubinkina V., Tyakht A., Alexeev D.** MetaFast: Fast Reference-free Graph-Based Comparison of Shotgun Metagenomic Data // **Bioinformatics**. 2016. V.32. No. 18, pp. 2760-2767. **IF: 8.136, IF: 4.981** (<https://academic.oup.com/bioinformatics/article/32/18/2760/1743520>).
7. Izreig S., Samborska B., Johnson R.M., **Sergushichev A.**, Ma E.H., Lussier C., Loginicheva E., Donayo A., Poffenberger M., Sagan S., Vincent E., **Artyomov M.**, Duchaine T., Jones R. The miR-17 Similar to 92 MicroRNA Cluster is a Global Regulator of Tumor Metabolism // **Cell Reports**. 2016. Vol. 16. No. 7, pp. 1915-1928. **IF: 7.87** (<https://www.sciencedirect.com/science/article/pii/S2211124716309548?via%3Dihub>).
8. **Artemov M., Sergushichev A., Schilling J.** Integrating Immunometabolism and Macrophage Diversity // **Seminars in Immunology**. 2016. V. 28. № 5, pp. 417-424 (<http://www.sciencedirect.com/science/journal/10445323/28/5>).
9. Dubinkina V., Ischenko D., **Ulyantsev V.**, Tyakht A., Alexeev D. Assessment of k-mer Spectrum Applicability for Metagenomic Dissimilarity Analysis // **BMC Bioinformatics**. 2016, 17:38. **Open Access. IF: 2.56, SJR: 1.72** (<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-015-0875-7>).
10. Vashukova E., Glotov A., **Fedotov P.**, Efimova O., Pakin V., Mozgovaya E., Pendina A., Tikhonov A., Koltsova A., Baranov V. Placental MicroRNA Expression in Pregnancies Complicated by Superimposed Pre-Eclampsia on Chronic Hypertension // **Molecular Medicine Reports**. 2016. V. 14. No. 1, pp. 22-32. **IF: 1.554** (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4918533/>).
11. **Сергушичев А.А.** Алгоритм кумулятивного вычисления статистики представления набора генов // Научно-технический вестник информационных технологий, механики и оптики. 2016. № 5, с. 956-959 (<https://ntv.ifmo.ru/file/article/15913.pdf>).
12. **Александров А.В., Шалыто А.А.** Метод исправления ошибок вставки и удаления в наборе чтений нуклеотидной последовательности // Научно-технический вестник информационных технологий, механики и оптики. 2016. № 1, с. 108-114 (<https://ntv.ifmo.ru/file/article/14554.pdf>).
13. **Комиссарова С., Чакова Н., Ниязова С., Казаков С., Жукова Е., Александров А., Глотов О., Глотов А.** Особенности клинических проявлений гипертрофической кардиомиопатии у пациентов с различными мутациями в генах саркомеров // Российский кардиологический журнал. 2016. № 1, с. 20-25 (<https://doi.org/10.15829/1560-4071-2016-1-20-25>).
14. **Казаков С.В., Шалыто А.А.** Анализ геномных и метагеномных данных в образовательных целях // Компьютерные инструменты в образовании. 2016. № 3, с. 5-15 (<http://cte.eltech.ru/ojs/index.php/kio/issue/view/137>).

Конференции

1. **Loboda A., Artyomov M., Sergushichev A.** Solving Generalized Maximum-Weight Connected Subgraph Problem for Network Enrichment Analysis / Algorithms in Bioinformatics. 16th International Workshop (WABI 2016). Lecture Notes in Computer Science. V. 9836, pp. 210-221 (https://link.springer.com/chapter/10.1007%2F978-3-319-43681-4_17).
2. **Казаков С.В., Шалыто А.А.** Сборка генома *de novo* из данных высокопроизводительного секвенирования на персональном компьютере / Материалы международной международной научной конференции «Компьютерные науки и информационные технологии». Саратов: СГУ. 2016, с. 178-181 (<https://drive.google.com/file/d/0Bw-3xW7BayMkb1FCcDVmdVVPY1k/view>).
3. **Сергушичев А.А.** Алгоритм для быстрого анализа представления генов / Материалы Всероссийской научной конференции по проблемам информатики СПИСОК-2016. СПб.: ВВМ. СПбГУ. 2016, с. 238-244 (<http://spisok.math.spbu.ru/2016/s15.asp>).
4. **Казаков С.В., Шалыто А.А.** Сборка генома *de novo* на персональном компьютере / Материалы Всероссийской научной конференции по проблемам информатики СПИСОК-2016. 2016, с. 245-250 (<http://spisok.math.spbu.ru/2016/s15.asp>).

Бакалаврская работа

Лобода А.А. Решение задачи нахождения связного подграфа максимального веса с учетом весов ребер. Бакалаврская работа. Университет ИТМО. 2016 (<http://is.ifmo.ru/diplomatheses/2016/bachelor/loboda/loboda.pdf>).

22 декабря защитили кандидатские диссертации Алексей Сергушичев и Сергей Казаков

Тема диссертации Леша «Методы вычислительного анализа метаболических моделей для интерпретации транскриптомных и метаболомных данных». Его научный руководитель – Максим Артемов. Несмотря на то, что Лешиным научным руководителем долгое время был я, в конце концов, истина восторжествовала, и Максим стал его руководителем. Автореферат, диссертация и презентация Алексей приведены по адресу: <http://is.ifmo.ru/disser/>. Там же размещен **отзыв выдающегося иммунолога, академика РАН С.А. Недоспасова**, а здесь находятся все документы, требующуюся для защиты: http://fppo.ifmo.ru/?page1=16&page2=52&page_d=1&page_d2=146244.

Тема диссертации Сергея «Автоматизация сборки генома и сравнительного анализа метагеномов для обучения геномной биоинформатике». Его научным руководителем был я. Автореферат, диссертация и презентация – все приведено по адресу: <http://is.ifmo.ru/disser/>. Там же размещен **отзыв академика РАН М.В. Дубины**. Все документы, требующуюся для защиты, размещены здесь: http://fppo.ifmo.ru/?page1=16&page2=52&page_d=1&page_d2=145757.

Сережа довел сборщик генома до возможности использования на персоналках без участия авторов. Этому сборщику в мире аналогов нет. Вернее один есть, но платный. После того, как нам удалось разработать сборщик, использующий мало оперативной памяти, мы пошли дальше. Если все в мире собирают геном на кластерах, в облаках и под *Linux*, то мы **решили делать сборщик для врачей: для персоналок и под Windows**. И сделали его и не только для этой операционки.

Сборщик целесообразно применять для сборки генов бактерий и вирусов (это оказалось очень актуально в период пандемии в 2020 г.) в условиях, когда ученые не хотят идти на *Amazon*, а могут открыть персоналку и начать работать.

Получение нашими ребятами отзывов от академиков можно считать официальным признанием нашего вклада в биоинформатику, в которую я «впутался» в 2009 г. в результате знакомства с академиком К.Г. Скрябиным. Некоторое время эта деятельность ничего, кроме скепсиса, у окружающих не вызывала, **но мы держались и удержались!** Поздравляю с этим ребят и себя тоже!

Перед защитой ребят пришли письма, которые привожу в переводе. **Леша получил письмо из онкологического национального исследовательского центра Испании:** «Мы делаем анализ с использованием Вашего программного пакета *fgsea*, и я хотел бы сказать за него большое спасибо. Это очень хороший пакет! Он сэкономил нам кучу работы и времени. Напишите, пожалуйста, как только выйдет статья в журнале, чтобы мы могли ее процитировать. *H. Tejero, PhD, Translational Bioinformatics Unit*».

Сергей получил письмо от Б. Олссона, PhD, руководителя исследовательской группы по биоинформатике из школы бионаук Университета Шведе (Швеция): «Мы успешно использовали *ITMO Genome Assembler* для сборки черновых геномов недавно отсеквенированных бактерий, а сейчас пишем статью в *Standards in Genomics*. Ваш сборщик хорошо подошел нам, так как мы должны были работать на платформе *Windows*. **Мы были рады найти *de novo* сборщик, который работает этой платформе, очень прост в установке и использовании, а также не требует много памяти и очень вычислительно эффективен.** Это позволяет устанавливать его на персональный компьютер. Мы считаем, что он заполняет важную нишу для *de novo* сборки генов».

На защите С. Казакова, который защищался по специальности 05.13.06. «Автоматизация технологических процессов и производств» (образование), кое-кто из членов совета сомневался в том, что его диссертация относится к образованию. Когда я вышел выступать как научный руководитель, то сразу согласился, что эта диссертация, действительно, не про образование, так как по «Закону об образовании», «образование – это воспитание плюс обучение», в то время как у Сергея нет ни слова про воспитание, а только про обучение, так как воспитанием обычно занимаюсь я. Однако так как в паспорте специальности нет ни слова про воспитание, то я сказал, что считаю

сделанный нами выбор специальности правильным. А еще я сказал, что диссертация полностью соответствует «духу» специальности. Сказанное «народу» понравилось.

Мой давний знакомый профессор **Владимир Ляндрес** из Израиля так прокомментировал наши «биологические» защиты: «**Толя! Здорово!**». Я ответил: «Спасибо на добром слове. **Я же должен доказывать себе (да и остальным тоже), что я не зря остался жить и работать в России, и что очень многого можно добиться и здесь!**» «Доказал убедительно», – написал профессор **Игорь Бессмертный**. Пришло также письмо от профессора **Валерия Вяткина**: «Анатолий Абрамович! Примите мои искренние поздравления! Хороший урок скептикам!».

Что нам дали исследования по сборке генома

Сегодня можно утверждать, что разработка сборщика генома позволила нашим ребятам войти в мир **биоинформатики**, что обеспечило возможность после знакомства с Максимом Артемовым из Донецка, который окончил химический факультет МГУ и аспирантуру в МТИ, а сейчас работает в университете Вашингтона в Сент-Луисе, сразу начать совместную работу.

Сначала туда на длительное время поехал Леша Сергушичев, а потом – Антон Александров. В результате стало понятно, что **программисты, получившие образование на нашей кафедре, значительно проще входят в мир биологии, чем биологи в мир программирования**. Максим быстро понял уровень наших ребят, и они с ним пошли дальше – **начали заниматься системной биологией** на каком-то совершенно невероятном для нас уровне (стали появляться публикации, которые мне даже не снились: в журналах *Science*, *Nature*, *Cell*, *Immunity*, *Molecular Cell* и т. д.). Конечно, в этих работах были не только наши ребята – в каждой из них по 10-15 авторов из МТИ, Гарварда, Сент-Луиса, институтов и клиник Бостона, но аффилиация «*ITMO University*» в публикациях в этих выдающихся изданиях стала появляться.

Это решило проблему финансирования нашей международной научной лаборатории «Компьютерные технологии» в рамках «Программы 5-100», так как для его получения, в частности, требовалось 40-50 публикаций в год, от которых сначала была необходима только их индексация в базах данных *Scopus* и *Web of Science*, а потом потребовали, чтобы эти работы были опубликованы в хорошо цитируемых изданиях.

Сначала в качестве показателя цитируемости у нас использовался импакт-фактор (*IF*), но потом и моими стараниями, в частности, университет перешел на *SJR* – показатель цитируемости в рассматриваемой области. Было время, когда мы брали на себя обязательство иметь 40 публикаций в год, **зарегистрированных** в указанных базах **в этом же году**, а руководство университета требовало обеспечить по ним еще и средний *SJR*, равный 0,9 – по всем публикациям надо было в сумме набрать 36.

Однако набрать эту сумму только за счет информатики (*Computer Science*) **невозможно**. Например, чемпион мира по программированию Максим Буздалов вместе с французским профессором Бенджамином Доером получили *Best Paper* на теоретической секции конференции *Genetic and Evolutionary Computation Conference (GECCO 2017)*. Но *SJR* материалов этой конференции, которые собирают по 600-700 человек со всего мира, равен *нулю*! У каждого из выпусков *Lecture Notes in Computer Science* у Шпрингера – 0,3. Если 40 умножить на 0,3, то получается 12, а совсем не 36. Если бы у нас не было публикаций даже не по биоинформатике, а по системной биологии (в этой области журналы имеют значительно более высокий *SJR* по сравнению с биоинформатикой, не говоря уже об информатике), то наша лаборатория бы уже давно вылетела из «Программы 5-100», и многие ребята от нас бы ушли.

Я, видимо, немножко сумасшедший. А может быть, и не немножко, если довел такого хорошего мальчика, как Леша Сергушечев, не до программирования, например, в *Google*, а до такого безумия, как *Fellow* Сколтех по биоинформатике и системной биологии (об этом ниже). Причем это звание ежегодно присваивается только пяти ученым до 35 лет со всей страны. В комиссию в 2017 г. входили шесть человек, из которых двое – совсем не простые ребята, которые приехали из Америки – **Михаил Гельфанд** (внук великого математика **Израиля Моисеевича Гельфанда**) и **Константин Северинов**, который командует тремя лабораториями – двумя в России и одной в Америке.

Леша тоже хотел уехать, но, в отличие от других, только в одно место на земле – в Бостон, где исследования по иммунитету рака выполняются на высшем уровне, а вовсе не в *Google* или куда возьмут. При этом отмечу, что он закончил кафедру «Компьютерные технологии», а не биоинформатику или системную биологию. Мы с ним уже не единожды в огромной конкурентной борьбе выиграли большие гранты по этой тематике. В результате он никуда не уехал, остался на постоянной работе на нашей кафедре. **Зачем ему куда-то ехать, когда ему здесь при нормальной зарплате никто не только не мешает делать то, что он хочет, а все ему только помогают!** Первое время он мог (пока он не стал у нас лидером этого направления, и у него не появилась «куча» забот) сколько угодно «сидеть» в Америке, и от него только и требовалось, что в статьях ставить аффилиацию *ITMO University* и все. Деятельность Леша и его ребят способствовала тому, что недавно наш университет попал в список восходящих звезд журнала *Nature*, а еще во многом благодаря статьям Кати Эсауловой в *Science* и Лешиной в *Nature* наш университет на несколько лет попал в Шанхайский рейтинг. Почему на несколько лет? А потому, что там и не только там понимают, что для того, чтобы опубликоваться в журналах такого уровня надо работать три-четыре года.

Семинар по системной биологии

16-21 мая Институт биоинформатики и Университет ИТМО провели **третий** семинар по системной биологии с участием приглашенных лекторов из *Washington University in St. Louis* и *Harvard University*. **Специальным гостем семинара был выдающийся ученый Марк Дейли**, о взаимодействии с которым написано ниже (<http://bioinformaticsinstitute.ru/sbw2016>).

О Леше Сергушичеве

У меня уехал выпускник (на этот раз в *Facebook*), который является соавтором статьи, которая рецензируется в *Nature*! Многие ребята, видимо, с ума посходили: **даже в такой ситуации** они либо не верят в себя, либо не хотят заниматься наукой, либо хотят много денег.

Однако, мы, все-таки, сопротивляемся этой тенденции и очень неплохо. Так, например, Алексей Сергушичев в своем недельном отчете как-то написал: «Приняли еще две статьи, где я четвертый в списке авторов в *Cell Reports* и в *Genome Research*». «Молодец, классно!», – ответил я. Благодаря усилиям Леша и окружающих его ребят Университет ИТМО уже в 2016 г. обеспечил заметный прирост в публикационной активности в журналах издательской группы *Nature* (<https://rb.ru/news/nature-index/>)! Об этом более подробно будет рассказано ниже.

Интервью с Лешей в газете «Санкт-Петербургские ведомости» размещено здесь: <https://spbvedomosti.ru/news/obshchestvo/programmist/>. А тем временем портал *The Village* первого сентября того же года реализовал проект (<http://www.the-village.ru/village/people/specials/243523-coders>), в рамках которого наши студенты и выпускники, должны были не только рассказать о себе, но и сказать несколько слов о своем кумире в области информационных технологий и сфотографироваться в его образе. Вот, например, что сказал Алексей: «Хочу лет через десять стать профессором в каком-нибудь университете. **А еще я мечтаю сделать что-нибудь крутое, чтобы в будущем у кого-нибудь был повод сняться в моем образе**». Интересно, и я знаю этот университет...

Леша определился с тем, чем заниматься в жизни – о нем вышла статья с красноречивым названием: «**Прирожденный исследователь**». *Newton*. 2016. Июль, № 24 (<http://is.ifmo.ru/aboutus/2016/born-researcher.pdf>). Однако после защиты кандидатской – в конце декабря 2016 г., он еще не знал, где в ближайшее время будет заниматься наукой.

Алексей имел высокорейтинговые статьи на стыке компьютерных наук, биоинформатики и системной биологии, написанные в соавторстве с сильными иностранными учеными (<http://news.ifmo.ru/science/it/news/5697/>). Как-то он сказал, что **его мечта – публикация статьи в *Nature*, в которой он будет первым автором**.

Когда я написал об этом в «Заметках...», то по этому вопросу возникла интересная переписка с Игорем Мазничей: «А это круто, когда мечта – это всего лишь публикация в каком-то журнале?» Я ответил: «Да, это очень круто, когда в таком журнале и первым автором. В настоящей науке свои законы». Он написал: «Тогда понятно, почему большое число открытий сделано людьми, далекими

от науки :-)). Вот мой ответ: **«Ни одно открытие не сделано не учеными, людьми, далекими от науки, сделаны только изобретения!»**

Я спросил Лешу, согласен ли он с моим последним утверждением? Он ответил, что в целом у него нет возражений, но, видимо, что-то можно открыть и случайно. Я согласился с ним и сказал: «Мне кажется, что **случайно можно открыть, например, книгу на той странице, которая нужна**, да и то, скорее всего, если ее на этой странице уже открывали». Лешин смех был вполне одобрительным.

И еще про людей далеких от науки: **надо стараться, чтобы «люди, далекие от науки, были далеки от нее».**

Как мы помогли Университету ИТМО стать «восходящей звездой» журнала *Nature*
В портале *Nature Index* (<http://www.natureindex.com/news-blog/who-are-the-research-worlds-rising-stars>) отметили, что сотрудники Университета ИТМО стали весьма активно публиковаться в широко известных в мире 68 журналах, связанных с журналом *Nature* (*Nature Partner Journal*) (<http://www.natureindex.com/faq#journals>), и назвали наш университет «восходящей звездой» этого портала (*Nature Index of Rising Stars*) (<http://news.ifmo.ru/archive/archive2/news/5867/>).

Там была отмечена и международная научная лаборатория (МНЛ) «Компьютерные технологии» (<https://www.natureindex.com/.../inte.../5746654a140ba05a228b456d>), так как наши ребята в соавторстве опубликовали статьи в следующих журналах из указанного списка: *Genome Research*, *Cell Host & Microbe*, *Immunity*, *Cell Metabolism*, *Molecular Cell*, *Nature Genetics* (<https://rb.ru/news/nature-index/>). Я поблагодарил за это Максима Артемова, и получил ответ: «Класс! То ли еще будет! Мы на этом останавливаться не намерены :-). Спасибо Вам большое за ребят и поддержку!». В 2017 г. у наших появились статьи еще и в *Science*, *Cell* и *Nature Microbiology*.

Интересно, что Франция – это страна, где одной статьи за четыре года в *Nature*, *Cell* или *Science* достаточно, чтобы классифицировать ученого как активно публикующегося.

При этом отмечу, что сотрудникам, успешно работающим в МНЛ, в нашем университете платят весьма прилично, но о том, сколько платят китайцам за публикации в журналах типа *Science* и *Nature*, вслух лучше не произносить (https://indicator.ru/news/2017/07/13/skolko-platyat-kitajskim-uchenym-za-publikacii?utm_referrer=https%3A%2F%2Fzen.yandex.com&utm_source=twsharing&utm_medium=social), как впрочем, и о средней зарплате профессора в ведущих китайских университетах, число которых не один или два, а порядка ста.

