

## ОТЗЫВ

на автореферат диссертации Солдато́ва Руслана Андреевича

«Методы предсказания структурных элементов РНК»

на соискание учёной степени кандидата физико-математических наук по специальности 03.01.09 –  
математическая биология, биоинформатика.

Диссертационная работа Солдато́ва Р.А. посвящена фундаментальной проблеме биоинформатики - предсказанию вторичной структуры РНК - и задаче повышения эффективности алгоритмов ее предсказания. В актуальности задачи не приходится сомневаться – в условиях экспоненциального увеличения объема данных о биологических объектах благодаря совершенствованию высокопроизводительных методов анализа биомакромолекул существует острая потребность в эффективных алгоритмах и программных инструментах обработки полученных данных. В особенности это касается методов анализа последовательностей нуклеиновых кислот, темпы роста производительности которых превысили даже известный закон Мура для вычислительной техники.

Как справедливо замечает автор во введении, за последние два десятилетия произошел впечатляющий прорыв в области биологии РНК, сопровождающийся открытием новых классов не кодирующей белок РНК (нкРНК) и соответствующим существенным увеличением веса РНК в регуляции функционального состояния клетки. Возможность участия этой биологической «темной материи» ранее не учитывалась в моделях регуляции функций клеток, и, фактически не существовало теоретических предпосылок для ее предсказания. Тем не менее, в настоящее время некодирующая РНК поразительным образом меняет наши представления о регуляции множества клеточных процессов.

Одной из наиболее существенных характеристик нкРНК, оказывающей определяющее влияние на ее биогенез, функционирование и деградацию, является ее вторичная структура. На фоне значительных объемов данных о первичных последовательностях нуклеиновых кислот, качественной и экспериментально подтвержденной информации о вторичных структурах крайне немного.

Автор в работе выбирает два подхода к предсказанию вторичной структуры: анализ исключительно на основе структурированности и сравнительно-геномный метод. Новизной предложенного в работе варианта первого способа является воплощенная автором в алгоритме RNASurface оценка Z-значений структурированности общей последовательности с эффективным рекурсивным вычислением Z-значений подпоследовательностей. Сокращение времени работы алгоритма по сравнению с многократным применением алгоритма Зукера достигается за счет использования линейной аппроксимации фонового распределения энергий для фиксированного динуклеотидного состава, и набора квадратичных регрессий для фиксированной длины анализируемой последовательности. Разработанный Солдатовым Р.А. метод определения локально-оптимальных структурированных сегментов внутри последовательности позволяет отказаться от необходимости априорного выбора размера окна сканирования, что обеспечивает рост доли предсказанных структурных РНК. Действительно, необходимость априорного выбора окна

сканирования создает определенные проблемы в связи с тем, что структурированные фрагменты могут значительно отличаться по длине первичной последовательности. В проведенном сравнении авторского алгоритма с алгоритмом RNALfoldz получено почти двукратное повышение эффективности предсказания особенно при малом числе ложных структур.

Во второй части работы автор использовал диффузию как модель эволюции структурированности РНК. Такой новаторский метод позволяет автору более эффективно оценивать эволюцию Z-значения в условиях отсутствия отбора на структурированность. Результатом использования данного подхода на примере известных некодирующих РНК *Drosophila melanogaster* при использовании диффузионной модели на филогенетическом дереве мух стало значительное улучшение надежности предсказания некодирующих РНК.

Хотелось бы спросить автора, использовался ли им разработанный алгоритм RNASurface для анализа структурированности у эукариот, и каковы на взгляд автора пределы применимости метода, основанного исключительно на анализе структурированности, в связи с проблемой ограниченной функциональной значимости структур РНК, рассчитанных термодинамически в идеализированных условиях, существенно отличающихся от внутриклеточных.

Выводы, изложенные в работе Солдатов Руслана Андреевича, представляются обоснованными. Основные результаты работы опубликованы в отечественных и зарубежных журналах с хорошим импакт-фактором, что отражает высокий уровень работы. Результаты неоднократно докладывались на международных конференциях. Полученные результаты открывают перспективу построения комплексных эффективных программных продуктов для анализа данных высокопроизводительного секвенирования. Автореферат отражает основное содержание работы, написан ясным научным языком, раскрывает актуальность и значимость полученных результатов, однако, содержит некоторое количество опечаток, которые, впрочем, существенно не влияют на научное качество работы. Диссертационная работа соответствует требованиям ВАК, предъявляемым к диссертациям на соискание ученой степени кандидата физико-математических наук, а её автор, Солдатов Р. А., заслуживает присуждения искомой степени по специальности 03.01.09 – математическая биология, биоинформатика.

В.н.с. отдела молекулярной и клеточной патофизиологии,

Доц. кафедры общей патологии и патофизиологии ГБОУДПО РМАПО

рук. ЦКП ФГБНУ НИИ общей патологии и патофизиологии

к.м.н.

А.А. Московцев

✓ Подпись заверяю: Москувцев Руслан  
Ученый секретарь  
ФГБНУ НИИ общей патологии и патофизиологии  
Скуратовская

к.м.н.

27.11.2015



Л.Н.Скуратовская