

ОТЗЫВ

официального оппонента на диссертационную работу

Виноградовой Светланы Владимировны

на тему: «Предсказание структурных элементов РНК с использованием

экспериментальных данных»,

представленную на соискание ученой степени

кандидата биологических наук по специальности 03.01.09 «математическая

биология, биоинформатика»

Основной целью работы Виноградовой Светланы Владимировны была разработка метода для учета экспериментальных данных при предсказании вторичной структуры РНК. Появление технологий высокопроизводительного секвенирования привело к значительному снижению стоимости секвенирования, что в свою очередь стало причиной накопления большого количества информации о геномах и транскриптомах живых организмов. Таким образом методы, позволяющие находить функциональные элементы на основе первичных данных и с учетом имеющихся на данный момент данных, полученных с помощью высокопроизводительных экспериментальных методов как никогда актуальны. С этой точки зрения работа Виноградовой Светланы Владимировны находится в русле текущих тенденций и является достаточно важной в практическом смысле.

Очевидно, что предсказание пространственной структуры молекулы РНК исходя только из последовательности составляющих ее нуклеотидов является в каком то смысле вырожденной задачей с множеством равновероятных решений. Таким образом использование дополнительной информации о структуре РНК может дать необходимый набор ограничений, который позволит повысить точность предсказания. Но тут нужно отметить, что существующие экспериментальные методики не дают информации достаточного качества, чтобы ввести строгие ограничения на структуру. Опираясь на них мы можем только с определенной долей уверенности говорить о том в каком состоянии находится тот или иной нуклеотид. Как мне кажется наибольшей заслугой автора данной работы является разработанный подход, позволяющий с достаточной гибкостью, учесть экспериментальные данные ряда достаточно широко распространенных методик

SHAPE-Seq и PARS при предсказании вторичной структуры РНК. Причем автором предложена унифицированная величина на основе оценки, которая в принципе позволит объединять результаты нескольких принципиально разных экспериментальных методик.

В ходе выполнения работы автором было показано на основе находящихся в открытом доступе данных, что вторичная структура РНК *in vivo* не обязательно находится в оптимальном с точки зрения термодинамики состоянии, которое позволяет ей выполнять свою функциональную роль. Было показано, что без использования экспериментальных данных данную особенность РНК не удастся учесть в полной мере.

Диссертационная работа Виноградовой С. В. Имеет традиционную структуру. Она состоит из обзора литературы, двух глав, содержащих методы и результаты, выводы и список использованной литературы.

В главе «Обзор литературы» автором приводятся основные сведения о типах РНК, вводятся понятия первичной, вторичной и пространственной структур РНК. Дается обзор основных методов предсказания вторичной структуры РНК, описываются экспериментальные методы изучения вторичной структуры ДНК – PARS, SHAPE и DMS.

Вторая глава «Свойства экспериментальных данных» посвящена описанию основных особенностей экспериментальных данных. Автором были проведен анализ данных, полученных с помощью экспериментальных методов SHAPE, PARS и ДМС. Показаны их недостатки, сделан вывод о невозможности непосредственного использования данных для внесения строгих ограничений на пространственную структуру молекулы РНК. На основе этого анализа автором вводится унифицированная для всех экспериментальных методик количественная мера, характеризующая структурные свойства нуклеотида на основе логарифма отношения правдоподобия.

В третьей главе «Поиск структурированных участков РНК» автором описывается подход по интеграции экспериментальных данных в алгоритм предсказания пространственной структуры РНК RNA Surface. Вводится понятие псевдосвободной энергии. На нескольких примерах показана эффективность работы модифицированной версии программы RNASurface.

Однако, диссертационная работа не лишена и ряда недостатков.

Для обозначения измеряемой в результате SHAPE эксперимента величины автор использует термин «реактивность», что не кажется достаточно удачным. Более приемлем был бы термин «реакционная способность».

В тексте пункта 1.3.1 приводится рис. 1.4 и указывается, что на нем показан пример расчета свободной энергии для структуры «стебель - петля». На нем приводятся значения энергий для различных участков структурного элемента, но нет никаких указаний каким образом они были получены.

Название пункта 1.5.2 «Полногеномные карты РНК» следовало бы назвать «Полнотранскриптомные карты РНК», так как анализ с использованием экспериментальных методов описанных в этом пункте обзора литературы проводится на уровне полного транскриптома, а не генома.