Появление указанной выше статьи у Антона Александрова с огромными показателями IF: 31.616. SJR: 23.762 рассматривалось мною как начало его публикаций в журналах такого уровня. Однако для Антона это было не началом, а концом научной карьеры. В октябре 2016 г., вместо представления диссертации к защите, он сообщил мне в присутствии Леши Сергушичева, что **разочаровался в науке** и поэтому уезжает на работу в *Google*. Я спросил его, не путает ли он слова, и не разочаровался ли он на самом деле не в науке, а в себе. Внятного ответа не последовало, да если он и был, это к делу уже не относилось (http://spbvedomosti.ru/news/obshchestvo/kuda_nbsp_zh_ne_nbsp_rasplativshis/).

При этом отмечу, что стать учеными хотят лишь немногие. **«Чаще всего это не страх того, что не вынести тяжелый труд или лишиться нормального человеческого будущего. Гораздо непреодолимее кажется неуверенность в собственных силах, своем предназначении»** (Николай Егорович Жуковский).

Об уровне наших студентов

Константин Зайцев в 2016 г. учился на шестом курсе. До этого он несколько лет проработал у нас в лаборатории и занимался биоинформатикой (<https://stepik.org/course/Python-основы-и-применение-512/>), что позволило ему на полгода (!) поехать на стажировку в Америку к Максиму Артемову. Там он решил поступать в аспирантуру, но не по биоинформатике или вычислительной биологии, а по ... **иммунологии**. И это, не имея базового образования по этому предмету.

При поступлении, естественно, надо было сдавать экзамен по специальности, который состоял из пяти часовых бесед один на один с известными в мире иммунологами, включая пионера в области онкологической иммунологии **Роберта Шрайбера** (https://ru.wikipedia.org/wiki/Шрайбер,_Роберт). По словам Кости, они остались довольны его ответами! На мой вопрос откуда он знает этот предмет,

молодой человек ответил, что за время работы у нас в лаборатории выучил его! Несмотря на это, еще не известно поступит ли Костя в эту аспирантуру, так как в Университете Вашингтона в Сент-Луисе результаты экзамена по специальности только один из показателей для оценки кандидата на поступление. **Среди других показателей есть и академическая успеваемость в бакалавриате**, с которой у Кости на младших курсах не все было здорово.

Это пример того, что в университете не все могут позволить себе «забить» на учебу. Указанный университет, естественно, в этом вопросе не единственный: я знаю компанию, в которой при среднем балле диплома менее 4,8 кандидат даже не рассматривается: считается, что он либо неспособный, либо разгильдяй. Однако не очень важно, поступит ли Костя сейчас в эту аспирантуру или нет, главное то, что он крутой и определился с тем, чем хочет заниматься в жизни!

Часть 3. 2017, 2018 гг.

<https://vk.com/@1077823-chast-3-2017-2018-gg-kak-bioinformatika-i-sistemnaya-biologi>

2017 год

В этом году Алексей Сергушичев возглавил **группу «Биоинформатика»** в международной научной лаборатории «Компьютерные технологии», хотя как отмечено выше, Ученый Совет университета еще в 2011 г. утвердил лабораторию «**Алгоритмы сборки геномных последовательностей**» при нашей кафедре.

Государственное задание

2017-2019 гг. Государственное задание Министерства образования и науки РФ. Тема: «**Разработка алгоритмов анализа метаболических и сигнальных сетей для идентификации модулей, регулирующих клеточную адаптацию**». Руководитель – А.А. Шалыто.

Первого января 2017 г. я получил сообщение от Леши: «Анатолий Абрамович, поздравляю с поддержкой (<http://минобрнауки.рф/документы/9389>) Вашего проекта по государственному заданию подведомственным Министерству образования и науки РФ образовательным организациям «**Разработка алгоритмов анализа метаболических и сигнальных сетей для идентификации модулей, регулирующих клеточную адаптацию**».

Из более 2500 заявок Комиссия Минобрнауки России при активном участии Совета по науке при Минобрнауки России и Проектного офиса научно-технологической инициативы отобрала 444 проекта 125 организаций.

Вот мой ответ Алексею: «Проект заявлялся под моим руководством, но его ответственный исполнитель – ты. Поэтому, кто из нас выиграл, еще неизвестно: по моему мнению, оба». Не менее приятно было получить короткое письмо от Максима Артемова: «Урррааа!». Интересно, что, когда я встретил Киру Вяткину, которая писала отзыв ведущего предприятия на диссертацию Сергея Казакова от лаборатории Певзнера, то поделился с ней этой приятной для нас новостью, а кроме этого добавил, что с момента начала работ по сборке генома в 2010 г. и до сих пор, я не прочел ни одной работы по биологии, что мне не мешает вести ребят и в этой области вперед. Кира на несколько секунд задумалась и сказала: «А Вам, Анатолий Абрамович, похоже, это и не надо – у Вас, и так, все хорошо получается!». Что тут скажешь?

Статьи в выдающихся журналах

1. Steed A., Christophi G., Kaiko G., Sun L., Goodwin V., Jain U., **Esaulova E.**, Artyomov M., Morales D., Holtzman M., Boon A., Lenschow D., Stappenbeck T. The Microbial Metabolite Desaminotyrosine Protects from Influenza Through Type I Interferon // **Science**. 2017. Vol. 357, Issue 6350, pp. 498-502. **IF: 37.0. SJR: 13.5** (<https://science.sciencemag.org/content/357/6350/498>).
2. Ulland T., Song W., Huang S., Ulrich J., **Sergushichev A.**, Beatty W., **Loboda A.**, Cairns N., Kambal A., Loginicheva V., Gilfillan S., Cella M., Virgin H., Unanue E., Wang Y., Artyomov M., Holtzman D., Colonna M. TREM2 is a Global Regulator of Microglia Energetic and Biosynthetic Metabolism During Steady State and in Alzheimer's Disease // **Cell**. 2017. V. 170. Issue 4, pp.649-663.e13. **IF: 30.41. SJR: 26.95!!!** ([http://www.cell.com/cell/pdf/S0092-8674\(17\)30830-9.pdf](http://www.cell.com/cell/pdf/S0092-8674(17)30830-9.pdf)).
3. Chatterjee S., Luthra P., **Esaulova E.**, Agapov E., Yen B., Borek D., Edwards M., Mittal A., Jordan D., Ramanan P., Moore M., Pappu R., Holtzman M., Artyomov M., Basler C., Amarasinghe G., Leung D.

Structural Basis for Human Respiratory Syncytial Virus NS1-mediated Modulation of Host Responses // **Nature Microbiology**. 2017. Vol 2. Article number: 17101 (<https://www.nature.com/articles/nmicrobiol2017101>).
4. Bi W., Greenwald N., Abedalthagafi M., Wala J. Gibson W., Agarwalla P., Horowitz P., Schumacher S., Esaulova E., Mei Y., Chevalier A., Ducar M., Thorner A., van Hummelen P., Stemmer-Rachamimov A., Artyomov M., Al-Mefty O., Dunn G., Santagata S., Dunn I., Beroukhi R. Genomic Landscape of High-grade Meningiomas // **Genomic Medicine**. 2017. 2. Article number: 15. Open Access. *Nature Partner Journal* (<https://www.nature.com/articles/s41525-017-0014-7>).

Конференция

Isomurodov J., Loboda A., Sergushichev A. Ranking Vertices for Active Module Recovery Problem / Proceedings of 4th International Conference on Algorithms for Computational Biology (AlCoB 2017). Springer. 2017, pp. 75-84. Portugal (<http://www.springer.com/gp/book/9783319581620>).

Бакалаврская работа

Зенкова Д.М. Реализация эффективного взаимодействия между платформой для анализа экспрессии генов *Morpheus* и библиотекой вычислительных методов *R/Bioconductor*. Бакалаврская работа. Университет ИТМО. 2017 (http://is.ifmo.ru/diploma-theses/2017/bachelor/zenkova/thesis_Dariya_Zenkova.pdf).

Skoltech Fellowship Program 2017

Шестого апреля Леша Сергушичев вошел в пятерку лучших молодых ученых страны по **системной биологии** и попал в *Skoltech Fellowship Program 2017* (<https://www.skoltech.ru/en/2017/04/the-winners-of-the-systems-biology-fellowship-program-2/>). Так были оценены его исследования по **созданию вычислительных методов и программного обеспечения, которые позволяют понять динамику регуляции биохимических реакций в процессе иммунного ответа** (<http://news.ifmo.ru/ru/science/it/news/6584/>). Продолжительность проекта – три года. При этом главное, что **победителям никуда не надо переезжать**, а следует только раз в год отчитываться.

Нина Яныкина на это сообщение отреагировала так: «Какая круть!». Я ответил: «Еще какая!», особенно учитывая то, что те, кто принимал решение по этому вопросу, были совсем «непростыми ребятами», такие как, например, упомянутые выше Михаил Гельфанд и Константин Северинов.

Игорь Бужинский, который, как и многие окружающие меня люди, нелегко «расстается с добрыми словами», написал: «Зашибись успех!»

После этого я стал называть Лешу «великим русским ученым», учитывая также и его «послужной список»: (<https://scholar.google.ru/citations?user=fcH0gPgAAAAJ&hl=en>).

Семинар по системной биологии

14-19 мая Институт биоинформатики и Университет ИТМО провели **четвертый** семинар по системной биологии в России с участием приглашенных лекторов из *Washington University in St. Louis* (Максим Артемов), *Harvard University* (Никита Артемов), Институт биоинформатики (Александр Предеус), Университетеб ИТМО (Алексей Сергушичев) и *Center for Genomic Regulation, Barcelona* (Герман Демидов). (<http://bioinformaticsinstitute.ru/sbw2017>).

До этого Леша Сергушичев и братья Артемовы **совместно с Институтом биоинформатики** провели международное мероприятие – семинар по системной биологии в Австралии (<https://register.gimr.garvan.org.au/systemsbiology/>).

Екатерина Эсаулова

Выпускница магистратуры и сотрудница нашей международной научной лаборатории Екатерина Эсаулова поступила в аспирантуру Университета Вашингтона в Сент-Луисе (100-е место в рейтинге QS).

Вот, что в свое время писала Катя. «Я училась на матмехе СПбГУ по специальности «Прикладная математика и информатика». На третьем курсе нужно было определиться с темой курсовой и диплома, и мой научный руководитель предложил биоинформатический проект. Тогда я совсем не знала, что это такое. Мне понравилась тема, я узнала, что в Санкт-Петербурге есть Институт биоинформатики и летняя школа по биоинформатике. Решила попробовать.

В итоге попала и в Институт, и в школу, а знания, полученные в них, так контрастировали с формальной, аккуратной и неизменной математикой, которая была на матмехе, что я была в восторге! Поняла, что биология – это то, к чему я хочу приложить свои умения. В Институте биоинформатики потрясающие преподаватели, которые помогали мне с биологией и программированием, а также вдохновляли на учебу, – я даже не думала о том, чтобы пропускать пары. Когда есть проект, который тебе интересен, все навыки начинают иметь смысл и, кажется, начинается настоящая учеба.

Я окончила программу «Алгоритмическая биоинформатика» Института биоинформатики в 2015 г. и хотела продолжить учиться и работать в этой области. Для этого я поступила в магистратуру по специальности «Прикладная математика и информатика» на кафедру «Компьютерные технологии» Университета ИТМО и встретила там нового научного руководителя – Алексея Сергушичева, который предложил мне работу, и я согласилась. Сейчас я *Visiting Scientist* в [Washington University in St. Louis](http://www.washington.edu) и занимаюсь системной биологией – анализом экспрессии генов, изучением предсказаний и объяснений ответа рака на иммунотерапию. Находиться в среде иммунологов, видеть, как проводят настоящие эксперименты, формулируют гипотезы, планируют исследования, встречаться с крупными учеными – очень интересно. Два года назад, когда я начинала, я не могла даже мечтать о таком. И еще хочу подчеркнуть, что все это **было бы сложно повернуть без тех людей, которых я встретила на кафедре КТ, которые мало того, что любят свою работу, но еще и активно помогают другим**».

Катя на защите магистерской диссертации говорила так быстро и четко, что, когда она закончила, я спросил у нее ни китайка ли она, так как видел, что многие из них выступают в таком же стиле. Выяснилось, что нет :-). После этого я спросил ее о том, может ли она также бодро выступать по-английски. Она ответила: «К сожалению, да» и увидев мое удивленное лицо, продолжила: «Они часто меня не понимают». Катя продолжит сотрудничество с нами в рамках международной научной лаборатории.

А с вопросом об сотрудничестве дело было обстоит так. Когда стало известно, что Катя уезжает работать в Америку, я спросил Лешу Сергушичева о том, будет ли она продолжать с нами работать. Леша **предложил «отстать» от Кати**, так как в бакалавриате у нас она не училась, а почти все время в магистратуре пробыла в Сент-Луисе. Обеспечение такой возможности при обучении в магистратуре я посчитал серьезным нашим вкладом в возможность дальнейшего сотрудничества и поэтому решил поговорить с Катей.

Разговор оказался на удивление коротким. Только я успел сказать, что Максим Артемов считает себя обязанным России за предоставленную ему возможность бесплатно получить прекрасное образование на химфаке МГУ, как **Катя ответила, что она тоже обязана, но нам, так как без нас у нее в биоинформатике ничего бы не получилось**.

В приведенном выше списке статей, опубликованных в 2017 г. с нашим участием, указаны три (!) публикации Кати – в журналах *Science*, *Nature Microbiology* и *Genomic Medicine*. Два последних журнала новые, и поэтому у них в этом году еще не было *IF* и *SJR*, но представляете какие значения этих показателей в них будут через пару лет!

В Катиной статье с нашей аффилиацией в журнале *Science* экспериментально доказано, что **кишечная микробиота мышей влияет на формирование иммунного ответа организма при заражении вирусом гриппа. Метаболиты, которые образуются в результате деятельности кишечных бактерий, стимулируют выработку интерферона. Этот белок подавляет размножение вируса.** (http://news.ifmo.ru/ru/science/life_science/news/6856/).

Сотрудничество с Катей продолжилось и в 2018 г. – благодаря ей аффилиация нашего университета снова появилась в одном из лучших журналов мира – *Cell*! После этого у Кати появился новый научный руководитель, и наше сотрудничество прервалось, что и не мудрено, так как частный *Washington University in St. Louis*, в отличие от других известных в мире университетов, не захотел иметь с нами двойную аспирантуру. Однако мы признательны Кате за ее четыре статьи в указанных классных журналах, тем более что статья в *Science* открыла нашему университету попадание в течение нескольких лет в Шанхайский рейтинг университетов мира, в котором статьи в журналах *Science* и *Nature* приносят «зачетные очки».

Про попадание в Шанхайский рейтинг более подробно. 14 августа 2018 г. наш университет в целом, а не по отдельным предметам, впервые стал виден в этом рейтинге (http://news.ifmo.ru/ru/university_live/ratings/news/7768/). У него пока 801-900 место. Ключевым фактором попадания в этот рейтинг стало наличие у сотрудников университета **пяти публикаций** в журналах *Science* (три публикации) и *Nature* (две публикации), которые вышли в 2016 и 2017 гг. Среди этих статей в *Science* и статья с участием Кати.

Я поблагодарил за это не только ее, но Максима Артемова – главного организатора всех наших последних побед в этой области. Катя написала мне: «Спасибо Вам за возможность работать над такими проектами!». Максим же ответил: «Кстати, вчера приняли Катину статью в журнал *Cell* (<https://www.sciencedirect.com/science/article/pii/S009286741831242X>). Это дополнительно к весенней Лешиной статье в *Nature* :-). Ну, и прочие по мелочи... (знаю я эти Мелочи, А.Ш.). Так что в следующем рейтинге, думаю, Университет ИТМО тоже не затеряется :-)). На это я ответил: «Максим! Спасибо за все! И я стараюсь по всему «фронту» (хотел написать по всем «статьям», но понял, что это неправильно – здесь сражаешься ты и Леша), а я, например, недавно долго уговаривал Жавлона Исомуродова поступать к нам в аспирантуру по вашей с Лешей тематике. Алексей, который при этом присутствовал, **сказал, что это был мастер-класс в области уговоров**». Про Жавлона написано ниже.

В октябре 2018 г. упомянутая выше статья в Катиним участием была опубликована в журнале *Cell*, у которого **IF: 31.398, SJR: 25.14**. Я поблагодарил Катю за это. Она написала: «Рада стараться». «А уж как мы рады твоим стараниям», – ответил я.

После этого я послал информацию об этом достижении Илье Куфтырёву, отвечающему в университете за рейтингование. Он ответил: «В прошлом году она была еще магистранткой. И вот благодаря таким «простым магистрантам» Университет ИТМО попал в Шанхайский рейтинг. Вот и думай после этого, «кто более матери – истории ценен?» :-). А к Вам Анатолий Абрамович, за то, что уговорили Катю продолжать по совместительству работать в нашем университете, применимы слова Роберта Рождественского: «...и на всей Земле не хватило мрамора, / чтобы вырубить парня в полный рост».

Жавлон Исомуродов

С 24 по 29 сентября Жавлон принимал участие *Heidelberg Laureate Forum (HLF)* – <http://www.heidelberg-laureate-forum.org>. Гейдельберг – старый университетский город в Германии. Это встреча крупнейших ученых в области математики и *Computer Science* (лауреатов премии Филдса, Тьюринга, Абеля и Неванлинны) с двумястами одаренными молодыми учеными (студентами, аспирантами и постдоками) со всего мира, прошедшими строгий отбор.

Беседа с Жавлоном после возвращения с форума опубликована здесь: <http://news.ifmo.ru/ru/education/official/news/7911/>. Интересно, что на первой фотографии в этой статье с молодежью общается лауреат премии Тьюринга **Джон Хопкрофт. Он почетный доктор Университета ИТМО с 2009 г.** (http://is.ifmo.ru/important/_hopcroft_itmo.pdf).

В одной из комнат нашего научного центра «Компьютерные технологии» (лаборатория в последнее время превратилась в центр) в треугольнике стоят столы Леша Сергушичева, Саши Лободы и Жавлона. Как-то я сказал двум последним, что у Вас здесь коэффициент умственных способностей, похоже, зашкаливает. Ребята скромно ответили, что так оно и есть, но только когда Леша на рабочем месте, а он в это время с Максимом Артемовым проводил семинар по системной биологии в Ницце!

Жавлон очень способный человек – еще в школе он стал бронзовым призером Международной (всемирной) олимпиады по математике. В студенчестве, работая в нашей лаборатории, он создал совместно с Лешей Сергушичевым и Сашей Лободой продукт **mcmcRanking**: «Tool for estimate probabilities of vertices being in active module using its likelihoods and it proposes methods for ranking vertices in order of importance. Estimating probabilities based on Markov chain Monte Carlo (MCMC) methods» (<https://github.com/ctlab/mcmcRanking/>).

В ноябре Жавлон Исомуродов проходил стажировку в качестве *Visiting Researcher в Institute for Molecular Medicine Finland* под руководством Марка Дейли и Никиты Артемова. Он занимался разработкой методов для типирования *Human Leukocyte Antigen (HLA)*. Тип *HLA* является

критической информацией, необходимой для лечения раковых заболеваний с помощью современных препаратов для иммунотерапии. В клинической практике типирование *HLA* требует высокой точности и обходится достаточно дорого. Жавлон использовал данные *Next Generation Sequencing* для построения алгоритма, позволяющего предсказывать тип *HLA* с достаточной для клинической практики точностью, что должно значительно удешевить иммунотерапию рака.

После моих уговоров Жавлон в аспирантуру поступил, сказав, правда, что ничего не обещает. Он не обманул :-)) – несмотря на успехи в науке через непродолжительное время он ушел работать в промышленность. Жаль...

О факультете биоинформатики в Университете ИТМО

«Разобравшись» с биологией :-)), я «взялся» за химию. Первое мое предложение в январе 2017 г. в этой области **Владимир Николаевич Васильев воспринял весьма скептически**, но через несколько дней я «пристал» к одному из двух братьев Виноградовых, которые организовали в нашем университете международную лабораторию по-новому для нас направлению – химии. Стоя рядом с проходной, я рассказал Александру чем мы занимаемся, и его заинтересовали наши программисты, а самое главное – исследования Леши Сергушичева. С лабораторией Виноградовых сотрудничает Евгений Пидко из Университета Эйнховена, который через несколько дней встретился с Алексеем и нашел область для их совместной деятельности.

