

## Мария Самсонова: «Математика - новый микроскоп для биологии»



Мария Самсонова: «Математика - новый микроскоп для биологии»

Интересный факт: если ежедневно по восемь часов набирать по одному слову в секунду, то потребуется 50 лет, чтобы напечатать геном человека. Неудивительно, ведь в геноме содержится весь наследственный материал организма. Только вот как изучать такие сложные биологические системы? С помощью математики и информатики!



Новой героиней спецпроекта «Настоящие великие» стала Мария САМСОНОВА – первый в Санкт-Петербурге доктор наук по биоинформатике, профессор Высшей школы прикладной математики и вычислительной физики СПбПУ. В Политехе она также возглавляет Научно-исследовательскую лабораторию математической биологии и биоинформатики, где изучают геномику растений и разрабатывают новые методы для их селекции.

Как связаны биология и математика, всегда ли вредны мутации, зачем растениям «мусорные» ДНК и почему так сложно создавать новые сорта – об этом и многом другом узнаете из интервью.

- Мария Георгиевна, почему вы решили развивать в Политехе направление математической биологии и биоинформатики? Казалось бы, не совсем профильные для нашего университета сферы.

- Не зря математику считают царицей всех наук. Сейчас для прогресса в биологии она играет такую же роль, какую для становления физики в XVII веке. Благодаря успехам биологии развивается математика, и наоборот, новые математические методы приводят к тому, что решаются задачи биологии. Эти области знания взаимно обогащают друг друга. Поэтому и появилась математическая биология.

Существует еще биоинформатика – это более узкая область математической биологии, которая занимается обработкой и анализом данных молекулярных и биологических экспериментов. То есть математическая биология может анализировать и экосистему,

и целый организм, и его определенную ткань, а биоинформатика работает только с данными молекулярного уровня.

В Политехе всегда была сильна прикладная математика. После учебы в СПбГУ я пошла работать в Институт высокопроизводительных вычислений и баз данных, а потом институт стал частью Политехнического университета. И это счастье для меня, потому что здесь в 2001 году я организовала Отдел компьютерной биологии, который сейчас перерос в целую научно-исследовательскую лабораторию.



- Сейчас биоинформатика считается профессией будущего. Вы согласны?

- Это действительно так. Биоинформатика – довольно молодая наука. Она появилась после того, как ученые расшифровали ДНК и научились анализировать белки. Поскольку и ДНК, и белки – это полимеры, то есть молекулы сложного строения, возникла потребность в компьютерных методах для их изучения. Так и зародилась биоинформатика, которая стала активно развиваться с 1980-х годов.

Биоинформатика и математическая биология – это и правда профессии будущего. Биологический эксперимент очень дорог: он требует времени и большого количества вложений. С помощью математики можно провести эксперименты *in silico* (*In silico* – термин, обозначающий компьютерное моделирование. Фраза создана по аналогии с *in vivo* (в живом организме) и *in vitro* (в пробирке). – Примеч. Авт.). И на основе этих

данных сделать предсказания, которые затем можно проверить с помощью направленного эксперимента.

Биоинформатика экономит время, трудовые ресурсы и деньги. Более того, позволяет моделировать эксперименты, которые проводить нельзя. Мы же не можем у человека получать мутации, зато можем смоделировать метаболический путь или какую-то другую систему организма и сформулировать гипотезы о механизмах возникновения какого-то заболевания. Поясню, очень часто болезнь возникает из-за нарушения работы определенных сигнальных или метаболических путей. Методы биоинформатики позволяют нам сделать модель болезни и посмотреть, какие нарушения к ней приводят и как их можно компенсировать. Потом ученые проверяют такие гипотезы в эксперименте.



- Это как создать цифровой двойник?

- В принципе, да. Хотя создать цифровой двойник болезни трудно, потому что организм сам по себе очень сложная система, более того, каждый организм уникален. Кстати, сейчас в создании вакцины от COVID-19 огромную роль играют биоинформатические методы. Потому что прежде чем создавать вакцины или

лекарства, необходимо посмотреть, как они будут работать в организме, то есть создать математическую модель.

- А вы участвовали в работе над вакциной?

- Нет, но мы планируем провести некоторые исследования, связанные с COVID-19. Врачи предоставят нам данные, а мы будем их анализировать. Это наш совместный проект с Институтом биомедицинских систем и биотехнологий Политеха, который возглавляет [Андрей Владимирович Васин](#).

- Геном человека, который состоит из трех миллиардов нуклеотидов, так называемых строительных блоков ДНК, был прочитан 20 лет назад. Какие задачи стоят перед биоинформатиками сейчас?

- То, что мы прочитали геном, еще ничего не значит. Даже если в организме найдут вредные мутации, в этом нет ничего страшного. Потому что гены работают не сами по себе, а совместно – есть целая цепочка превращений. Так что вредная мутация в одной части гена и такая же вредная мутация в другой могут взаимно компенсировать друг друга.



- Минус на минус дает плюс.