На мой взгляд в обзоре литературы недостаточно подробно рассказано о существующих на данный момент вычислительных методах предсказания вторичной структуры РНК с использованием экспериментальных данных. А именно на стр. 36 приводится формула иллюстрирующая подход, используемый Вэшитлом, но к сожалению, нет никаких объяснений величин, входящих в данную формулу. Как мне кажется этому вопросу следовало бы уделить больше внимания, так как работа посвящена как раз улучшению существующих методов именно в этой части.

В пункте 2.1.1 описывается два фильтра по покрытию используемый для отбора образцов с удовлетворительным качеством. Первый фильтр удаляет те транскрипты, которые покрывались в среднем меньше чем 1 ридом, второй же отбирал транскрипты с минимальным покрытием 10. Из текста не ясно зачем нужен первый фильтр, так как второй вносит более строгие ограничения.

На рис. 2.2 приведены распределения PARS значений для разных классов молекул РНК. Исходя из организации рисунка можно было бы предположить, что автор хотел, чтобы читатель смог сравнить их друг с другом, но к сожалению, они приведены в разных масштабах, что делает сравнение затруднительным. На вертикальной оси всех приведенных на данном рисунке графиков отсутствуют подписи.

На рис. 2.3 приводятся ROC кривые для оценки качества классификации нуклеотидов на основе PARS реактивностей с использованием разных

фиксированных порогов. Значения порогов и принципы их выбора в тексте диссертации не приводятся. То же самое относится к рис.2.5, и к выбору порогов для классификации на основе SHAPE реактивности.

На рис. 2.8 и для 5', и для 3' областей приведены по три графика, исходя из контекста соответствующих разным образцам, но информацию о них автор не приводит.

На рис. 2.9 а и б и рис. 2.10 а и б по вертикальной оси отложены значения log-сигнала, хотя нигде в тексте диссертации не вводится данное понятие. На рис. 2.9 с не указано, что отложено по вертикальной оси, а также в тексте диссертации не указано каким образом он был получен, указывается только что был проведен анализ периодичности, но какой именно из текста диссертации остается не ясно.

На графиках, изображенных на рис. 3.2 отсутствуют подписи на вертикальных осях.

В работе автор указывает на наличие зависимости псевдо-МСЭ от МСЭ и покрытия. Данное замечание наводит на мысль, что для экспериментальных методик, рассматриваемых в работе может иметь место зависимость покрытия от МСЭ, иными словами от пространственной структуры РНК. Определение такой зависимости было бы интересно с практической точки зрения и позволило бы в перспективе оценивать применимость того или иного экспериментального метода к разным классам РНК.

Название пункта 3.1.4 «Полногеномный поиск на основе экспериментальных данных» наводит на мысль об использовании разработанного программного обеспечения для поиска структурированных элементов непосредственно в геномной последовательности. Хотя из текста данного раздела скорее следует, что автором был произведен анализ части транскриптома.

Недостатки, приведённые в отзыве, затрудняют чтение и понимание работы. Однако, они никоим образом не ставят под сомнение ее высокий научный уровень и не снижают значимость полученных результатов. Личный вклад соискателя несомненен, а содержание отражено в двух публикациях по теме диссертации. Автореферат полноценно отражает содержание диссертации. Диссертационная работа Виноградовой Светланы Владимировны "Предсказание структурных элементов РНК с использованием экспериментальных данных" является

законченным научно-квалифицированным исследованием и соответствует всем требованиям "Положения о порядке присуждения ученых степеней", утвержденного Постановлением Правительства РФ от 24 сентября 2013 г. №842. Автор заслуживает присуждения ученой степени кандидата биологических наук по специальности 03.01.09 – “математическая биология, биоинформатика”.

кандидат физико-математических наук,
научный сотрудник
лаборатории системной биологии и вычислительной генетики
Федерального государственного бюджетного учреждения науки
Института общей генетики им. Н. И. Вавилова
Российской академии наук(ИОГен РАН)



/Касьянов А.С./

27 апреля 2016 года

Подпись
удостоверяю

Ученый секретарь ИОГен РАН
доктор биологических наук

Огаркова О.А.