Уже 31 января делая пленарный доклад на научной и учебно-методической конференции Университета ИТМО, Евгений упомянул об Алексее и сказал, что **у нас в университете хороший факультет биоинформатики**. Интересно, что и в следующем пленарном докладе, выступающий из Университета ИТМО тоже сказал, что хочет сотрудничать с факультетом биоинформатики. После этого я поинтересовался, что они имеют в виду, говоря о факультете, и оказалось – группу Леши Сергушичева, высокорейтинговые статьи которого «возвышают» его группу до факультета.

Интересно, что на этой конференции выяснилось, что **В.Н. Васильев уже знал о моем «движении» в сторону братьев Виноградовых и Пидко**, и воспринял «мой поход» в химию значительно более благожелательно, чем когда я в свое время «двинулся» в биологию. Меня не удивило то, что ректор уже знает об этом, так как я уже много лет под впечатлением об истории, озвученной **Аркадием Ивановичем Вольским**, который устраиваясь на работу к **Юрию Владимировичу Андропову**, начал рассказывать ему о себе, но Генеральный секретарь его остановил, задав вопрос: **«Неужели Вы думаете, что знаете о себе больше, чем мы о Вас?»** А еще Пидко сказал мне, что ему очень нравится обстановка в нашем университете – никто с ним сразу не говорит ни о деньгах, ни о статьях, как это происходит обычно.

Совещание по ИТ-образованию в Вышке

20 апреля 2017 г. в **Высшей Школе Экономики (ВШЭ)** прошел круглый стол на тему **«Будущее ИТ-образования в России»**, которая с моей «легкой руки» сузилась до вопроса: **«Кому готовить ИТ-спецназ?»** (<https://postgrespro.ru/blog/company/202540>). Он был организован компанией *Postgres Professional*, ВШЭ и Центром развития ИТ-образования МФТИ. Подготовку профессионалов для ИТ-отрасли собрались обсудить представители ведущих российских ВУЗов (МГУ, ВШЭ, МФТИ, Университета ИТМО) и ИТ-компаний, развивающих собственные образовательные проекты (<https://postgrespro.ru/blog/news/201373>, <https://postgrespro.ru/events/202543>).

При этом первый доклад был сделан деканом факультета «Компьютерные науки» Вышки Иваном Аржанцевым. Это очень сильный факультет, много научных лабораторий, они тоже занимаются наукой в рамках программы «5 в 100».

Я там выступал около получаса (<https://www.youtube.com/watch?v=tDaXU04-zH8>). В своем выступлении (https://www.youtube.com/watch?v=qc_mtnGwpA&t=100s) я обратил внимание на то, что у нас с москвичами разные точки зрения на организацию образовательного процесса и не только его. Все эти московские ВУЗы (МГУ, Вышку, МФТИ) «обслуживает», как это было у нас в 90-е годы, сборная команда преподавателей Москвы, которые работают в «Яндексе» и многочисленных институтах Академии наук, и они ездят с одного места на другое, и имеют поэтому по одной публикации в год, потому что им некогда – они должны стремительно перемещаться по Москве, а там огромные пробки :-)).

И еще. Например, в Вышке есть лаборатория биоинформатики и ей по совместительству руководит Михаил Гельфанд, а не выпускник или сотрудник Вышки. А у меня же цель совершенно иная: **я хочу, чтобы у нас учеными и преподавателями были, по возможности, наши выпускники** – они обычно очень сильные, и мы их за годы учебы хорошо узнаем. При этом отмечу, что у нас тоже есть лаборатория биоинформатики, но ей руководит, не «пришелец», а наш выпускник Леша Сергушичев, который, кстати, «из рук» Гельфанда «получил» *Research Fellowships Skoltech* по... системной биологии.

Саша Лобода

В 2017 г. Саша (<https://scholar.google.ru/citations?user=AYBD9cAAAAJ&hl=ru&oi=ao>) уехал на длительную стажировку в *Broad Institute of MIT and Harvard* (Cambridge, MA) в лабораторию популяционной генетики Марка Дейли (Mark Daly), который в то время имел индекс Хирша по *Google Scholar* **168** (!) – 96-ой в истории человечества, и 200 000 цитирований. На первом месте по этому индексу был **Зигмунд Фрейд**, у которого он был равен 272. Время идет. Фрейду весьма трудно держаться :-)) на первом месте, и к 22.01.2020 г. он опустился на пятое место (<https://www.webometrics.info/en/hlargerthan100>), а Марк с индексом Хирша **189** и числом цитирований 254 892 поднялся на 83 место! 30.06.2023 г. Марк уже имел следующие значения этих показателей: **231** и 400 083!

При этом отмечу, что Марк по результатам стажировки Саши был готов дать ему рекомендацию в аспирантуру, но она не потребовалась, так как молодой человек поступил в аспирантуру к нам, а мы уже много лет его хорошо знаем! Сейчас он работает сразу над несколькими проектами. Один из них связан с выводом регуляторных сетей и был начат еще в Университете ИТМО. Вторая работа состоит в разработке метода, который позволяет подбирать генотипы здоровых людей для анализа полногеномного поиска ассоциаций, не передавая конфиденциальную информацию о пациентах. Цель еще одного Сашиного проекта – создание метода для определения функциональной зависимости между генами и некодирующими мутациями.

В 2018 г. появился препринт, написанный в таком (!) составе: *Artomov M., Loboda A., Artyomov M., Daly M.* A platform for case-control matching enables association studies without genotype sharing (<https://www.biorxiv.org/content/early/2018/11/14/470450>). Еще никто из наших ребят не имел такого (!) соавтора, как Daly, а мой «внук» – Саша Лобода (если Лешу Сергушичева считать «сыном») – заимел! Классно! Потом этого соавтора приобрел и Леша Сергушичев.

С 15 по 19 октября Саша участвовал на конференции «American society of human genetic (ASHG) 2019 Annual Meeting» в Хьюстоне (<https://www.ashg.org/2019meeting/>), где выступил с постерным докладом *Loboda A., Artomov M., Daly M., Sergushichev A.* Algorithm for gene regulatory network inference recovers biological insights from large-scale gene expression data, который был посвящен тому, как выводить сеть прямых регуляций генной активности по данным РНК-секвенирования (<https://eventpilotadmin.com/web/page.php?page=Session&project=ASHG19&id=30309028>). У Саши там был и второй доклад: *Artomov M., Loboda A., Artyomov M., Daly M.* Public platform with 42,291 exome control samples enables association studies without genotype sharing (<https://eventpilotadmin.com/web/page.php?page=Session&project=ASHG19&id=140003>), который сделал брат Максима Артемова – Никита (*Mykyta Artomov*).

Все публикации Саши приведены здесь: https://www.researchgate.net/profile/Alexander_Loboda, а он сам с сентября 2020 г. – в Университете ИТМО. В конце 2022 г. Саша стал там кандидатом наук и *PhD* (<https://dissovet.itmo.ru/dissertation/?number=749244>)!

Снова о Леше Сергушичеве

После защиты диссертации Алексей не исключал отъезда постдоком за границу, но **не куда-нибудь, что устраивает многих**, а только в Бостон, так как это мировая столица по исследованиям иммунологии рака – там находятся МТИ, Гарвард, известные институты и клиники. **Ну, что тут скажешь? Этому, вроде бы, я и учил...**

Зная желание Алексея, я так просто не сдался, и снова старался убедить его в том, что в Университете ИТМО он лет на десять раньше станет профессором, и что если Алексей, все-таки, уедет, то хорошо бы ему остаться у нас работать по совместительству. Кстати, из разговора с его научным руководителем Максимом Артемовым выяснилось, что далеко не везде Леше позволят это

сделать, но такое место, где Леше сразу разрешат с нами сотрудничать, существует – это Университет Вашингтона в Сент-Луисе, в котором за последние годы Алексей зарекомендовал себя как отличный ученый.

А пока я задавал Леше провокационные вопросы: «А, где ты найдешь студентов такой силы, как те, которые у тебя сейчас работают?» или «А ты понимаешь, что на новом месте ты минимум два-три года будешь завоевывать **свое** место под солнцем, а когда это получится, придется идти куда-то дальше и там снова завоевывать **свое** место?». При этом мне снова вспомнилась женщина, которая сказала: «**Зачем мне путешествовать, если я уже здесь?**».

Все взвесив, особенно то, что мы с ним недавно выиграли конкурс по его тематике на получение госзадания на три года, где он должен был быть ответственным исполнителем, Алексей 17 января отправил мне и коллегам письмо: «Я более-менее пришел к тому выводу, что все-таки пока остаюсь в Университете ИТМО как на основном месте работы, и в своих действиях буду исходить из того, что еще как минимум года два-три я буду находиться здесь».

На это Максим Артемов написал: «Поздравляю Университет ИТМО и Лешу! Будут очень продуктивные пара-тройка лет, начиная с этого момента! В течение недели мы **засабмитим статью в Nature** (!) в коллаборации с *Marco Colonna*. Это *competitive topic* и параллельная статья от конкурирующей группы только что ушла на ревью туда же – редакторы ждут теперь нас. Так что – за взятие новых высот с новой степенью!».

Кстати, Максим в марте 2017 г. выступил с лекцией в ... Эрмитаже, на которой рассказывал о живописи и ... иммунологии (<http://news.ifmo.ru/ru/education/cooperation/news/6524/>). До этого он выступал с гостевыми лекциями в Институте биоинформатики: «Системная биология» (13.09.2014, <https://www.youtube.com/watch?v=6bIbI5jSnNI&t=39s>); «Иммунотерапия раковых опухолей с точки зрения системной биологии» (16.07.2015г., <https://www.youtube.com/watch?v=og4eRWp2vMk&t=62s>); «Метаболическая регуляция иммунного ответа» (16.01.2016 г., <https://www.youtube.com/watch?v=ILXWMAeFaf0&t=419s>).

А вот, что написал я после письма, полученного от Алексея: «Спасибо, Леша, за ответ, тем более в западном стиле – письменно. Есть теория малых дел, с помощью которых можно делать большие. Когда уходил Федя Царев, я со слезами на глазах просил тебя остаться на три года. Ты пообещал и выполнил. Я со своей стороны тоже сделал все, что мог. Теперь моих слез не потребовалось, и ты сам решил остаться на два-три года! Классно! **А там, смотришь, немного останется до защиты докторской, а потом и до получения должности профессора в Университете ИТМО, а затем и этого звания**».

При этом отмечу, что этому решению Алексея предшествовал разговор за несколько часов до его защиты, в котором приняли участие я, Леша и Максим Артемов. Во-первых, Максим поблагодарил меня, что я безропотно согласился признать его научным руководителем Лешу несмотря на то, что почти до самой защиты считалось, что его руководитель я.

После этого Максим сказал: «Анатолий Абрамович! Отпустите Лешу – он по этой тематике перерос Вас и начинает перерастать меня». На это я ответил, что я не бабушка от Достоевского, а Алексей – не ее внучка, и, тем более, не моя. У меня и Лешу нет юбок, которые мы скрепили бы булавкой, поэтому он абсолютно свободен и может ехать куда ему угодно. Однако нужно ли ему это делать – большой вопрос? После этого Максим стал расписывать перспективы, которые откроются перед Алексеем, если он окажется Там.

На это я спросил Максима сколько ему лет и какую должность он занимает. Оказалось 34 года, и он *Assistant Professor*. Максим мне этого не говорил, но большинство наших соотечественников, занимающих эту должность, как, впрочем, и должность *Associate Professor*, первое слово в их названиях обычно опускают. Это во многом связано с тем, что в США «профессором» называют любого преподавателя университета или колледжа, кроме аспирантов, привлекаемых к учебному процессу. Но я думаю, что причина не только в этом...

После этого я высказал предположение, что так как Максим в этой должности уже находится весьма продолжительное время (обычно требуется проработать в ней пять-шесть лет), то он скоро станет *Associate Professor*, что и произошло, а вот еще через пять-шесть лет, когда ему будет за сорок, станет

ли он *Full Professor*, даже учитывая все его научные достижения – еще вопрос? (Максима можно поздравить – этот вопрос успешно разрешился). Потом я сказал, что у Леши в Университет ИТМО это может получиться значительно быстрее. С этим Максим не стал спорить, и на этой оптимистической ноте мы пошли на защиту Леши.

У Алексей пока все складывается, как я и предсказал. Как только он получил диплом кандидата наук, то практически сразу был избран доцентом (*Associate Professor*) на нашей кафедре, минуя, должность ассистента, в то время как в Америке после защиты диссертации человек, который хочет продолжить академическую карьеру, обычно переходит в другой университет: либо на должность *Assistant Professor* (а то и на должность *Research Assistant Professor*) в относительно слабый университет, либо на три-четыре года постдоком в хороший университет, и только после этого – на должность *Assistant Professor* в другой хороший университет. Когда я сказал Леше, что мы помогли ему сэкономить несколько лет жизни, он без затей ответил: «Поэтому здесь и хожу!». Надеюсь, что не только по этой причине...

Шестого апреля Алексей вошел в пятерку лучших молодых ученых в стране по **системной биологии** (<https://www.pm.skoltech.ru/>) и попал в **Skoltech Fellowship Program 2017** (<http://www.skoltech.ru/en/2017/04/the-winners-of-the-systems-biology-fellowship-program-2/>), о чем я уже рассказывал.

В августе 2017 г. у Леши и его ученика **Саши Лободы**, как отмечено выше, вышла статья в журнале **Cell** с обалденными показателями: **IF: 30.41. SJR: 26.95!**, а вот его не менее обалденный список публикаций в *US National Library of Medicine National Institutes of Health*: <https://pubmed.ncbi.nlm.nih.gov/?term=Sergushichev+A>.

Седьмого сентября 2020 г. я посмотрел профиль Алексея на *Research Gate*. Он у него весьма высок – 31.64 (https://www.researchgate.net/profile/Alexey_Sergushichev).

При этом очень важно, что мама Леши понимает, насколько больших успехов добился ее сын! Я спросил Алексея, как она отнеслась к тому, что он имея возможность уехать почти куда угодно, остался на постоянной работе в российском университете. Молодой человек гордо ответил: «У меня мама – доцент. Она понимает каков уровень моих публикаций». Побольше бы таких мам!

А вот крохотка под названием «Их нравы», которую я опубликовал в Интернете (<http://is.ifmo.ru/belletristic/nravy>): «Я написал Леше Сергушичеву, что начинает выходить журнал *Nature Machine Intelligence*. Это его не заинтересовало, так как, по его словам, **ему и обычного Nature хватает :-)**!».

Геннадий Короткевич

Гена (https://ru.wikipedia.org/wiki/Короткевич,_Геннадий_Владимирович) написал под руководством Леши Сергушичева магистерскую диссертацию на тему «**Алгоритмы для эффективного анализа представленности функциональных наборов генов**». В задании на работу сказано: «Требуется разработать алгоритмы для эффективного анализа представленности функциональных наборов генов, превосходящие использованные в прототипе – *Fast Gene Set Enrichment Analysis (FGSEA)*, который не позволяет вычислить точные *P*-значения функциональных наборов генов, а также сколько угодно малые их *P*-значения с хорошей относительной точностью».

В работе Геннадий разработал точный полиномиальный алгоритм вычисления *P*-значения функциональных наборов генов при условии, что ранги генов – целые числа, и приближенный алгоритм, позволяющий вычислять сколько угодно малые *P*-значения с хорошей абсолютной точностью. Выполнено сравнение этих алгоритмов между собой и с алгоритмом *FGSEA*. Установлено, что возможности, предоставляемые предложенным алгоритмом недоступны для ранее созданных алгоритмов представленности.

Еще до защиты магистерской я спросил Лешу: «Ты знаешь, чем Гену занять после диплома?», а потом **предложил познакомить Геннадия с Марком Дейли** – выдающимся ученым, математиком, занимающимся генетикой. Ему 50 лет. Он по цитируемости входит в сотню ученых мира, в том числе и умерших – на первом месте Фрейд. Марк в 2019 г. переехал на работу в *Institute for Molecular Medicine Finland* в Хельсинки, что упрощает коммуникацию с ним.

Леша знаком с Марком, и он с моей «подачи» написал ему письмо, в котором была приведена информация об успехах Гены, включая ссылку на статью о нем в «Википедии». **Марк восхитился возможностью работать с таким талантом и предложил встретиться.** При этом **я сказал Геннадию, что у него появляется шанс ненавязчиво второй раз войти в историю человечества**, решив еще две-три научные задачи к тем десяткам тысяч, которые он уже решил. Финляндия под боком, можно жить здесь, а туда ездить к научному руководителю. Наши ребята уже неоднократно так делали и делают.

Вот письмо от Леша Сергушичева Марку, которое «спровоцировал» я: «Mark, I understand you're still moving back and forth and it's probably not the best time, but I want to ask you an advice or may be a little favor. Here at ITMO we have a student, Gennady Korotkevich, who graduates this year from our master program (he is a groupmate of Alex Loboda). What's interesting about Gennady is that he is really a genius and is the best in competitive programming in the world: you can check out his Wikipedia page that lists major competitions he won (https://en.wikipedia.org/wiki/Gennady_Korotkevich). What's also great about him is that he is quite normal personally (we've seen a lot of top programmers here and with many of them it can be quite challenging to talk to, so this means a lot).

However, he's unsure on what to do after graduating. He, of course has a lot of job offers from IT companies like Google, but apparently this isn't what he wants. On the other hand, we tried to engage him in doing science, but it doesn't work that well either. Max was trying to work with him at some point, but it didn't get anywhere. It seems that he is really good at solving smallish problems that can be done in a few hours (by him). Bug he did help me a lot in the time of my PhD: after an hour talk with him he suggested a simple algorithm for solving the problem of cumulative GSEA statistic calculation, with which I was struggling for a few days. In the end it resulted in the FGSEA method that is hundreds time faster compared to the Broad's one due to an algorithmic improvement (<https://www.biorxiv.org/content/early/2016/06/20/060012>). After that we've been trying to take it even further, but it's not going that well, probably due to it being more of a research-style problem and me not being engaging enough.

So, my question is: would you be able to meet and talk with him at some point, may be give him some advice? I understand you are frequent at Helsinki now, maybe you would have a free hour? I believe that it would be really great for everyone if you'd find a common language with him. Best, Alexey».

Теперь приведу ответ на это письмо: «Hey Alexey – I would love to talk with Gennady – he sounds fantastic! I am just now on my way back to Boston from Helsinki – will be back in a few weeks, but glad also to make contact by phone sooner if that would be useful. I'll also try to think of some problems that might interest him :) Best wishes Mark».

Их встреча в Хельсинки состоялась, и теперь у Гены есть возможность ненапряжно войти в историю человечества второй раз. Воспользуется ли он этим шансом? Мы с Лешей свое дело в этом вопросе сделали...

Геннадий Короткевич о решении задач по биоинформатике

Вот, что сказал Гена на эту тему (http://news.ifmo.ru/ru/university_live/achievements/news/8550/): «Мой научный руководитель – Леша Сергушичев (он продуктивно занимается биоинформатикой), иногда встречается с задачами, которые имеют какой-то алгоритмический подтекст. Понятно, что комплексно научная деятельность – это не только алгоритмы и теория. Биоинформатика очень практическая область, но Леша в какой-то момент нашел задачу и спросил, как к ней можно подступиться, и я пару месяцев над ней в фоновом режиме думал. Получилось так, что задача была про биоинформатику, но та постановка, которую я получил от Леша, была достаточно абстрактной, что мне интереснее, чем ее приложение. Может быть, помогло то, что Леша сам участвовал в чемпионатах по программированию.

Мне кажется, что люди, которые занимаются наукой, даже если не участвовали раньше в олимпиадах по информатике и программированию, хорошо понимают, что такое алгоритмические задачи. Для этого не нужно быть чемпионом мира. Многие люди, которые занимаются биологией, физикой, сталкиваются с программированием, и должны уметь программировать. Конечно, не обязательно на очень высоком уровне. И наоборот, есть много вещей в биологии, которые люди из области программирования, даже чемпионы, не смогут сразу понять.