- Совершенно верно. Мало прочесть геном – надо понять, как работает процесс в

совокупности, как устроена система целиком. Это очень сложно, и человечество еще не научилось это делать. Еще одна важная проблема заключается в том, что ген – это не конечный этап передачи информации от ДНК к признаку организма. Из гена синтезируется матричная РНК (РНК, содержащая информацию о первичной структуре белков. – Примеч. Авт.), а матричная РНК превращается в белок. Сейчас все секвенируют геномы, но будущее за протеомикой – наукой, характеризующей репертуар белков в организме. Белки – это ферменты и строительные молекулы, из которых созданы мы все, которые могут включать и выключать гены. Все в нашем организме определяется белками.

- Секвенирование генома - это когда меняют цвет глаз?

- Нет, это уже редактирование генома. А секвенирование означает, что вы прочитываете, из каких нуклеотидов состоит ваша ДНК.

- А включать и выключать гены - это что значит?

- Выключенный ген – это ген, который существует, но по определенным причинам молчит. То есть он выключен не в результате мутации, а потому что на нем находятся белки, которые останавливают его работу. Самое главное, что некоторые такие комплексы могут передаваться потомству. У животных это не распространено, а у растений такие изменения наследуются и называются эпигенетическими.



- Вы как раз занимаетесь исследованием геномов сельскохозяйственных культур. Почему это важно?

- Окультуривание растений и интенсификация сельского хозяйства в ходе зеленой революции в середине 20 века привели к тому, что генетическое разнообразие культурных растений уменьшилось. Из-за этого нам сложнее создавать новые сорта, потому что генетическое разнообразие – это исходный материал для селекции.

Что такое генетическое разнообразие? Это замены нуклеотидов в ДНК. В каждом человеке одна нуклеотидная замена (называется однонуклеотидным полиморфизмом) встречается один раз на 300 нуклеотидов. Это значит, каждого человека отличает примерно  $10^7$  замен или полиморфизмов. Такие замены в генах растений – материал для селекции и их адаптации к условиям произрастания. Чем больше замен в гене и чем разнообразнее эти замены в разных растениях, тем богаче материал для селекции и тем скорее растение будет адаптироваться к новым климатическим условиям. Так вот, зеленая революция привела к тому, что гены культурных растений стали похожи друг на друга и замен нуклеотидов в них стало меньше. А значит, селекция становится неэффективной и растения теряют способность к адаптации, что очень плохо, поскольку климат сейчас начал сильно меняться. Генетическую изменчивость сохранили дикие родичи сельскохозяйственных растений, и актуальной задачей

является вовлечение этой генетической изменчивости в процесс создания новых сортов.

Кроме того, сейчас в России семена многих сельскохозяйственных культур завозные, и тут уже встает вопрос продовольственной безопасности. Перестройка привела к тому, что многие сорта были утеряны, селекционная база была развалена.

А вообще, растения интересны тем, что из-за множества мобильных элементов у них огромные геномы.

- И что это значит?

- Мобильные генетические элементы – это последовательности ДНК, которые могут перемещаться внутри генома. Геном любого организма состоит из смысловой части, которая кодирует белки, и из еще одной части, достаточно большой, которая пока мало изучена. Раньше считалось, что это junk DNA, то есть в переводе с английского «мусорная» ДНК. Геном растений буквально набит мобильными элементами, которые, как выясняется, играют важную регуляторную роль. Откуда у растений так много мобильных элементов, сейчас разбираются ученые. Видимо, они играли большую роль в процессе эволюции.



- А чем так важна регуляторная функция?

- Она показывает, включен ген или нет. Какие-то гены работают всегда, например те, которые кодируют белки и ферменты клеточной машины, синтезирующей белки в любой клетке. А какие-то включаются при определенных условиях, например, когда клетка начинает делиться.

- И вы создаете математические модели, прогнозируя, когда произойдет включение генов?

- Да, математика вообще считается новым микроскопом для биологии.

- Работаете в партнерстве с сельхозпроизводителями?

- Нам в лаборатории интересны более фундаментальные исследования, но, конечно, с практическим приложением. В основном мы работаем по грантам РФФИ и РНФ, были несколько ФЦП по сое с компаниями-производителями.

- Вы помогаете им выводить генномодифицированные сорта?

- Нет, мы исследуем дикие образцы растений, чтобы мобилизовать в них изменчивость и применить ее в селекции. Генная модификация же заключается в том, что вы вводите в организм чужеродную ДНК, поэтому все генномодифицированные организмы содержат в геноме фрагменты чужой, часто бактериальной, ДНК. Есть еще редактирование гена, когда направленно меняется его последовательность, при этом следов чужеродной ДНК в геноме не остается. Поэтому пока организмы с отредактированными генами не считают генномодифицированными.

Мы не получаем генномодифицированные растения и не редактируем гены, а изучаем естественную вариацию в геномах и стараемся понять, как на основе нее формируются признаки организма. Если подходить механистически, то когда мы поймем, как устроен тот или иной признак или орган растения, например цветок, можно научиться его конструировать. Это безумно сложно, потому что в биологии большую роль играет стохастика – случайность. Как случайность переходит в закономерность, это удивительно и интересно. Над разгадкой этой тайны работают ученые по всему миру.

- Уверена, все получится и ваш коллектив сыграет в этом процессе значимую роль. Спасибо за интересный разговор!