Недавно опять встретил Лешу, и он предложил еще одну задачу, которую, может быть, полезно было бы решить. С точки зрения науки, люди, которые занимаются олимпиадами, хороши в том, чтобы решать очень трудные задачи, если они «удачно» сформулированы.

Мне кажется, что можно было бы придумать удобный формат для того, чтобы ученые могли формулировать задачи, которые у них возникают во время исследований, и передавать их людям, которые решили в жизни много алгоритмических задач. Занятие наукой – это не только решение трудных задач. Это сложный процесс: эксперименты, идеи, публикации, бюрократия. Говорят, что в бизнесе мало трудных задач, а здесь – много. Хотя, кажется, в науке и платят поменьше».

Потом научным руководителем Гены стал я, но еще не вечер...

Bioinformatics Contest 2017

Студенты и сотрудники нашей кафедры совместно с Институтом биоинформатики провели первый **Bioinformatics Contest 2017** (<http://contest.bioinf.me/>), в котором победу одержал Геннадий Короткевич! (<http://mon.stepik.org/>, <http://contest.bioinf.me/results>). Задачи были составлены командой из Университета ИТМО в составе: Алексей Сергушичев, **Виталий Демьянюк**, **Анна Малова**, **Григорий Шовкопляс**, **Артем Васильев**, **Виталий Аксенов** и **Илья Збань**.

Контеcт состоял из двух раундов: квалификационного, который проходил в течение недели с 23 по 29 января, и финального, проходившего 18 февраля в течение 24 часов (<http://news.ifmo.ru/ru/science/it/news/6495/>).

В соревнованиях приняло участие более 3000 человек из 86 стран: 685 – США, 477 – Россия, 126 – Великобритания, 117 – Индия, 105 – Германия, 85 – Франция, 74 – Канада, 51 – Испания, 41 – Бангладеш, 39 – Польша, 37 – Бразилия, Украина, 34 – Нидерланды, 32 – Китай, 30 – Беларусь, 29 – Швеция, 27 – Турция, Швейцария, 25 – Чехия, Южная Корея, 20 – Австралия, 19 – Япония, 18 – Пакистан, 16 – Египет, Сербия, Дания, Вьетнам, Южная Африка, 15 – Бельгия, Италия, 14 – Португалия, Болгария, 13 – Мексика, Сингапур, Ирландия, Тайвань, 12 – Эстония, Индонезия, 11 – Греция, 10 – Австрия, Хорватия, Таиланд, 9 – Аргентина, Иран, Венгрия, 8 – Новая Зеландия, Румыния, Финляндия, Армения, 7 – Норвегия, Чили, Колумбия, Перу, 6 – Израиль, Гонконг, 5 – Малазия, Филиппины, Макао, Словения, 4 – Литва, Тунис, 3 – Саудовская Аравия, Казахстан, Словакия, 2 – Катар, Сирия, Боливия, Молдова, Азербайджан, Шри Ланка, 1 – Нигерия, Узбекистан, Уругвай, Сант Винсент и Гренадины, Иордания, Грузия, Ямайка, Гана, Судан, Туркменистан, Люксембург, Босния и Герцеговина, Малави, Исландия, Арабские эмираты, Куба.

В финал вышло 379 участников, но только 177 из них получили зачетные баллы. (http://ct.ifmo.ru/ru/viewnews/3660/Bioinformatics_Contest_2017_pobeda_za_nami.htm).

2018 год

Статьи в выдающихся журналах

1. Bambouskova M., L. Gorvel L., Lampropoulou V., **Sergushichev A.**, Loginicheva E., Johnson K., Korenfenld D., Mathyer M., Kim H., Huang L., Duncan D., Bregman H., Keskin A., Santeford A., Apte R., Sehgal R., Johnson B., Amarasinghe G., Soares M., Satoh T., Akira S., Hai T., Strong C., Auclair K., Roddy T., Biller S., Jovanovic M., Klechevsky E., Stewart K., Randolph G., Artyomov M., Electrophilic Properties of Itaconate and Derivatives regulate Ikbz/Atf3 in Inflammatory Axis // **Nature**. 2018. Vol. 556. No. 7702, pp. 501-504, doi:10.1038/s41586-018-0052-z. **IF: 40.137, SJR: 18.13** (<https://www.nature.com/articles/s41586-018-0052-z>).
2. Gubin M., **Esaulova E.**, Ward J., Malkova O., Runci D., Wong P., Noguchi T., Arthur C., Meng W., Alspach E., Medrano R., Fronick C., Fehlings M., Newell E., Fulton R., Sheehan K., Oh S., Schreiber R., Artyomov M. High-Dimensional Analysis Delineates Myeloid and Lymphoid Compartment Remodeling during Successful Immune-Checkpoint Cancer Therapy // **Cell**. Vol.175. 2018.Issue 4, pp.1014-1030.e19. **IF: 31.398, SJR: 25.14** (<https://www.sciencedirect.com/science/article/pii/S009286741831242X>).
3. Gorman M., Caine E., **Zaitsev K.**, Begley M., Weger-Lucarelli J., Uccellini M., Tripathi S., Morrison J., Yount B., Dinno K., Rückert C., Young M., Zhu Z., Robertson S., McNally K., Ye J., Cao B., Mysorekar I., Ebel G., Baric R., Best S., **Artyomov M.**, Garcia-Sastre A., Diamond M. An Immunocompetent Mouse Model of Zika Virus Infection // *Cell Host&Microbe*. 2018. 23(5), pp. 672-685.e6. **IF: 14.946, SJR: 8.34** (<https://vk.com/@itmoru-uchenye-uluchshili-model-dlya-izucheniya-virusa-zika>). В статье изложена новая модель для изучения патогенеза вируса Зика (https://news.itmo.ru/ru/science/life_science/news/7552/).

4. Howard N., Marin N., Ahmed M., Rosa B., Martin J., Bambouskova M., **Sergushichev A.**, Loginicheva E., Kurepina N., Rangel-Moreno J., Chen L., Kreiswirth B., Klein R., Torrelles J., Balada-Llasat J., Amarasinghe G., Mitreva M., Artyomov M., Hsu F., Mathema B., Khader S. Mycobacterium Tuberculosis Carrying a Rifampicin Drug Resistance Mutation Reprograms Macrophage Metabolism Through Cell Wall Lipid Changes // **Nature Microbiology**. 2018. № 3, pp.1099-1108. **IF: 14.182, SJR: 7.14** (<https://www.nature.com/articles/s41564-018-0245-0>).
5. Kim K., Shim D., Lee J., **Zaitsev K.**, Williams J., Kim K., Jang M., Jang H., Yun T., Lee S., Yoon W., Prat A., Seidah A., Choi J., Lee S., Yoon S., Nam J., Seong J., Oh G., Randolph G., Artyomov M., Cheong C., Choi J. Transcriptome Analysis Reveals Nonfoamy Rather Than Foamy Plaque Macrophages Are Proinflammatory in Atherosclerotic Murine Models // **Circulation Research**. 2018. V. 123. Issue 10, pp. 1127-1142. **IF: 15.211, SJR: 5.72** (<https://www.ahajournals.org/doi/full/10.1161/CIRCRESAHA.118.312804>).
6. Brähler S., Raju S., Saunders B., **Zaitsev K.**, Artyomov M., Zinselmeyer B., Murphy K., Miner J., Shaw A. et al. Opposing Roles of Dendritic Cell Subsets in Experimental Glomerulonephritis // **Journal of the American Society of Nephrology**. 2018. 29 (1), pp. 138-154. **IF: 8.966. SJR: 4.36** (<https://jasn.asnjournals.org/content/29/1/138>).
3. Freylikhman O., Kiselev A., **Kazakov S.**, **Sergushichev A.**, Panferova Y., Tokarevich N., Kostareva A. Draft Genome Sequence of *Coxiella burnetii* Historical Strain Leningrad-2, Isolated from Blood of a Patient with Acute Q Fever in Saint Petersburg, Russia // **Microbiology Resource Announcements** (Genome Announc.). 2018. 6(3). pii: e01464-17 (<https://www.ncbi.nlm.nih.gov/pubmed/29348343>).
4. Nair S., Huynh J.P., Lampropoulou V., Loginicheva E., Esaulova E., Gounder A.P., Boon A.C., Schwarzkopf E.A., Bradstreet T.R., Edelson B.T., Artyomov M.N., Stallings C.L., Diamond M.S. Irg1 Expression in Myeloid Cells Prevents Immunopathology During M. Tuberculosis Infection // **Journal of Experimental Medicine**. 2018, Vol. 215, No. 4, pp. 1035-1045 (<https://pubmed.ncbi.nlm.nih.gov/29511063/>).
9. Kiselev A., Kornishina T., **Sergushichev A.**, Smolina N., Klyushina A., Pervunina T., Bang M.L., Sjöberg G., Sejersen T., Kostareva A. New Variant in CMYA5 Gene is Associated with Early-Onset Restrictive Cardiomyopathy and Autism-Spectrum Disorder // **Cardiovascular Research**. 2018. Vol. 114, Issue suppl.1, pp. S19 (https://academic.oup.com/cardiovascres/article/114/suppl_1/S19/4981035).
10. Kostina A., Kiselev A., Bjorck H., Irtyuga O., **Sergushichev A.A.**, Baranov Y., Eriksson P., Kostareva A., Malashicheva A. Notch Signaling Pathway is Attenuated in Aortic Endothelial Cells of Patients with Aortic Pathologies Associated with Bicuspid Aortic Valve // **Cardiovascular Research**. 2018, Vol. 114, Issue suppl.1, pp. S83 (https://academic.oup.com/cardiovascres/article/114/suppl_1/S83/4981236).
11. **Alexeev N.**, Alekseyev M. Combinatorial Scoring of Phylogenetic Trees and Networks Based on Homoplasy-Free Characters // **Journal of Computational Biology**. 2018. Vol. 25, No. 11. **IF: 1.191** (<https://doi.org/10.1089/cmb.2018.0082>).
12. Artyomov M., **Loboda A.**, Artyomov M., Daly M. A platform for case-control matching enables association studies without genotype sharing // **bioRxiv**. Preprint (<https://europepmc.org/article/PPR/PPR62037>).

В последней работе среди наших соавторов Mark Daly: Google Scholar. h-index - 200!!! 28.06.2020.

Конференция

Zabelkin A., Alexeev N. Estimation of the True Evolutionary Distance Under the INFER Model / RECOMB-CG 2018. Lecture Notes in Computer Science. Vol. 11183, pp 72-87 (https://link.springer.com/chapter/10.1007%2F978-3-030-00834-5_4).

Свидетельства о регистрации программ

1. **Лобода А.А., Сергушичев А.А.** Программа для решения обобщенной задачи поиска связного подграфа максимального веса // Свидетельство о регистрации программы для ЭВМ. № 2018 612213. Дата регистрации – 14.02.2018, http://is.ifmo.ru/certificates/2018_02_14_loboda_sergushichev.
2. **Сергушичев А.А., Артемов М.Н., Зенкова Д.М., Каменев В.В., Саблина М.М., Шалыто А.А.** Программное средство для визуального и интерактивного анализа экспрессии генов // Свидетельство о регистрации программы для ЭВМ. № 2018 666621. Дата регистрации – 19.12.2018, <http://is.ifmo.ru/certificates/2018-12-19-sergushichev-et-al.PDF>.

ITMO Genome Assembler

В ноябре 2018 г. Сергей Казаков выложил на *GitHub* дополнительную информацию о нашем сборщике генома (<https://github.com/ctlab/itmo-assembler>): «**ITMO Genome Assembler** is a software for *de novo* assembling small and middle-sized genomes with ease. It had been developed by the team of researchers from *Genome Assembly Algorithms Laboratory* (now – bioinformatics group in *Computer Technologies Laboratory*) at *ITMO University* with the participation of the «Bioengineering» center of Russian Academy of Sciences. Active development of the software had been conducted from 2010 to 2016. Currently the main functionality of the assembler is supported, but without further improvements. Supported sequencing technologies – *Illumina/Solexa*, *Ion Torrent* (with indel errors) and *Sanger*. Assembler can use both paired reads (paired-end, mate-pair) as well as unpaired ones».

Говоря кратко и по-русски: Сборщик генома **ITMO Genome Assembler** – это программное средство для *de novo* сборки геномных последовательностей из чтений, полученных с секвенаторов *Illumina/Solexa*, *Ion Torrent* и *Sanger*.

Bioinformatics Contest 2018

Студенты и сотрудники нашей кафедры совместно с Институтом биоинформатики провели в конце февраля второй *Bioinformatics Contest* (<http://contest.bioinf.me/>), в котором Геннадий Короткевич занял третье место! (<http://mon.stepik.org/2018-finals.html>). Гена не победил на этот раз, во-первых, потому, что, по крайней мере, одна из задач финала (*Recombination of Plasmids*) составлялась так, чтобы ему из-за специфики было крайне трудно решить ее в полном объеме, а, во-вторых, Геннадий за отведенное время (сутки) участвовал еще в ... трех контестах. Задачи были составлены командой из Университета ИТМО под руководством Алексея Сергушичева. В соревнованиях участвовало более 3000 человек из 103 стран. При этом в финале 188 участников получили зачетные баллы.

Пятый семинар по системной биологии

В середине мая Максим Артемов, Алексей Сергушичев и Со провели под Санкт-Петербургом пятый семинар по системной биологии (<https://bioinf.me/education/sbw/2018>).

Участие в работе летней школы по биоинформатике

В конце июля проходила летняя школа по биоинформатике, проводимая, как обычно, Институтом биоинформатики. Школа в 2018 г. была посвящена применению биоинформатики в исследованиях рака. 100 участников из 30 городов и шести стран прошли конкурсный отбор, чтобы прослушать лекции ведущих ученых и воплотить в жизнь научные проекты. Наши (Алексей Сергушичев, Константин Зайцев, Никита Алексеев и Ярослав Баринов) и не только они, уже не первый год делились с участниками школы своим опытом и знаниями (<http://news.ifmo.ru/ru/education/cooperation/news/7802/>).

Обсуждение биоинформатики на канале «КритМышь»

Седьмого августа молодые сотрудники лаборатории «Компьютерные технологии» Константин Зайцев и Ярослав Баринов на канале «КритМышь» Университета ИТМО (канал для тех, кто мыслит критически) обсудили проблемы машинного обучения, нейросетей, искусственного интеллекта и биоинформатики. Все в этой беседе было хорошо за исключением самого первого момента, когда ведущий, представляя Костю, сказал, что он аспирант Университета Вашингтона в Сент-Луисе и сотрудник международной научной лаборатории «Компьютерные технологии» Университета ИТМО, и Костя кивком головы согласился с этим. Это, видимо, связано с тем, что обычно желаемое очень хочется выдать за действительное – аспирантом американского университета он не был и не стал им, хотя и собирался (https://vk.com/itmo.ctlab?z=video-166411935_456239036%2F3652c188ffc2eeb7bc%2Fpl_wall_-76507013).

Участие в биохакатоне

Шестого декабря в Сеченовском Университете завершился биохакатон по эволюционной биоинформатике. Участникам необходимо было выяснить, каким образом мобильные элементы (в частности, ретротранспозоны) оказывают влияние на процесс эволюции, на какие группы генов и

сигнальные пути они могут действовать. Наши студентки Анастасия Мурзина и Карина Пац входили в команду, занявшую второе место.

Nature Index

Издательский дом *Springer Nature* опубликовал индекс публикационной активности *Nature Index*, охватывающий период с 01.01.2018 г. по 31.12.2018 г. (http://news.ifmo.ru/ru/university_live/ratings/news/8590/). Рейтинг отражает этот вид деятельности научных учреждений всего мира на основе анализа статей в 82 ведущих научных журналах, связанных с *Nature*. Место в рейтинге определялось по значению показателя *Fractional count* – это дробь, в числителе которой число авторов от данной организации, а в знаменателе – общее число организаций, авторы которых участвовали в написании этой статьи.

Университет ИТМО в этом году на шестом месте в России (в прошлом мы были четвертыми). До нас – РАН (!), МГУ, НИЦ «Курчатовский институт», СПбГУ, МФТИ, после нас – Объединенный институт ядерных исследований, Сколтех, Казанский федеральный университет, МИСиС. Значение показателя *Article count* (число статей в указанных журналах) у нас 75 (53 статьи по физическим наукам, 18 – по химии, четыре – по наукам о жизни).

Еще в этом рейтинге упоминается Альметрика – *Altmetric Score (AS)*, которая учитывает упоминание статьи на разных сайтах, сетях и т. д. Каждое цитирование берется с определенным коэффициентом. Статья с участием Леши Сергушичева в *Nature* имеет значение показателя *AS* равное 170. У следующей статьи, связанной с нашим университетом, значение этого показателя – 93 (<https://www.natureindex.com/institution-outputs/russia/itmo-national-research-university-itmo-university/513906d134d6b65e6a000eba>).

Статьи о старении и молекулярном дизайне

1. Mamoshina P., Kochetov K., Putin E., Cortese F., Aliper A., Lee W., Uhn L., Skjodt N., Kovalchuk O., Scheibye-Knudsen M., Zhavoronkov A. Population Specific Biomarkers of Human Aging: A Big Data Using South Korean, Canadian and Eastern European Patient Population // **The Journals of Gerontology. Series A**. 2018. Vol.73. Issue 11, pp. 1482-1490. IF: 5, 95. SJR: 2.79 (<https://academic.oup.com/biomedgerontology/article/73/11/1482/4801287/>). Разработан алгоритм определения возраста по результатам анализа крови с учетом пола и национальности. <https://news.ifmo.ru/ru/science/it/news/7260/>.
2. Putin E., Asadulaev A., Ivanenkov Y., Aladinskiy V., Sanchez-Lengeling B., Aspuru-Guzik A., Zhavoronkov A. Reinforced Adversarial Neural Computer for De Novo Molecular Design // **Journal of Chemical Information and Modeling**. 2018. Vol.58. Issue 6, pp.1194-1204. IF: 3.76 (<https://pubmed.ncbi.nlm.nih.gov/29762023/>). Пятый и шестой авторы из Гарварда, программный продукт которых наши ребята улучшили.
3. Putin E., Asadulaev A., Vanhaelen Q., Ivanenkov Y., Anastasia V., Aladinskaya A., Aliper A., Zhavoronkov A. Adversarial Threshold Neural Computer for Molecular De Novo Design // **Molecular Pharmaceutics**. 2018. Vol.15. Issue 10, pp. 4386-4397. IF: 4.44 (<https://pubs.acs.org/doi/10.1021/acs.molpharmaceut.7b01137>).

Магистратура по специальности «Биоинформатика и системная биология».

20 апреля появилось сообщение о том, что в нашем университете на базе международной научной лаборатории «Компьютерные технологии» в сотрудничестве с Институтом биоинформатики под руководством Алексея Сергушичева планируется открытие в 2018 г. англоязычной магистерской программы «Биоинформатика и системная биология» (<http://news.ifmo.ru/ru/education/official/news/7466/>). Вот мнение магистранта из Индии об этой программе приведено здесь: <http://news.ifmo.ru/ru/education/students/news/8494/>.

Появление этой программы стало возможным благодаря опыту, накопленному в ходе проведения семинаров и школ в этой области. Программа рассчитана на то, чтобы обучить студентов с базовым биологическим образованием методам биоинформатики и дать представление о связи между вычислительной и экспериментальной биологией.

При этом будет осуществляться подготовка специалистов, способных применять методы биоинформатики для решения задач, возникающих в различных областях биологии и медицины. В

программе рассматриваются: 1. Эволюционные процессы: происхождение видов, появление и распространение антибиотикорезистентности; 2. Диагностика и изучение молекулярных механизмов генетических заболеваний; 3. Структурное моделирование для задач фармакологии; 4. Персонализированный подбор терапии для пациентов с раковыми и другими заболеваниями.

Первый набор будет в 2018-2019 учебном году. На программе 13 мест, из которых восемь – бюджетные. Первый выпуск – 2020 г.

Геннадий Короткевич остался в Университете ИТМО

27 июня в День молодежи состоялась ежегодная Церемония чествования лучших выпускников вузов Санкт-Петербурга. Университет ИТМО представлял магистр кафедры «Компьютерные технологии» Университета ИТМО Геннадий Короткевич – самый титулованный спортивный программист в мире (http://news.ifmo.ru/ru/university_live/achievements/news/7642/). «Поступление в Университет ИТМО было осознанным. Годы, проведенные на кафедре «Компьютерные технологии», принесли мне много успехов. Моя дальнейшая деятельность также будет связана с Альма матер – я принял решение поступать в аспирантуру», – сказал Геннадий. И, естественно, поступил. Вот так!

Как отмечено выше, магистерская работа Геннадия, выполненная под руководством Леши Сергушичева, была посвящена биоинформатике. При поступлении в аспирантуру руководителем Гена выбрал меня. Я, естественно, согласился, но думаю, что это временно, так как наметил ему значительно более «крутого» руководителя – Марка Дейли, с которым он 17 августа в присутствии Леши встретился в Хельсинки.

В приложении к газете «Коммерсант» (Наука, Приложение № 39 от 27.09.2018, с. 35) была опубликована статья о Геннадии Короткевиче с характерным названием, предложенным мною: «Он уже вошел в историю человечества» (<https://www.kommersant.ru/doc/3751359>). А еще я горжусь приведенными в статье словами Гены, сказанными им на моем юбилее (<https://cloud.mail.ru/public/JMhv/NwabsCoyb/%D0%A2%D0%BE%D1%81%D1%82%2015.mp4>): «Я сижу за столом чемпионов, что очень для меня ценно. Важно и то, что я могу находиться в своей компании. Это всегда было для меня смыслом жизни и тем, что меня очень вдохновляло». Кроме того, горжусь последним абзацем этой статьи, предложенным мной: «Гена молод, гениален и красив, он уже вошел в историю человечества и сам способен решить, нужно ли ему входить в нее еще раз».

Часть 4. 2019, 2020 гг.

<https://vk.com/@1077823-chast-4-kak-bioinformatika-i-sistemnaya-biologiya-poyavilis>

2019 год

Статьи в выдающихся журналах

1. Zhavoronkov A., Ivanenkov Y., Aliper A., Veselov M., Aladinskiy V., Aladinskaya A., Terentiev V., Polykovskiy D., Kuznetsov M., Asadulaev A., Volkov Y., Zholus A., Shayakhmetov R., Zhebrak A., Minaeva L., Zagribelnyy B., Lee L., Soll R., Madge D., Xing L., Guo T., Aspuru-Guzik A. Deep Learning enables Rapid Identification of Potent DDR 1 Kinase Inhibitors // *Nature Biotechnology*. 2019. 37, pp.1038-1040. 2-year IF: 31.864, 5-year IF: 45.117, SJR 2018: 14.57 (<https://www.nature.com/articles/s41587-019-0224-x>).
2. Caine E., Scheaffer S., Arora N., Zaitsev K., Artyomov M., Coyne C., Moley K., Michael S. Diamond M. Interferon Lambda Protects the Female Reproductive Tract Against Zika Virus Infection // *Nature Communications*. 2019. Vol. 10. Article number: 280. IF: 12.134, SJR: 6.585 (<https://www.nature.com/articles/s41467-018-07993-2>).
3. Pelgrom L., Patente T., Sergushichev A., Esaulova E., Otto F., Ozir-Fazalalikhani A., van der Zande HJP, van der Ham A., van der Stel S., Artyomov M., Everts B. LKB1 Expressed in Dendritic Cells Governs the Development and Expansion of Thymus-Derived Regulatory T Cells // *Cell Research*. 2019. 29(5): 406-419. Doi: 10.1038/s41422-019-0161-8. 5-year IF: 18.448, SJR: 6.01 (<https://www.nature.com/articles/s41422-019-0161-8>).
4. Cella M., Gamini R., Sécca C., Collins P., Zhao S., Peng V., Robinette M., Schettini J., Zaitsev K., Gordon W., Bando J., Yomogida K., Cortez V., Fronick C., Fulton R., Lin L., Gilfillan S., Flavell R.,

- Shan L., **Artyomov M.**, Bowman M., Oltz E., Jelinsky S., Colonna M. Subsets of ILC3–ILC1-Like Cells Generate a Diversity Spectrum of Innate Lymphoid Cells in Human Mucosal Tissues // *Nature Immunology*. 2019. June. 5-year IF: 21.974, SJR: 13.3 (<https://www.nature.com/articles/s41590-019-0425-y>).
5. Bajpai G., Bredemeyer A., Li W., **Zaitsev K.**, Koenig A., Lokshina I., Mohan J., Ivey B., Hsiao H., Weinheimer C., Kovacs A., Epelman S., Artyomov M., Kreisel D., Lavine K. Tissue Resident CCR2 – and CCR2+ Cardiac Macrophages Differentially Orchestrate Monocyte Recruitment and Fate Specification Following Myocardial Injury // *Circulation research*. 2019. 124 (2), 263-278. IF: 15.86, SJR: 5.72 (<https://www.ncbi.nlm.nih.gov/pubmed/30582448>).
6. Jordan S., Tung N., Casanova-Acebes M., Chang C., Cantoni C., Zhang D., Wirtz T., Naik S., Rose S., Brocker C., **Gainullina A.**, Maier B., LeRoith D., Gonzalez F., Meissner F., Ochando J., Rahman A., Chipuk J., Artyomov M., Frenette P., Piccio L., Berres M., Gallagher E., Merad M. Dietary Intake Regulates the Circulating Inflammatory Monocyte Pool // *Cell*. 2019. 178, pp.1102-1114. IF: 36.216, SJR: 25.976 (<https://www.sciencedirect.com/science/article/pii/S0092867419308505>).
7. Zhang S., Weinberg S., Deberge M., **Gainullina A.**, Schipma M., Kinchen J., Ben-Sahra I., Gius D., Yvan-Charvet L., Chandel N., Schumacker P., Thorp E. Efferocytosis Fuels Requirements of Fatty Acid Oxidation and the Electron Transport Chain to Polarize Macrophages for Tissue Repair // *Cell Metabolism*. 2019. Vol. 29. No 2, pp. 443-456. IF: 22.415, SJR: 10.69 (<https://www.ncbi.nlm.nih.gov/pubmed/30595481>).
8. **Zaitsev K.**, Bambouskova M., Swain A., Artyomov M. Complete Deconvolution of Cellular Mixtures Based on Linearity of Transcriptional Signatures // *Nature Communications*. 2019. Vol. 10. Article number: 2209. IF: 12.134, SJR: 6.585 (<https://www.nature.com/articles/s41467-019-09990-5>).
9. Dmitrieva R., ... **Sergushichev A.**, Sitnikova M., Kostareva A. Skeletal Muscle Resident Progenitor Cells Coexpress Mesenchymal and Myogenic Markers and Are Not Affected by Chronic Heart Failure-Induced Dysregulations // *Stem Cells International*. 2019. Open Access. 11 p. (<https://doi.org/10.1155/2019/5690345>).
10. Park E., ..., **Zaitsev K.**, ... Yokoyama W. Toxoplasma Gondii Infection Drives Conversion of NK Cells into ILC1-Like Cells // *eLIFE*. 2019. 8. e47605. SJR: 6.62 (<https://pubmed.ncbi.nlm.nih.gov/31393266/>).
11. Nikitin D., ... **Mursina A.**, Pats K., ... Buzdin A. Retroelement-Linked H3K4me1 Histone Tags Uncover Regulatory Evolution Trends of Gene Enhancers and Feature Quickly Evolving Molecular Processes in Human Physiology // *Cells*. 2019. 8 (10). 1219. IF: 5.656, SJR: 2.742 (<https://doi.org/10.3390/cells8101219>).
12. Lelyavina T., **Sergushichev A.**, ... Dmitrieva R. Clinical Response to Personalized Exercise Therapy in Heart Failure Patients with Reduced Ejection Fraction Is Accompanied by Skeletal Muscle Histological Alterations // *International Journal of Molecular Sciences*. 2019. Vol. 20. Issue 21, 5514. SJR: 1.31 (<https://doi.org/10.3390/ijms20215514>).
13. Kiselev A., Vaz R., **Sergushichev A.**, Dmitrieva R., ... Kostareva A. Truncating variant in MYOF gene is associated with limb-girdle type muscular dystrophy and cardiomyopathy // *Frontiers in Genetics*. 2019. 10: 608 (<https://doi.org/10.3389/fgene.2019.00608>).
14. Glotov A.S., **Kazakov S.V.**, Vashukova E.S., Pakin V.S., Danilova M.M., Nasykhova Y.A., Masharsky A.E., Mozgovaya E.V., Ereemeeva D.R., Zainullina M.S., Baranov V.S. Targeted Sequencing Analysis of ACVR2A Gene Identifies Novel Risk Variants Associated with Preeclampsia // *Journal of Maternal-Fetal and Neonatal Medicine*. Vol. 32. 2019. Issue 17. IF: 1.826 (<https://doi.org/10.1080/14767058.2018.1449204>).
15. O`Braen S. et al. Genome-wide sequence analyses of ethnic population across Russia // *Genomics*. 2019. V.112. Issue 1, pp. 442-458. 5-Year IF: 3.270. SJR: 1.688. (<https://www.sciencedirect.com/science/article/pii/S0888754318307419?via%3Dihub>). Среди авторов с аффилиацией ИТМО появился выдающийся ученый Стефан О`Брайен: **Google Scholar. h-index - 154!**
16. **Korotkevich G.**, Sukhov V., **Sergushichev A.** FGSEA. Fast Gene Set Enrichment Analysis / Bioconductor. Open-Source Software for Bioinformatics. 2019 (<http://bioconductor.org/packages/devel/bioc/html/fgsea.html>). Среди авторов этой статьи Геннадий Короткевич! Лиха беда начало!
- А еще здесь стоит обратить внимание на выдающийся старт **Насты Гайнуллиной!**

Конференции

1. Avdeyev P., **Atamanova M.**, Alekseyev M. A Uniform Theory of Adequate Subgraphs for the Genome Median, Halving, and Aliquoting Problems / 6th International Conference on Algorithms for Computational Biology (AlCoB 2019). Lecture Notes in Bioinformatics (LNCS 11488). Algorithms for Computational Biology. 2019, pp. 97-111. **SJR: 0.295** (<https://www.springer.com/gp/book/9783030181734>).
2. **Gainullina A., Sergushichev A., Randolph G., ImmGen Consortium, Artyomov M.** ImmGen Open-Source Mononuclear Phagocyte Profiling Analysis / Proceedings of 9th Moscow Conference on Computational Molecular Biology (MCCMB'19). М.: ИППИ РАН, 2019 (<http://mccmb.belozersky.msu.ru/2019/thesis/MCCMB2019/abstracts/105.pdf>).

Bioinformatics Contest 2019

Второго февраля начался предварительный тур **третьей** международной олимпиады по биоинформатике – *Bioinformatics Contest 2019* (<https://bioinf.me/en/contest>). Задачи были составлены сотрудниками нашей кафедры: Алексей Сергушичев, Виталий Демьянюк, Григорий Шовкопляс, Артем Васильев, Виталий Аксенов, Илья Збань, Никита Алексеев, Мария Атаманова, Дмитрий Якутов. Еще один автор задач – **Сергей Аганезов**, закончивший наш университет. Он сейчас работает в университете Джона Хопкинса (США). Спонсоры: *JetBrains*, *Yandex* и *Serokell* (это компания нашего недавнего выпускника Арсения Сероки).

23, 24 февраля наши ребята провели *Final Round Bioinformatics Contest 2019* (<http://mon.stepik.org/>), для участия в котором было допущено более 400 участников. Из них 247 участников набрали очки. **С большим отрывом победил Геннадий Короткевич**, который набрал 6310 баллов из возможных 6650 баллов. Вот как прокомментировал в «ВКонтакте» победу Гены организатор этих соревнований Леша Сергушичев: «Идет Гена по лесу, видит проходит констест. Сел за него и ... победил». Второе место в этом констесте занял **Николай Будин**.

24 марта появилось **большое интервью** (<https://www.listennotes.com/podcasts/the-bioinformatics/31-bioinformatics-contest-fUVycsnwYBa/>) в один час 46 минут на the Bioinformatics chat #31: *Bioinformatics Contest 2019 with Alexey Sergushichev (one of the organizers of the contest) and Gennady Korotkevich* (the 1st prize winner), в котором они разобрали решение трех задач квалификационного раунда (*Bee Population, Sequencing Errors, Transposable Elements*) и пяти задач финального раунда (*Cancer and Chromosome Rearrangements, Epigenomic Marks, Bacterial Communities, Minimal Genome, Endangered Species*). В интервью прозвучали фамилии авторов задач из Университета ИТМО: Виталия Аксенова, Григория Шовкопляса, Дмитрия Якутова, Ильи Збаня и Артема Васильева.

Открытая лекция профессора Марка Дейли в Университете ИТМО

Леша Сергушичев с Максимом Артемовым и другими учеными в мае 2019 г. провели **шестой** семинар по системной биологии под Санкт-Петербургом (<https://bioinf.me/education/sbw/2019>). **В нем принял участие Марк Дейли** – выдающийся ученый, один из пионеров вычислительной генетики, с которым наши ребята активно сотрудничают. В настоящее время он директор Института молекулярной медицины Финляндии (*Institute for Molecular Medicine Finland*) в Хельсинки. На раннем этапе проекта секвенирования генома человека Марк предложил генетическую модель построения карт гаплотипов – *HarMap*, которые позволили провести первый полногеномный поиск ассоциаций (GWAS). В его лаборатории были найдены одни из первых генетических ассоциаций для болезни Крона, аутизма и шизофрении.

Марк Дейли ([https://en.wikipedia.org/wiki/Mark_Daly_\(scientist\)](https://en.wikipedia.org/wiki/Mark_Daly_(scientist))) – один из наиболее цитируемых ученых в области генетики человека и входит в **Топ 100 самых цитируемых ученых всех времен и народов по оценкам Google Scholar**. На 14.09.2020 г. его индекс Хирша был равен 202 при числе цитирований – 293 854 (<https://scholar.google.com/citations?user=SW3aNagAAAAJ&hl=en>). На 27.08.2020 г. в рейтинге *Highly Cited Researchers (h>100) According to Their Google Scholar Citations Public Profiles* (<https://www.webometrics.info/en/hlargerthan100>) Марк занимал 85 место с индексом Хирша – 196 и числом цитирований – 280 584. В этом рейтинге на эту дату индекс Хирша по *Google Scholar*, равный 100 и более, имели 4167 ученых.

25 мая Марк провел открытую лекцию в Университете ИТМО на тему *The big data and computational opportunities in the electronic health records and genome data*. До лекции Максим

представил меня Марку. Я подарил Дейли буклет нашего факультета. В моем экземпляре буклета (на странице с портретами Артемова и Сергушичева) он написал: «**Congratulations on ITMO great success!** Hoping for yeast future collaboration. Mark Daly» (https://vk.com/albums1077823?z=photo1077823_456241916%2Fphotos1077823). Интересно, что, когда я сказал ему, что у меня плохой английский, Марк заметил, что его русский не лучше! Хорошо, когда общаются люди без комплексов... Кстати, Билл Гейтс не знает иностранных языков, и ничего, как-то ;-) обходится...

После встречи мы сфотографировались: **В. Ульянов, А. Сергушичев, М. Дейли, М. Артемов и я.** Съемка проходила на фоне фотографии Г. Короткевича и Н. Нигматуллина. Марк обратил на это внимание и попросил включить Гену в кадр, так как знаком с ним. Потом оказалось, что в кадр еще попала фотография А. Станкевича, А. Васильева и Б. Минаева (https://vk.com/albums1077823?z=photo1077823_456242082%2Fphotos1077823).

Костя Зайцев – победитель программы *Skoltech Fellowship Program 2019*

28 мая (в мой день рождения) в Сколково прошла торжественная церемония, посвященная победителям программы *Skoltech Fellowship Program 2019*. (<https://www.skoltech.ru/en/2019/05/systems-biology-program-2019-award-ceremony-and-seminar/#>). Наш выпускник Костя Зайцев (<https://news.itmo.ru/ru/science/it/news/8536/>) вошел в пятерку лучших молодых ученых страны по системной биологии! (<https://www.pm.skoltech.ru/>). Тема его проекта: *Identification and separation of sources of transcriptional variability in single-cell RNA-seq data* (<http://news.ifmo.ru/ru/science/it/news/8536/>). А вот какой у Кости был старт! Обратите внимание на названия журналов, где опубликованы указанные ниже статьи:

1. Caine E., Scheaffer S., Arora N., **Zaitsev K.**, Artyomov M., Coyne C., Moley K., Michael S. Diamond M. Interferon lambda protects the female reproductive tract against Zika virus infection // **Nature Communications**. 2019. Vol. 10. (Article number: 280). IF: 12.134, SJR: 6.585 (<https://www.nature.com/articles/s41467-018-07993-2>).

2. **Zaitsev K.**, Bambouskova M., Swain A., Artyomov M. Complete deconvolution of cellular mixtures based on linearity of transcriptional signatures // **Nature Communications**. 2019. Vol. 10. (Article number: 2209). IF: 12.134, SJR: 6.585 (<https://www.nature.com/articles/s41467-019-09990-5>). В статье изложен новый метод определения типов клеток в образце ткани. Определены типичные признаки активности генов разных типов клеток и на этой основе построена модель для определения разных клеток в смешанных образцах. Такой подход позволяет, например, понять как соотношение разных клеточных типов связано с выживаемостью онкологических больных (<http://news.ifmo.ru/ru/science/it/news/8526/>).

3. Cella M., Gamini R., Sécca C., Collins P., Zhao S., Peng V., Robinette M., Schettini J., **Zaitsev K.**, Gordon W., Bando J., Yomogida K., Cortez V., Fronick C., Fulton R., Lin L., Gilfillan S., Flavell R., Shan L., Artyomov M., Bowman M., Oltz E., Jelinsky S., Colonna M. Subsets of ILC3–ILC1-like cells generate a diversity spectrum of innate lymphoid cells in human mucosal tissues // **Nature Immunology**. 2019. June 17. 5-year IF: 21.974, SJR: 13.3 (<https://www.nature.com/articles/s41590-019-0425-y>).

4. Bajpai G., Bredemeyer A., Li W., **Zaitsev K.**, Koenig A., Lokshina I., Mohan J., Ivey B., Hsiao H., Weinheimer C., Kovacs A., Epelman S., Artyomov M., Kreisel D., Lavine K. Tissue resident CCR2 – and CCR2+ cardiac macrophages differentially orchestrate monocyte recruitment and fate specification following myocardial injury // **Circulation research**. 2019. 124 (2), 263-278. IF: 15.86, SJR: 5.72 (<https://www.ncbi.nlm.nih.gov/pubmed/30582448>).

Напомню, что 2017 г. одним из победителей программы *Skoltech Fellowship Program* был Леша Сергушичев!

Когда показываешь ссылки на эти публикации ИТ-талантливым молодым людям, некоторые из них говорят, что биология их не интересует. При этом они даже не удосуживаются прочесть фрагмент из интервью Гены Короткевича (http://news.ifmo.ru/ru/university_live/achievements/news/8550/), приведенный выше, в котором он говорит, что Леша Сергушичев «очистил» ему задачу от биологии, и она стала чисто математическо-программистской.

Один старый художник когда-то говорил мне, что нет разницы между Озерками и Таити – и там, и там жизнь людей, и поэтому для создания выдающихся произведений не надо, как это сделал Гоген, ехать на край света. И нашим молодым людям нет необходимости для занятий наукой ехать на «Таити», а достаточно подняться на второй этаж кафедры, где сидят Леша и Костя, и заняться с ними наукой. Они знают, как публиковаться в журналах с такой цитируемостью, которая не снилась не только мне, но и почти всем руководителям, к которым молодые люди собрались за рубеж!

В это время один мой молодой коллега высказал сомнение в моем вкладе в развитие биоинформатики на нашей кафедре – он все заслуги в этом вопросе отнес на счет Алексея Сергушичева (<https://scholar.google.ru/citations?user=fcH0gPgAAAAAJ&hl=en>). Я транслировал это высказывание Леше, который прокомментировал его так: **«Вы называете меня «великим русским ученым», так вот «великим ученым» я, возможно, стал бы и без Вас, но русским и по этой тематике – только благодаря Вам».** Неплохо сказано, не правда ли? Особенно если учесть, что Леша родом из Вологды...

Семинар по секвенированию единичных клеток

В начале июня в Томске прошла первая в России международная научная школа, посвященная новому исследовательскому подходу – секвенированию единичных клеток. Этот способ открывает большие возможности в изучении особенностей развития рака и причин устойчивости опухолей к терапии. В качестве лекторов школы выступили Павел Морозов из Университета Рокфеллера (Нью-Йорк, США), Константин Оконечников из Центра исследования рака (Гейдельберг, Германия), **Алексей Сергушичев и Константин Зайцев из Университета ИТМО** (Санкт-Петербург, Россия), Софья Касацкая из Института биорганической химии РАН (Москва, Россия), ученые из Томска. В Томске имеется секвенатор *NextSeq500*, он позволяет анализировать не только геном, экзом, эпигеном, транскриптом, но и отдельные гены клеток различных живых организмов (<http://www.tsu.ru/news/v-tomske-prokhorit-pervaya-v-rf-shkola-sekvenirova/>).

Участие в работе летней школы «Биоинформатика в исследованиях развития и старения»

С 29 июля по третье августа в Москве проходила летняя школа «Биоинформатика в исследованиях развития и старения» (<https://vk.com/@itmo.ctlab-bioinformatika-v-issledovaniyah-razvitiya-i-stareniya>), в рамках которой Алексей Сергушичев прочел лекцию «Анализ данных *RNA-seq*» (https://www.youtube.com/watch?time_continue=4&v=qjVykn0i5s8). В работе школы принимал участие Костя Зайцев.

Как мы «сохранили» Костю Зайцева в Университете ИТМО

А теперь я расскажу, как можно быть в тренде мировой науки, никуда не уезжая из родного российского университета. В 2016 г. компания «Филип Моррис» выделила деньги на пять грантов на три года каждый на научные исследования, что было названо *Systems Biology Skoltech Fellowship Program* (<https://www.pm.skoltech.ru/>). При этом победители могли оставаться там, где работают: здесь нет и не было принудительной мобильности – надо только отчитываться один раз в год.

Получить этот грант очень трудно, так как в «приемную комиссию» входят, как отмечено выше, Михаил Гельфанд и Константин Северинов. Так вот, в 2017 г. нам надо было «удержать» на кафедре Лешу Сергушичева, который сейчас работает у нас и публикуется с ведущими в этой области учеными мира. Победа в Сколково (<http://www.skoltech.ru/en/2017/04/the-winners-of-the-systems-biology-fellowship-program-2/>) упростила решение этой задачи.

А в 2019 г. произошло следующее. Наш выпускник Костя Зайцев проводил исследования в Университете Вашингтона в Сент-Луисе, но потом поступил неосторожно – выехал в Россию. Когда он захотел вернуться назад, с визой возникли большие проблемы. Прошло достаточно много времени, и Леша сказал мне, что Костя собирается уехать в Германию... Я спросил Алексея знает ли он кого-то из ученых в этой области в Германии, и услышал ответ: «Нет». После этого я предположил, что Костя может попасть к профессору, которому абсолютно наплевать на Лешу и на Университет ИТМО. Леша сказал, что это вполне вероятно, и услышал от меня простой вопрос: «А какой тогда нам прок от этого?» После этого Леша сделал ничем не обоснованное предположение, что Косте может быть там хорошо.

Потом я в очередной раз пояснил Леше, что несмотря на то, что очень уважаю труд водителей трамваев, я таковым не являюсь, и цель моей жизни не состоит в том, чтобы бессмысленно для

страны, университета, кафедры и себя способствовать перевозке молодых талантливых людей на большие расстояния. Выслушав эту тираду, Леша спросил меня: «А что же делать?». Мой ответ был на удивление прост: «**Думать!**».

И они придумали: Костя выиграл грант в Сколково, и 28 мая он был вызван туда на награждение, о чем написала газета «Коммерсант» <https://www.kommersant.ru/doc/4066991>). И так, Костя может три года ни от кого не зависеть, оставаться в нашем коллективе с очень сильными сотрудниками и студентами, проводить исследования, с теми людьми, с кем работал раньше, а совместно с Лешой Сергушичевым и их другом и коллегой Максимом Артемовым организовывать школы и семинары по биоинформатике и системной биологии в России и других странах.

А еще Леша и он общаются с Марком Дейли, о величии которого сказано выше. У Марка в лаборатории *Broad Institute of MIT and Harvard* работают брат Максима – Никита и стажирется наш аспирант Саша Лобода.

Так в рамках предложенной мною инициативы «Сохраним в университетах лучших!» (<https://www.itweek.ru/management/article/detail.php?ID=108777>) нам удалось «сохранить» и Костю! По этому вопросу пока все.

Nature Index Russia

31 июля стал известен *Nature Index Russia* за период 1.08.2018-31.07.2019 г. (<https://www.natureindex.com/country-outputs/russia>), характеризующий публикационную активность научных организаций и вузов в изданиях, связанных с журналом *Nature*. На этот раз Университет ИТМО занял седьмое (!) место в России после РАН (в целом!), МГУ, СПбГУ, Курчатовского центра, МФТИ и Объединенного института ядерных исследований. После нас – Новосибирский государственный университет и Сколтех.

Лекция о биоинформатике и изучении происхождения людей

27 августа Владимир Ульянов прочел лекцию «Биоинформатика: как геномные данные помогают изучать происхождение людей и механизмы развития заболеваний» (https://vk.com/itmo.ctlab?z=video-28625572_456239690%2Fpl_-76507013_-2). Лекция прошла в рамках лектория Университета ИТМО в Новой Голландии (<http://news.ifmo.ru/ru/announce/53436/>).

Леша Сергушичев на «Эхо Москвы в Санкт-Петербурге»

24 сентября Алексей участвовал на «Эхо Москвы в Санкт-Петербурге» в программе «Пифагоровы штаны» на тему «Компьютер помогает биологам решать различные задачи и открывать универсальные механизмы работы тех или иных генов. Как это работает для клеток иммунной системы?».

Лекции в Computer Science Center

22 ноября Леша Сергушичев прочел в *Computer Science Center* (<https://compscicenter.ru/>) лекцию на тему «Соединяем точки: использование графов для интерпретации биологических данных» (https://www.youtube.com/watch?time_continue=5&v=JY6yxRv0vHI&feature=emb_logo).

За четыре дня до этого там же Екатерина Носкова, работающая у нас в международной научной лаборатории «Компьютерные технологии», прочла лекцию на тему «Вывод демографических историй из геномных данных» (https://www.youtube.com/watch?time_continue=189&v=S_cEJ5pL9IU&feature=emb_logo).

2020 год

С первого января при нашей «кафедре» в рамках международного научного центра «Компьютерные технологии» действует научно-образовательный центр геномного разнообразия под руководством Алексея Сергушичева. Сотрудники этого центра перешли к нам из СПбГУ (<http://dobzhanskycenter.spbu.ru/ru/59-kategoriya-ru-ru/sotrudniki/rukovodstvo/62-o-brajen-stefan-dzh-r.html>).

Научный руководитель центра – выдающийся ученый Стивен О'Брайен (https://en.wikipedia.org/wiki/Stephen_J._O'Brien). На 27.08.2020 г. в *Google Scholar* Stephen J. O'Brien имел индекс Хирша – 152, а число цитирований – 98 527 (<https://scholar.google.com/citations?user=XwnFYYYAAAAJ&hl=en>). В этот же день он в рейтинге *Highly Cited Researchers (h>100) According to Their Google Scholar Citations Public Profiles*

(<https://www.webometrics.info/en/hlargerthan100>) занимал 445 место со следующими показателями: индекс Хирша – 155, а число цитирований – 103 479. Его фотографии с Сергушичевым, Ульяновцевым и мною размещены здесь: https://vk.com/id1077823?z=photo1077823_457243976%2Fphotos1077823 и https://vk.com/id1077823?z=photo1077823_457243975%2Fphotos1077823.

Новая структура должна стать центром биоинформатики, в котором будут вестись работы по сборке геномов различных живых организмов, а также изучению генетического разнообразия животных и людей (<https://indicator.ru/mathematics/bioinformatika-eto-novyi-mikroskop.htm>).

Пятого марта на портале Университета ИТМО была опубликована статья об этом: «В Университете ИТМО создан научно-образовательный центр геномного разнообразия: в нем готовятся изучать геном балтийской нерпы, амурского тигра и других животных» (<https://news.itmo.ru/ru/science/it/news/9233/>).

За день до ее публикации там же вышла статья «Ученые Университета ИТМО создали алгоритм, позволяющий легче и точнее строить историю популяций людей и животных» (<https://news.itmo.ru/ru/science/it/news/9222/>). В ней описан программный продукт, реализующий этот алгоритм, который позволяет быстро и эффективно анализировать геномные данные, и на их основе строить наиболее вероятные модели демографической истории тех или иных популяций растений, животных и людей. Эта статья посвящена научной работе: *Noskova G., Ulyantsev V., Koepli K-P., O'Brien S., Dobrynin P.* GADMA: Genetic Algorithm for Inferring Demographic History of Multiple Populations from Allele Frequency Spectrum Data // *GigaScience*. Vol. 9. 2020. Issue 3. IF: 4.688. SJR: 4.73. <https://doi.org/10.1093/gigascience/giaa005>. При этом отмечу, что из пяти авторов этой работы, четверо сотрудники вновь образованного центра: Екатерина Носкова, Владимир Ульяновцев, Павел Добрынин и Стивен О'Брайен.

Леша Сергушичев рассказывает о зарождении у нас биоинформатики

26 марта исполнилось 120 лет Университету ИТМО. В этот день на портале университета появилась статья «Как в университет ИТМО зарождались лазерные технологии, а когда начала развиваться биоинформатика – ученые вспоминают историю вуза» (https://news.itmo.ru/ru/university_live/achievements/news/9286/). Приведу рассказ Алексея Сергушичева, директора научно-образовательного центра геномного разнообразия, руководителя группы биоинформатики международного научного центра «Компьютерные технологии»: «Первые исследования в области биоинформатики начались с конца 2009-начала 2010 г. В конце ноября 2009 г. в Университет ИТМО по инициативе Анатолия Абрамовича Шалыто приезжал Егор Борисович Прохорчук из команды академика РАН Константина Георгиевича Скрябина. На встрече со студентами и выпускниками кафедры «Компьютерные технологии» он рассказал про геном и связанные с этим задачи биоинформатики. В результате в качестве дипломных работ этой темой начали заниматься трое наших очень сильных студентов: призер чемпионата мира по программированию Михаил Дворкин (http://is.ifmo.ru/diploma-theses/dvorkin_genom.pdf) и два чемпиона мира – Владислав Исенбаев (http://is.ifmo.ru/genom/isenbaev_thesis.pdf) и Евгений Капун (http://is.ifmo.ru/genom/kapun_thesis.pdf). Анатолий Абрамович для их мотивации написал два стиха: 1. Будет геном – будет диплом! 2. Не будет генома – не будет диплома! Стихи «сработали».

Позже этой задачей под руководством чемпиона мира по программированию Федора Царева занялся я (<http://is.ifmo.ru/diploma-theses/2011-bachelor-sergushichev/>, <http://is.ifmo.ru/diploma-theses/2013/master/sergushichev/sergushichev.pdf>) и мои одноклассники – Сергей Казаков (<http://is.ifmo.ru/diploma-theses/2013/master/kazakov/kazakov.pdf>) и Антон Александров (<http://is.ifmo.ru/diploma-theses/2013/master/alexandrov/alexandrov.pdf>). К моей и Сергея магистерским «руку приложил» Анатолий Абрамович, который в дальнейшем был научным руководителем наших кандидатских диссертаций (в самом конце руководителем Алексея по взаимной договоренности стал Максим Артемов, А.Ш.).

Ключевым этапом в развитии этой области стал 2013 г. В конце этого года мне удалось съездить на стажировку в Университет Вашингтона в Сент-Луисе к тогда еще начинающему исследователю, а теперь профессору Максиму Артемову, с которым в 2012 г. в Санкт-Петербурге познакомился Федор Царев. Тогда же, мой одноклассник Владимир Ульяновцев, а сейчас директор центра

«Компьютерные технологии», съездил на стажировку в группу Дмитрия Алексеева в ФНКЦ ФХМ. В этом же году наш университет очень удачно вошел в «Программу 5-100», которая позволила нам оставаться в университете и заниматься этой перспективной областью. Чуть позже мы стали вести и образовательную деятельность в виде ежегодного семинара по системной биологии, который в 2018 г. перерос в магистерскую программу по биоинформатике и системной биологии.

В результате у нас сформировалось несколько направлений, по которым наши исследования выполняются на мировом уровне: анализ данных в иммунологии и раке, метагеномика, сравнительная геномика, по которым у нас выходят статьи как в ведущих профильных журналах по биоинформатике, так и в ведущих журналах мира, включая *Nature*, *Science* и *Cell*, а нашими программными инструментами пользуются тысячи ученых по всему миру. Сейчас одна из главных задач – организационная: за последнее время наша команда сильно выросла, что требует адаптации форматов работы.

С точки зрения науки я бы выделил несколько направлений. Во-первых, это наши исследования по методам интерпретации результатов полногеномного анализа ассоциаций, выполняемых совместно с лабораторией выдающегося ученого Марка Дейли в *Broad Institute of MIT and Harvard*. Хотя это и довольно давно существующая задача, сейчас появляется шанс сделать прорыв в этой области за счет использования больших объемов публичных данных и их анализа с помощью эффективных вычислительных методов. Во-вторых, это систематический анализ открытых баз данных экспрессии генов. У нас уже есть отдельные инструменты по поиску, визуализации и анализу таких данных, но теперь все эти инструменты необходимо объединить в единую платформу, чтобы дать возможность биологам со всего мира простой и полезный доступ к имеющимся массивам данных. И в-третьих, это развитие недавно появившегося у нас направления по геномике животных и, в частности исчезающих видов, под руководством известного профессора Стивена О'Брайена и наращивание экспертизы в области геномики человека».

Говорят Сергушичев и Ульяновцев

Третьего апреля в статье «Биоинформатика – это новый микроскоп» на портале *Indicator* (<https://indicator.ru/mathematics/bioinformatika-eto-novyi-mikroskop.htm>) Леша Сергушичев и Володя Ульяновцев рассказали свои версии истории того, как на кафедре «Компьютерные технологии» Университета ИТМО начали заниматься биоинформатикой.

Вот, что говорит Алексей: «Это направление появилось на нашей кафедре около 10 лет назад. Началось все со знакомства с научной группой академика РАН Константина Скрябина (с самим Константином Георгиевичем, А.Ш.). Они занимались исследованием генома и захотели привлечь к этим исследованиям ребят, которые занимаются компьютерными науками (ничего этого они не хотели – это я предложил Скрябину взаимодействовать с нами, и он согласился, А.Ш.). Мы в то время были студентами. На кафедре было развито не так много научных направлений, многие уходили работать в программисты. И вот по инициативе профессора Анатолия Шалыто команда Скрябина предложила нам задачу по сборке генома из отдельных прочтений. Она звучала довольно понятно для программистов: фрагменты генома – строки, и их надо склеить с помощью алгоритмов. С этой задачи мы постепенно начали заниматься биоинформатикой. Решить ее, как сейчас нам видится, удалось не слишком хорошо, но у нас сформировалась команда биоинформатиков. **В 2012 г. мы через Федю Царева познакомились с Максимом Артемовым**, который в то время только-только стал профессором (его должность называлась *Assistant Professor*, А.Ш.) в Университете Вашингтона в Сент-Луисе. Я съездил к нему на стажировку на несколько месяцев, и мы начали работать вместе. С тех пор у нас вышло много совместных статей, в том числе в таких журналах, как *Nature* и *Cell*. Тогда же мы, через Володю Ульянцева, начали сотрудничать с лабораторией биоинформатики Дмитрия Алексеева в ФНКЦ ФХМ в Москве. С первых публикаций в топовых журналах всем стало понятно, что в биоинформатике университет может сильно проявить себя, и это направление получило всестороннюю поддержку».

Теперь вспоминает Володя: «Многие из нас начинали с олимпиадного движения. Первые курсы я делал задачи по программированию и постепенно перетек в науку – занимался исследованиями, связанными с автоматным программированием (<http://is.ifmo.ru/works/2008/Vestnik/53/01-automata-based-programming.pdf>). В это время ребята защищали свои магистерские по сборщику генома, а я на это смотрел немного со стороны. Отчасти мне было радостно, что я занимаюсь

Computer Science, который имеет шансы попасть в долгоиграющую историю, но в то же время и обидно, что прямо сейчас мои исследования мало кому нужны. Выбор был непростой, и в аспирантуре после стажировки в ФНКИ ФХМ я начал заниматься и тем, и другим».

Статьи в журналах

1. Zhou Y., Song W. Andhey P., ..., **Zaitsev K.**, ..., **Artyomov M.**, Colonna M. Human and Mouse Single-Nucleus Transcriptomics Reveal *TREM2*-Dependent and *TREM2*-Independent Cellular Responses in Alzheimer's Disease // *Nature Medicine*. Vol. 26. 2020, pp.131-142. **IF: 30.641. SJR: 17.01** (<https://doi.org/10.1038/s41591-019-0695-9>).
2. Williams¹ J., Winkels H., Durant C., **Zaitsev K.**, Ghosheh Y., Ley K. Single Cell RNA Sequencing in Atherosclerosis Research // *Circulation Research*. 2020. Vol.126 (9), pp. 1112-1126. **IF: 15.862. SJR: 5.72** (<https://www.ahajournals.org/doi/10.1161/CIRCRESAHA.119.315940>).
3. Wang Y., **Zaitsev K.**, Lu Q., Shan Li, Schaiff W., Kim K, Droit L., Wilen C., Desai C., Balce D., Orchard R., Orvedahl A., Park S., Kreamalmeyer D., Handley S., Pfeifer J., Baldrige M., Artyomov M., Stallings C., Virgin H. Select Autophagy Genes Maintain Quiescence of Tissue Resident Macrophages and Increase Susceptibility to *Listeria Monocytogenes* // *Nature Microbiology*. 2020. **IF: 14.3. SJR: 7.54** (<https://doi.org/10.1038/s41564-019-0633-0>). Эта статья посвящена тому, что отключение гена в миелоидных клетках практически полностью защищает мышей от заражения листерозом (https://news.itmo.ru/ru/science/life_science/news/9113/).
4. **Krashenninnikova K.**, Diekhans M., Armstrong J., Dievskii A., Paten B., **O'Brien S.** halSynteny: a Fast, Easy-to-Use Conserved Synteny Block Construction Method for Multiple Whole-Genome Alignments // *GigaScience*. Vol. 9. 2020. Issue 6. **IF: 6.95. SJR: 4.726** (<https://doi.org/10.1093/gigascience/giaa047>).
5. **Noskova G.**, **Ulyantsev V.**, Koepli K-P., **O'Brien S.**, **Dobrynin P.** GADMA: Genetic Algorithm for Inferring Demographic History of Multiple Populations from Allele Frequency Spectrum Data // *GigaScience*. Vol. 9. 2020. Issue 3. **IF: 4.688. SJR: 4.73** (<https://doi.org/10.1093/gigascience/giaa005>). Этому исследованию посвящена статья «Ученые Университета ИТМО создали алгоритм, позволяющий легче и точнее строить историю популяций людей и животных». Соответствующий этому алгоритму программный продукт позволяет быстро и эффективно анализировать геномные данные, и на их основе строить наиболее вероятные модели демографической истории тех или иных популяций растений, животных и людей (<https://news.itmo.ru/ru/science/it/news/9222/>).
6. Avdeyev P., **Alexeev N.**, Rong Y., Alekseyev M. A Unified IPL Framework for Core Ancestral Genome Reconstruction Problems // *Bioinformatics*. 2020. 36 (10), pp. 2993-3003. **IF: 4.531. SJR: 4.55** (<https://pubmed.ncbi.nlm.nih.gov/32058559/>). Этому исследованию посвящена статья «Ученые создали алгоритм для изучения эволюционной истории видов с полногеномными дупликациями» (<https://news.itmo.ru/ru/science/it/news/9203/>https://news.itmo.ru/ru/science/it/news/9203/?fbclid=IwAR0Zh2m3pUChAU4916bpw_nw4b9CX11zBLc6INJyKBzbEfrTx2bwlxjEAI).
7. Esaulova E., Cantoni C., Shchukina I., **Zaitsev K.**, Bucelli R., Wu G., Artyomov M., Cross A., Edelson B. Single-Cell RNA-seq Analysis of Human CSF Microglia and Myeloid Cells in Neuron inflammation // *Neurology: Neuroimmunology and Neuroinflammation*. 2020, Vol. 7, No 4, pp. 1-12. **IF: 7.353. SJR: 2.1** (<https://nn.neurology.org/content/7/4/e732>).
8. Röhl S., Rykaczewska U., Seime T., Suur B., Diez M., Gådin J., **Gainullina A.**, **Sergushichev A.** et al. Transcriptomic Profiling of Experimental Arterial Injury Reveals New Mechanisms and Temporal Dynamics in Vascular Healing Response // *JVS: Vascular Science*. 2020. V.1, pp.13-37. **SJR: 1.293** (<https://doi.org/10.1016/j.jvssci.2020.01.001>).
9. **Гайнуллина А.Н., Шалыто А.А., Сергушичев А.А.** Метод совместной кластеризации в графовом и корреляционном пространствах // Моделирование и анализ информационных систем. 2020. Т. 27 (2), с.180-193. (**Gainullina A.N., Shalyto A.A., Sergushichev A.A.** Method of the Joint Clustering in Network and Correlation Spaces. *Modeling and Analysis of Information Systems*. 2020. 27(2), pp. 180-193. (In Russ.) (<https://doi.org/10.18255/1818-1015-2020-2-180-193>).
10. **Гайнуллина А.Н., Сухов В.Д., Шалыто А.А., Сергушичев А.А.** Применение метода независимых компонент для определения начального приближения при поиске активных модулей в биологических графах // Научно-технический вестник информационных технологий механики и оптики. 2020. № 6, с. 888-892 (<https://ntv.ifmo.ru/file/article/20016.pdf>).

11. **Лобода А.А., Сергушичев А.А.** Вывод генных регуляторных сетей по данным экспрессии генов при помощи байесовских сетей // Научно-технический вестник информационных технологий механики и оптики. 2020. № 6, с. 835-840 (<https://ntv.ifmo.ru/file/article/20007.pdf>).
12. **Гайнуллина А.Н., Артемов М., Сергушичев А.А.** Модификация метода совместной кластеризации в графовом и корреляционном пространствах // Научно-технический вестник информационных технологий механики и оптики. 2020. № 6, с. 807-814 (<https://ntv.ifmo.ru/file/article/20004.pdf>).
13. **Alexeev N., Isomurodov J., Sukhov V., Korotkevich G., Sergushichev A.** Markov Chain Monte Carlo for Active Module Identification Problem // BMC Bioinformatics. Vol. 21. Article number: 261 (2020). Open Access. Selected articles from the 15th International Symposium on Bioinformatics Research and Applications (ISBRA-19): bioinformatics (<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-020-03572-9>).
14. **Aguilar S.V., Aguilar O., ... Gainullina A., et al.** The Immunological Genome Project. ImmGen at 15 // *Nature Immunology*. V. 21, 700-703 (2020) (<https://doi.org/10.1038/s41590-020-0687-4>).
15. **Rohl S., Rykaczewska U., Seime T., Suur B., Diez M., Gadin J., Gainullina A., Sergushichev A. et al.** Transcriptomic profiling of experimental arterial injury reveals new mechanisms and temporal dynamics in vascular healing response // *JVS: Vascular Science*. V.1, pp. 13-37 (2020) (<https://jvsvs.org/action/showPdf?pii=S2666-3503%282020%2930001-8>).
16. **Williams J., Zaitsev K., Kim K., Ivanov S., Saunders B.T., Schrank P., Kim K., Elvington A., Kim S., Tucker C., Wohltmann M., Fife B., Epelman S., Artyomov M., Lavine K., Zinselmeyer B., Choi J., Randolph G.** Limited proliferation capacity of aorta intima resident macrophages requires monocyte recruitment for atherosclerotic plaque progression // *Nature Immunology*. V. 21. 2020, pp. 1194-1204 (<https://www.nature.com/articles/s41590-020-0768-4>).
17. **Mogilenko D., Shpynov O., Andhey P., Arthur L., Swain A., Esaulova E., Brioschi S., Shchukina I., Kerndl M., Yao Z., Laha A., Zaitsev K., Burdett S., Gillfilan S., Stewart S., Colonna M., Artyomov M.** Comprehensive profiling of aging immune system reveals clonal GZMK+ CD8 T cells as conserved hallmark of inflammaging // *Immunity*. V. 54. Issue 1. 2021, pp. 99-115. e12 (<https://doi.org/10.1016/j.immuni.2020.11.005>).
18. **Perepelina K., Klauzen P., Khudiakov A., Zlotina A., Fomicheva Y., Rudenko D., Gordeev M., Sergushichev A., Malashicheva A., Kostareva A.** Generation of two iPSC lines (FAMRCi006-A and FAMRCi006-B) from patient with dilated cardiomyopathy and Emery–Dreifuss muscular dystrophy associated with genetic variant LMNAp.Arg527Pro // *Stem Cell Research*. 43(2020), 101714 (<https://www.sciencedirect.com/science/article/pii/S1873506120300167?via%3Dihub>).
19. **Zlotina A., Melnik O., Fomicheva Y., Skitchenko R., Sergushichev A., Shagimardanova E., Gusev O., Gazizova G., Loevets T., Vershinina T., Kozyrev I., Gordeev M., Vasichkina E., Pervunina T., Kostareva A.** A 300-kb microduplication of 7q36.3 in a patient with triphalangeal thumb-polysyndactyly syndrome combined with congenital heart disease and optic disc coloboma: a case report // *BMC Medical Genomics*. 13, Article number 175 (2020) (<https://bmcmmedgenomics.biomedcentral.com/articles/10.1186/s12920-020-00821-x>).
20. **Иванова О.А., Игнатьева Е.В., Леявина Т.А., Галенко В.Л., Комарова М.Ю., Борцова М.А., Ситникова М.Ю., Костарева А.А., Сергушичев А.А., Дмитриева Р.И.** Анализ транскриптома скелетной мускулатуры выявил влияние физических тренировок на молекулярные механизмы регуляции роста и метаболизма мышечной ткани у пациентов с хронической сердечной недостаточностью // *Российский кардиологический журнал*. 2020. № 10, с. 80-86 (<https://doi.org/10.15829/1560-4071-2020-4132>).
21. **Ignatieva E.V., Ivanova O.A., Komarova M.Y., Khromova N.V., Polev D.E., Kostareva A.A., Sergushichev A.A., Dmitrieva R.** LMNA mutations G232E and R482L cause dysregulation of skeletal muscle differentiation, bioenergetics and metabolic gene expression profile // *Genes*. 2020. 11(9). 1057 (<https://doi.org/10.3390/genes11091057>).
22. **Perepelina K., Kostina A., Klauzen P., Khudiakov A., Rabino M., Crasto S., Zlotina A., Fomicheva Y., Sergushichev A., Oganessian M., Dmitriev A., Kostareva A., Pasquale E.D., Malashicheva A.** Generation of two iPSC lines (FAMRCi007-A and FAMRCi007-B) from patient with Emery–Dreifuss muscular dystrophy and heart rhythm abnormalities carrying genetic variant LMNA p.Arg249Gln // *Stem Cell Research*. V. 47. 2020. 101895 (<https://doi.org/10.1016/j.scr.2020.101895>).
23. **Vershinina T., Fomicheva Y., Muravyev A., Jorholt J., Kozyreva A., Kiselev A., Gordeev M., Vasichkina E., Segrushichev A., Pervunina T., Sjoberg G., Skyttner-Rahmani S., Sejersen T., Kostareva A.**

Genetic Spectrum of Left Ventricular Non-Compaction in Pediatric Patients // Cardiology. 2020-10-13 (<https://www.x-mol.com/paper/1316154892677779456>).

Доклад на конференции

Gainullina A., Sergushichev A., Artyomov M. Network perspective on metabolic diversity among mononuclear phagocytes / Fourth International Conference «Bioinformatics: from Algorithms to Applications» (BiATA 2020). <https://www.lektorium.tv/node/37188>.

Свидетельство о регистрации программы

Сергушичев А.А., Гайнуллина А.Н., Артемов М. Программа для идентификации метаболических модулей по транскриптомным данным со сложным экспериментальным дизайном // Свидетельство о государственной регистрации программы для ЭВМ. № 2020 619734. Дата регистрации – 21.08.2020 г.

Интервью Карины Пац

12 марта на портале университета появилось интервью с сотрудницей научного центра «Компьютерные технологии» Карины Пац, которая проводит исследования механизмов работы ядерных рецепторов в Назарбаев Университете под руководством профессора Фердинанда Молнара (<https://news.itmo.ru/ru/education/trend/news/9247/>).

Статья Екатерины Носковой

27 мая Екатерина Носкова на портале «Хабр» опубликовала статью «Истории развития популяций, генетика и генетические алгоритмы» (<https://habr.com/ru/company/JetBrains-education/blog/502244/>).

Победитель программы ITMO Fellowship Анна Жук

Третьего июля на портале университета была опубликована статья (https://news.itmo.ru/ru/science/life_science/news/9542/) о новом победителе программы ITMO Fellowship Анне Жук, которая будет работать в нашем международном научном центре и заниматься сравнительной геномикой.

Новая специализация в магистратуре – «Алгоритмическая биоинформатика»

Девятого июля на портале нашего университета появилась статья об открытии новой специализации «Алгоритмическая биоинформатика» на рамках англоязычной магистерской программы «Биоинформатика и системная биология», возглавляемой Алексеем Сергушичевым (<https://news.itmo.ru/ru/education/trend/news/9562/?fbclid=IwAR1UM2JISxEj3TKOmpR2zFgXTxn5U2chyipeW-rsW0XSmNZUtlzpQXBool>). Она предназначена для выпускников бакалаврских программ в области математики и программирования, которые хотят заниматься биоинформатикой.

Награда на конференции GECCO 2020

12 июля на закрытии конференции GECCO 2020 было объявлено, что в рамках 17th Annual Human-Competitive Results («Humies») 2020 **Bronze Awards: Noskova G., Ulyantsev V., Koepli K-P., O'Brien S., Dobrynin P.** GADMA: Genetic Algorithm for Inferring Demographic History of Multiple Populations from Allele Frequency Spectrum Data» (https://news.itmo.ru/ru/university_live/achievements/news/9585/).

Результаты экспертизы по государственному заданию

23 июля, наконец-то, закончилась научно-исследовательская работа: Государственное задание Министерства образования и науки РФ. Тема: «Разработка алгоритмов анализа метаболических и сигнальных сетей для идентификации модулей, регулирующих клеточную адаптацию», которая под моим руководством выполнялась 2017-2019 гг. В заключении РАН, в частности, сказано: «Работа признана фундаментальным исследованием. Тема признана актуальной, а разработанные в ходе ее выполнения методология и подходы имеют высокую научную ценность. Понимание того, как метаболизм регулирует иммунный ответ и развитие раковых опухолей, открывает новые подходы для модуляции этих процессов. Признано, что финансирование темы за счет государственного бюджета было целесообразным, кадровый потенциал темы – высокий. Научный руководитель оценен положительно. Тема соответствует заявленным направлениям научно-технологического развития РФ». Заключение подписали вице-президент РАН

В.П. Чехонин (иммунохимик) и академик-секретарь отделения физиологических наук РАН **В.А. Ткачук** (биохимик).

Победа в международном соревновании по биоинформатике

24 августа на портале университета появилась статья «Команда биоинформатиков университета ИТМО одержала победу в международном турнире *MEDIC*» (https://news.itmo.ru/ru/life_science/science/news/9674/). В ходе этих соревнований участники должны были создать алгоритм, который на основании данных о микробиоте кишечника (составе бактерий, живущих в нем) определял бы страдает ли пациент воспалительными заболеваниями кишечника. Сотрудники нашей международной научной лаборатории Артем Иванов и Владимир Ульянцев победили сразу в двух подтурнирах.

Одним из перспективных методов диагностики болезней кишечника является **метагеномный анализ его микробиоты**. Соревнование *Metagenomics Diagnosis for Inflammatory Bowel Disease Challenge* проводилось на сайте организаторов (<https://www.intervals.science/resources/sbv-improver/medic>). В этих соревнованиях приняло участие 26 команд из разных стран мира: десять – из Италии, по шесть – из Греции и Японии, по три – из Финляндии и США. Оба подтурнира проходили в стандартном формате для соревнований по машинному обучению: сначала участникам давались размеченные данные микробиотного исследования 150 пациентов. На них нужно было обучить алгоритм правильно определять состояние человека – здоров ли он, болен ли язвенным колитом или страдает от болезни Крона. После этого обученный анализатор должен был проверить набор таких же данных, только без отметок, и сам сделать вывод, к какой из трех категорий относится каждый пациент.

В первом задании было известно, какие бактерии населяют кишечник и в каком количестве. Эти данные не надо было никак преобразовывать, и поэтому наши ребята сразу запускали алгоритмы машинного обучения. Они попробовали различные алгоритмы бустинга, один из которых сработал. *Второе задание* было гораздо более интересным – предоставлялись лишь сырые данные прочтения метагенома – короткие последовательности ДНК микробиоты пациентов, о которых ничего не было известно. Командам самим предстояло выделить признаки, которые являются ключевыми для постановки диагноза, и на их основе сделать предсказание на неразмеченных данных. Наши ребята использовали разработанную при нашем участии программу *MetaFast*, позволяющая по информации о ДНК выделять некоторые связанные структуры, которые и могут являться ключевыми. На основе этих структур наши участники осуществили машинное обучение и создавали классификатор».

В первом соревновании места распределились так: 1. Team CTLAB@ITMO, ITMO University, Russian Federation: Artem Ivanov, Vladimir Ulyantsev; 2. Team GG-GUMC, Georgetown University Medical Center, USA: Garrett Graham; 3. Team CDS-Lab, HPC and Networking Institute, CNRS, Italy: Mario Rosario Guarracino, Maurizio Giordano, Ichcha Manipur, Ilaria Granata, Lucia Maddalena, Marina Piccirillo, а во втором: 1. Team CTLAB@ITMO, ITMO University, Russian Federation: Artem Ivanov (TL), Vladimir Ulyantsev; 2. Team Mignon, University of Luxembourg: Enrico Glaab; 3. Team GiGi, University of Padova, Italy: Barbara Di Camillo, Marco Cappellato, Mehdi Poursheikhali Asghari, Sebastian Daberdaku, Giacomo Baruzzo, Filippo Pietrobon, Ilaria Patuzzi (<https://www.pmiscience.com/whats-new/winners-of-sbv-improver-machine-learning-diagnostic-challenge-announced>). Помимо соревновательной, *MEDIC* имел также и научную составляющую: по итогам турнира команды-победители приглашались к сотрудничеству и созданию совместных публикаций.

Статья с Анастасией Гайнуллиной, и ее защита диссертации

В июле была опубликована статья *Гайнуллина А.Н., Шалыто А.А., Сергушичев А.А.* Метод совместной кластеризации в графовом и корреляционном пространствах // Моделирование и анализ информационных систем. 2020. Том 27 (2), с.180-193. (*Gainullina A.N., Shalyto A.A., Sergushichev A.A.* Method of the Joint Clustering in Network and Correlation Spaces. *Modeling and Analysis of Information Systems*. 2020. 27(2), pp. 180-193. (In Russ.) <https://doi.org/10.18255/1818-1015-2020-2-180-193>). Известно, что одним из широко используемых подходов является решение задачи поиска активного модуля, в которой в графе биологических взаимодействий выделяется связный подграф, лучше всего отражающий разницу между двумя рассматриваемыми биологическими состояниями. В настоящей работе этот подход расширяется на случай большего числа

биологических состояний и формулируется задача совместной кластеризации в графовом и корреляционном пространствах. Для решения этой задачи предлагается итеративный метод, принимающий на вход граф G и матрицу X , в которой строки соответствуют вершинам графа. На выходе алгоритм выдает набор подграфов графа G так, что каждый подграф является связным и строки, соответствующие его вершинам, обладают высокой попарной корреляцией. Эффективность метода подтверждается экспериментальным исследованием на смоделированных данных.

24.12.2020 г. Анастасия защитила в Университете ИТМО диссертацию на соискание ученой степени кандидата технических наук и *PhD-dissertations* на тему «Метод графового анализа транскриптомных данных для обнаружения метаболической регуляции иммунных клеток» Руководитель – А.А. Сергушичев (<http://dissovet.itmo.ru/dissertation/?number=227442>).

При этом отмечу, что «Биоархиве» находится статья, в которой **Анастасия первый автор**, а Леша Сергушичев и Максим Артемов – последние, и это при том, что **в статье 44 автора из 19 организаций мира: Gainullina A. et al.** Open-Source ImmGen: Network Perspective Onmetabolic Diversity Among Mononuclear Phagocytes // bioRxiv, 2020. <https://www.biorxiv.org/content/10.1101/2020.07.15.204388v1.article-info>. После этого я стал уважительно называть Настю – 44/19.

В это время на портале *Research Gate* эти трое имели такие показатели: **Артемов** – 42.99, 172 публикации, 23249 прочтений, 7987 цитирований; **Сергушичев** – 32.15, 55, 9081, 2445; **Гайнуллина** – 15.89, 10, 1848, 116!

Про Лешу и Настю я написал текст «44/19» (<https://vk.com/@1077823-4419>). А еще я написал текст «Я горжусь тобой, Леша!» (<https://news.itmo.ru/ru/blog/336/>).

Первая серьезная статья с участием Геннадия Короткевича

18.11.2020 г. была опубликована первая серьезная статья с участием Гены (*Alexeev N., Isomurodov J., Sukhov V., Korotkevich G., Sergushichev A.* Markov Chain Monte Carlo for Active Module Identification Problem // BMC Bioinformatics. Vol. 21. Article number: 261 (2020). Open Access. Selected articles from the 15th International Symposium on Bioinformatics Research and Applications (ISBRA-19): bioinformatics (<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-020-03572-9>).

На портале нашего университета про эту статью вышел текст «Ученые представили метод, который позволяет оценить вероятность связи гена и болезни» (<https://news.itmo.ru/ru/science/it/news/9933/>). С помощью этого метода **можно определять, какова возможность того, что тот или иной ген или гены влияют на различные процессы в организме, в том числе на развитие заболеваний. Сравнивая гены человека, у которого есть интересующая исследователей болезнь или другой процесс в организме, с генами тех, у кого этой болезни нет, можно увидеть активность каких генов отличается, а затем на основе этих данных появляется возможность предположить, что некоторые из отличающихся генов могут иметь большее значение для развития интересующей исследователей болезни.**

Потом Геннадий стал соавтором еще одной работы: *Korotkevich G., Sukhov V., Budin N., Shpak B., Artyomov M., Sergushichev A.* Fast gene set enrichment analysis // bioRxiv. 2021, pp. 1-40. <https://doi.org/10.1101/060012>.

Лекция Никиты Алексева

14 декабря Никита Алексеев прочел для студентов лекцию «Комбинаторные задачи в биоинформатике» (https://www.youtube.com/watch?v=u2tqq9KPVZg&feature=emb_logo).

Группа биоинформатики в JetBrains Research

В конце декабря компания *JetBrains* опубликовала информацию о деятельности исследовательских групп в *JetBrains Research* (https://research.jetbrains.org/ru-ru/groups/?fbclid=IwAR0UNyAQ-8YRq2iVXSkdECU8JZdLIWBYsPZWbW_wiX78KoQ1DZ_DCFsdd6k). Одна из них – группа биоинформатики в составе А. Сергушичев (руководитель), К. Зайцев, С. Казаков и А. Лобода (<https://research.jetbrains.org/ru-ru/groups/biogroup/>). Она занимается применением методов информатики для решения биологических и медицинских проблем. Целью группы является разработка эффективных вычислительных методов для актуальных задач биологии и медицины. Проекты покрывают достаточно широкий спектр задач биологии и медицины: от обработки данных метагеномного ДНК-секвенирования до анализа экспрессии генов и метаболизма. Группа базируется

на кафедре «Компьютерные технологии» Университета ИТМО и поддерживается компанией *JetBrains*. Коллектив активно сотрудничает с лабораторией Максима Артемова в Университете Вашингтона в Сент-Луисе.

Основные проекты. 1. Сравнительный анализ метагеномных данных. В рамках проекта разрабатывается программа *MetaFast*. В этом методе применяется компромиссный вариант между использованием метагеномной *de novo*-сборки и спектров *k*-меров. Такая комбинация позволяет объединить достоинства этих подходов: интерпретируемость и вычислительную эффективность, при отсутствии необходимости наличия заранее известных референсных геномных последовательностей. **2. Анализ открытых баз экспериментальных данных.** Такие базы, как *TCGA* и *GEO Omnibus*, хранят данные большого числа экспериментов. Для поиска связей в этих данных разработан веб-сервис *GeneQuery*, позволяющий искать эксперименты с похожей регуляцией генов. Поиск неочевидных связей с другими экспериментами упрощает процесс генерации гипотез. **3. Методы анализа для изучения метаболической регуляции.** В последнее время была отмечена роль метаболической регуляции, особенно в контексте иммунной системы и развития опухолевых клеток. Данный проект посвящен разработке методов для изучения метаболической регуляции с помощью анализа данных транскриптомики и метаболомики. В рамках проекта был разработан набор методов, доступных в виде веб-сервиса *GAM* для интегративного сетевого анализа таких данных.

Первый выпуск магистерской программы по специальности «Биоинформатика и системная биология»

10 июня на кафедре появились первые выпускники магистерской программы по биоинформатике и системной биологии. Как считает ее руководитель Алексей Сергушичев, все студенты отлично справились и успешно защитили квалификационные работы. Он говорит, что это были трудные два года для студентов, но и еще более три трудных года для ее организаторов: первые дискуссии о том, чтобы в Университет ИТМО была магистратура по биоинформатике начались в январе 2017 г. Алексей поблагодарил и поздравил всех тех, кто участвовал в этом процессе, и, в частности, Виталия Аксенова – с него все в этом вопросе началось, Николая Вяхи и его команду из Института биоинформатики – они сильно помогали, Ксению Спиридонову – первого координатора программы и Максима Артемова, с которым наши ребята организуют ежегодные семинары по системной биологии, из которых выросла программа, и кто вдохновил Лешу пройти через все это.

28 августа на портале Университета была опубликована статья, в которой Ольга Иванова рассказывает об обучении в магистратуре по специальности «Биоинформатика и системная биология», которую она закончила в этом году (<https://news.itmo.ru/ru/education/trend/news/9689/>). Ольга в своей магистерской диссертации, выполненной в Центре Алмазова под руководством Ренаты Дмитриевой и Алексея Сергушичева, исследовала влияние мутаций в определенном гене на дифференцировку стволовых клеток мышечной ткани. При этом для изучения экспрессии генов она использовала биоинформатический метод «Анализ данных РНК-секвенирования», который позволил обнаружить те молекулярные пути, которые подверглись изменениям. «Мне интересна тема старения. Мутации, которые мы изучаем, на самом деле с этой темой связаны. Так, существует мутация в этом гене, которая приводит к «прогерии» – «синдрому Хатчинсона-Гилфорда (синдрому преждевременного старения)». В среднем пациенты с этой мутацией живут 13 лет, и уже у совсем маленьких детей наблюдаются все признаки старения органов, тканей, кожи, мышц. Три миллиарда пар нуклеотидов кодируют весь наш организм. И одна точечная мутация в одной из этих пар приводит к тому, что стареет все тело – и никто не понимает, почему. Мне было бы очень интересно в этом разобраться», – в заключение сказала Ольга.

Кстати, в статье «Как расшифровать геном и изобрести новое лекарство, не вставая из-за компьютера» (https://news.itmo.ru/ru/science/it/news/9760/?fbclid=IwAR2jYNbjWMZ8qiyMjg-qArWs2CnUpKSLZ-sFhPgJc7Koz_82TWWZQv_Uacs) Ольга совместно с Антоном Замятиным попыталась дать ответы на следующие восемь вопросов: **1.** Что такое биоинформатика? **2.** Как она появилась? **3.** Зачем она нужна? **4.** Где она используется? **5.** А где еще? **6.** Чем занимается биоинформатик? **7.** Биоинформатик – это программист или биолог? **8.** Как биоинформатика может изменить наше будущее?

Как это должно быть, или история одной выпускницы первого выпуска магистратуры
08.12.2020 г. на портале Университета ИТМО была опубликована статья «Я почувствовала, что мне нужно новое дыхание в науке»: выпускница ИТМО – о поиске себя, междисциплинарности и пути в биоинформатику» (<https://news.itmo.ru/ru/education/trend/news/9962/>), в которой окончившая в этом году магистратуру нашей кафедры «Биоинформатика и системная биология» на английском языке (<http://news.ifmo.ru/ru/education/official/news/7466/>) кандидат наук (!) Наталья Петухова описывает, как училась у нас и сразу после этого возглавила «Научно-исследовательский центр биоинформатики» в Первом медицинском университете им. И. П. Павлова (<https://www.1spbgmu.ru/ru/nauka/nauchnie-podrazdeleniya/106-glavnaya/4879-nauchno-issledovatel'skij-tsentr-bioinformatiki>).

Она рассказывает: «Каждый курс был уникальным и очень нужным – полезно было все. Сейчас я понимаю, что в магистратуре нам дали полное и разностороннее представление о том, что может биоинформатика. Наверное, самым главным было то, что нам не только давали много теоретического материала, но и много практики: постоянные домашние задания и жесткие дедлайны. Это был абсолютно оправданный и полезный подход, так как мы должны были сами быстро искать решения, сами пробовать и общаться друг с другом. Самое главное, что я получила в ИТМО – широкий научный кругозор. Нам преподавали профессора и ученые из многих мировых школ и университетов, нам дали широкий взгляд на различные научные задачи и подходы. Главное, мы можем применять свои знания в абсолютно различных областях. Я получила в университете именно то, зачем шла в биоинформатику. Я хочу воспользоваться случаем и поблагодарить мою вторую *Alma mater* – **Университет ИТМО и всю его потрясающую команду**, которая очень многое мне дала и вырастила меня как нового специалиста. Все, что я имею сейчас, все мои знания в биоинформатике, я имею благодаря вам. **Надеюсь, что мы будем развивать эту научную сферу вместе».**

Последняя фраза, по-моему, ключевая! Я считаю, что именно так, и, никак иначе, должно заканчиваться образование! В противном случае университет что-то не дорабатывает и никогда не перейдет в «генераторный» режим!

В первом семестре 2020/2011 года Леша Сергушичев читает магистрантам лекции по курсу «Системная биология», Володя Ульяновцев – по курсу «Метагеномика».

Цель, области и некоторые направления исследований

Цель

Разработка эффективных вычислительных методов для актуальных задач биологии и медицины.

Области исследований

1. Анализ данных в иммунологии и раке.
2. Метагеномика.
3. Сравнительная геномика.

Направления исследований по биоинформатике и системной биологии

1. Исследования по методам интерпретации результатов полногеномного анализа ассоциаций, выполняемых совместно с лабораторией выдающегося ученого **Марка Дейли** в *Broad Institute of MIT and Harvard*. В этой области имеется шанс сделать прорыв за счет использования больших объемов публичных данных и их анализа с помощью эффективных вычислительных методов.
2. Систематический анализ открытых баз данных **экспрессии генов**. Разработаны инструменты по поиску, визуализации и анализу таких данных, которые должны быть объединены в единую платформу для того, чтобы дать возможность биологам простой и полезный доступ к имеющимся массивам данных.
3. Геномика животных и, в частности, исчезающих видов, под руководством всемирно известного профессора **Стивена О'Брайена** и наращивание экспертизы в области геномики человека.

Некоторые проекты

Эти проекты связаны с исследованиями для достаточно широкого спектра задач: от обработки данных геномного и метагеномного ДНК-секвенирования до анализа экспрессии генов и метаболизма.

1. Разработка веб-приложения **Phantassus** для визуализации и анализа данных экспрессии генов (<https://ctlab.itmo.ru/phantassus/>).
2. Разработка метода **Fast Gene Set Enrichment Analysis (FGSEA)** для быстрого анализа представленности функциональных наборов генов по данным экспрессии. В этой работе принимал участие Геннадий Короткевич (<https://bioconductor.org/packages/release/bioc/html/fgsea.html>).
3. Разработка вычислительных методов для изучения метаболической регуляции. В последнее время установлена роль метаболической регуляции, особенно в контексте иммунной системы и развития опухолевых клеток. Данный проект посвящен разработке методов для изучения метаболической регуляции с помощью анализа данных транскриптомики и метаболомики. В рамках проекта разработан набор методов, доступных в виде веб-сервиса для анализа метаболитических сетей «**Shiny GAM: integrated analysis of genes and metabolites**» (<https://artyomovlab.wustl.edu/shiny/gam>) для интегративного сетевого анализа таких данных.
4. Разрабатывается продукт **mcmcRanking**: «Tool for estimate probabilities of vertices being in active module using its likelihoods and it proposes methods for ranking vertices in order of importance. Estimating probabilities based on Markov chain Monte Carlo (MCMC) methods» (<https://github.com/ctlab/mcmcRanking/>).
5. В рамках проекта **сравнительный анализ метагеномных данных** разрабатывается программа **MetaFast** (<https://github.com/ctlab/metafast>). В методе, на котором базируется эта программа, применяется компромиссный вариант между использованием метагеномной *de novo* сборки и спектров *k*-меров. Это позволяет объединить преимущества этих подходов: интерпретируемость и вычислительную эффективность при отсутствии необходимости в наличии заранее известных референсных геномных последовательностей.
6. Базы экспериментальных данных, такие как **TCGA** и **GEO Omnibus**, хранят в себе данные большого числа экспериментов и интересным является **поиск связей** в этих данных, для обеспечения которого разрабатывается веб-сервис **GeneQuery** (<http://genome.ifmo.ru/genequery>), позволяющий искать эксперименты с похожей регуляцией генов. Поиск неочевидных связей с другими экспериментами упрощает процесс генерации гипотез. В этой работе, в частности, принимал участие наш магистрант **Иван Арбузов**.

Некоторые результаты, полученные при нашем участии

1. Показано как кишечная микробиота влияет на иммунный ответ против гриппа. Работа опубликована в журнале **Science**: Steed A., Christophi G., Kaiko G., Sun L., Goodwin V., Jain U., Esaulova E., Artyomov M., Morales D., Holtzman M., Boon A., Lenschow D., Stappenbeck T. The Microbial Metabolite Desaminotyrosine Protects from Influenza Through Type I Interferon // **Science**. 2017. Vol. 357, Issue 6350, pp. 498-502. **IF: 37.0, SJR: 13.5** (<https://science.sciencemag.org/content/357/6350/498>).
2. Определен молекулярный механизм действия итаконата на клетки иммунитета. Эти исследования помогут в изучении псориаза и других аутоиммунных заболеваний (https://museum.itmo.ru/news_page/442/). Работа опубликована в журнале **Nature**: Bambouskova M., Gorvel L., Lampropoulou V., Sergushichev A., Loginicheva E., Johnson K., Korenfeld D., Mathyer M., Kim H., Huang L., Duncan D., Bregman H., Keskin A., Santeford A., Apte R., Sehgal R., Johnson B., Amarasinghe G., Soares M., Satoh T., Akira S., Hai T., Strong C., Auclair K., Roddy T., Biller S., Jovanovic M., Klechevsky E., Stewart K., Randolph G., Artyomov M., Electrophilic Properties of Itaconate and Derivatives regulate Ikbz/Atf3 in Inflammatory Axis // **Nature**. 2018. Vol. 556. No. 7702, pp. 501-504. **IF: 40.137, SJR: 18.13** (<https://www.nature.com/articles/s41586-018-0052-z>).
3. Разработана новая животная модель для изучения патогенеза вируса Зика (<https://news.itmo.ru/ru/news/7552/>). Работа опубликована в журнале **Cell Host & Microbe**: Gorman M., Caine E., Zaitsev K., Begley M., Weger-Lucarelli J., Uccellini M., Tripathi S., Morrison J., Yount B., Dinnon K., Rückert C., Young M., Zhu Z., Robertson S., McNally K., Ye J., Cao B., Mysorekar I., Ebel G., Baric R., Best S., Artyomov M., Garcia-Sastre A., Diamond M. An Immunocompetent Mouse Model of Zika Virus Infection // **Cell Host & Microbe**. 2018. 23(5), pp. 672-685.e6. **IF: 14.946, SJR: 8.34** (<https://vk.com/@itmoru-uchenye-uluchshili-model-dlya-izucheniya-virusa-zika>).
4. Создан программный продукт, позволяющий быстро и эффективно анализировать геномные данные, и на их основе строить наиболее вероятные модели демографической истории тех или иных популяций растений, животных и людей. Статья опубликована в журнале **GigaScience**. Noskova G., Ulyantsev V., Koepli K-P., O'Brien S., Dobrynin P. GADMA: Genetic Algorithm for Inferring Demographic History of Multiple Populations from Allele Frequency

Spectrum Data // *GigaScience*. Vol. 9. 2020. Issue 3. IF: 4.688. SJR: 4.73 (<https://doi.org/10.1093/gigascience/giaa005>). Этому исследованию посвящена статья «Ученые Университета ИТМО создали алгоритм, позволяющий легче и точнее строить историю популяций людей и животных» (<https://news.itmo.ru/ru/science/it/news/9222/>).

5. Разработан алгоритм определения возраста по результатам анализа крови с учетом пола и национальности (<https://news.itmo.ru/ru/science/it/news/7260/>). Работа опубликована в журнале *The Journals of Gerontology: Mamoshina P., Kochetov K., Putin E., Cortese F., Aliper A., Lee W., Uhn L., Skjodt N., Kovalchuk O., Scheibye-Knudsen M., Zhavoronkov A. Population Specific Biomarkers of Human Aging: A Big Data Using South Korean, Canadian and Eastern European Patient Population // The Journals of Gerontology: Series A*. 2018. Vol. 73. Issue 11, pp. 1482-1490. IF: 5, 95. SJR: 2.79 (<https://academic.oup.com/biomedgerontology/article/73/11/1482/4801287/>).

6. Создана программа **MetaCherchant** для анализа метагенома, которая выявила новые причины возникновения супербактерий (https://news.itmo.ru/ru/science/life_science/news/7090/). Работа опубликована в журнале *Bioinformatics: Olekhnovich E., Vasilyev A., Ulyantsev V., Kostryukova E., Tyakht A. MetaCherchant: Analyzing Genomic Context of Antibiotic Resistance Genes in Gut Microbiota // Bioinformatics*. 2018. Vol. 34. № 3, pp. 434-444. IF: 7.307. SJR: 4.92 (<https://pubmed.ncbi.nlm.nih.gov/29092015/>).

7. Разработаны алгоритмы сборки данных секвенирования: сборка генома на основе совмещения подходов использования графов де Брёйна и графа перекрытий. Алгоритмы были доработаны для поддержки чтений *Ion Torrent*. Разработаны алгоритмы и программы для сборки транскриптома; сборки скэффолдов. Создан общедоступный сборщик, позволяющий проводить сборку генома на персональных компьютерах под управлением OS Windows и не только.

8. Создано программное средство для анализа вариаций и поиска маркеров, ассоциированных с кардиомиопатией.

9. Создано программное средство для совместного анализа данных секвенирования и экспрессии генов, а также различных эпигенетических данных определенного типа рака.

10. Выполнено исследование иммуногенности раковых эпитопов для создания персонализированных противораковых вакцин.

11. Разработаны алгоритмы анализа транскрипционных и метаболических данных в контексте метаболических сетей реакций.

12. Разработаны алгоритмы для сравнительного анализа метагеномных данных.

13. Выполнена сборка геномов редких бактерий.

14. Создан программный продукт, реализующий этот алгоритм, который позволяет быстро и эффективно анализировать геномные данные, и на их основе строить наиболее вероятные модели демографической истории тех или иных популяций растений, животных и людей.

Некоторые организации, с которыми проводились или проводятся совместные работы

1. *Washington University St. Louis (WashU)*. Лаборатория Максима Артемова.
2. *Broad Institute of MIT and Harvard* (Cambridge, MA). Лаборатория Марка Дейли.
3. Институт биоинформатики.
4. Федеральный центр сердца, крови и эндокринологии им. В.А. Алмазова (Санкт-Петербург). Институт молекулярной биологии и генетики под руководством Анны Костаревой.
5. Научно-исследовательский институт акушерства, гинекологии и репродуктологии им. Д.О. Отта (Санкт-Петербург). Лаборатория пренатальной диагностики наследственных и врожденных заболеваний под руководством Андрея Глотова.
6. Научно-исследовательский институт физико-химической медицины (Москва). Лаборатория биоинформатики под руководством Дмитрия Алексеева.
7. Казанский государственный университет (КФУ). Институт фундаментальной медицины и биологии, отдел микробиологии. Работы проводятся совместно с Анастасией Тойменцевой.

Информационные ресурсы

За новостями по рассматриваемой тематике можно следить в социальных сетях *Facebook* (<https://www.facebook.com/ITMObioinf/>) и «ВКонтакте» (https://vk.com/bioinf_itmo). Еженедельный семинар по биоинформатике проходит в *Telegram*-канале (https://t.me/ITMO_bioinf_seminar).

Вместо заключения

10 июля 2020 г. я неожиданно понял, почему «втянул» наших ребят в тематику сборки генома, которая переросла в биоинформатику и системную биологию. **Это оказывается произошло «генетически» :-).**

Оказалось, что *PhD*-диссертация К.Э. Шеннона, которую он защитил в 1940 г., называлась *An Algebra for Theoretical Genetics*, <https://dspace.mit.edu/handle/1721.1/11174>. Эта работа по духу очень напоминала его статью 1938 г. по теории ... переключательных схем (*A Symbolic Analysis of Relay and Switching Circuits*) (<https://ieeexplore.ieee.org/document/5057767>). Я же свою научную деятельность начинал с перечисления параллельно-последовательных контактных схем с целью построения многофункциональных логических модулей из элементов с двусторонней проводимостью (Глава 13 моей книги Логическое управление. Методы аппаратной и программной реализации алгоритмов. СПб.: 2000. 780 с. (http://is.ifmo.ru/books/log_upr/1)). Перечисления я выполнил правильно, так как результаты, полученные мною, были в статье *Riordan J., Shannon C.E. The Number of Two-Terminal Series-Parallel Networks / J.Math. Phys.1942. 21, 83-93* (https://oeis.org/A000084/a000084_1.pdf). Потом эти статьи появились на русском языке в книге Шеннон К.Э. Работы по теории информации и кибернетике. М.: Изд-во иностранной литературы. 1963, 829 с.

В 1971 г. я их прочел, и моя научная жизнь понеслась...

А еще, напомним, что с Сергушичевым, Казаковым и Александровым я познакомился, во время курсовика, в рамках которого они с помощью ни какого-нибудь, а **генетического программирования** строили автомат, управляющий моделью самолета, выполняющего «мертвую петлю» (http://is.ifmo.ru/works/2013/alexandrov_samolet.pdf), что позволило им плавно перейти :-) **к сборке генома!**

Старт работам по биоинформатике и системной биологии в нашем коллективе и Университете ИТМО я дал, историю этого «старта» написал. «Летите» дальше, если вам это надо, сами...

22.09.2020.