

**ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА Д 002.077.04  
НА БАЗЕ ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО  
УЧРЕЖДЕНИЯ НАУКИ ИНСТИТУТА ПРОБЛЕМ ПЕРЕДАЧИ  
ИНФОРМАЦИИ ИМ. А.А. ХАРКЕВИЧА РОССИЙСКОЙ АКАДЕМИИ НАУК  
ПО ДИССЕРТАЦИИ НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ КАНДИДАТА  
БИОЛОГИЧЕСКИХ НАУК**

Аттестационное дело № \_\_\_\_\_  
решение диссертационного совета  
от 12 октября 2020, протокол № 6

О присуждении Клиник Галине Викторовне,  
гражданке Российской Федерации,  
ученой степени кандидата биологических наук.

Диссертация «Расположение аминокислотных замен на эволюционном дереве как показатель изменчивости однопозиционного адаптивного ландшафта» по специальности 03.01.09 – Математическая биология, биоинформатика, принята к защите 16 марта 2020 года, протокол № 4, диссертационным советом Д 002.077.04 на базе Федерального государственного бюджетного учреждения науки Института проблем передачи информации им. А.А. Харкевича Российской академии наук (127051, г. Москва, Большой Каретный переулок, д. 19, стр. 1., приказ о создании № 978/нк от 16 декабря 2013 года).

Соискатель **Галина Викторовна Клиник**, гражданка Российской Федерации, 1990 года рождения, в 2012 году окончила биологический факультет Московского Государственного Университета им. М.В. Ломоносова, по специальности «Биохимия». В период подготовки диссертации, с 01.10.2013 по 01.10.2016 обучалась в аспирантуре на факультете биоинженерии и биоинформатики МГУ им. М.В. Ломоносова, а с 21.01.2015 работает в ИППИ РАН в должности и.о. младшего научного сотрудника.

**Диссертация выполнена** в лаборатории молекулярной эволюции Федерального государственного бюджетного учреждения науки Института проблем передачи информации им. А.А. Харкевича Российской академии наук.

**Научный руководитель – Георгий Александрович Базыкин**, доктор биологических наук, заведующий лабораторией молекулярной эволюции Федерального государственного бюджетного учреждения науки Института



проблем передачи информации им. А.А. Харкевича Российской академии наук.

**Официальные оппоненты:**

**Владимир Александрович Трифонов**, гражданин РФ, доктор биологических наук, заведующий лабораторией сравнительной геномики в Институте молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск,

**Инна Сергеевна Поволоцкая**, гражданка РФ, кандидат биологических наук, заведующая лабораторией клинической геномики и биоинформатики в Обособленном структурном подразделении "Научно-исследовательский клинический институт педиатрии имени академика Ю.Е.Вельтищева" ФГАОУ ВО РНИМУ им.Н.И.Пирогова,

**дали положительные отзывы на диссертацию.**

**Ведущая организация** – Федеральное бюджетное учреждение науки «Центральный НИИ эпидемиологии Роспотребнадзора» дала **положительное заключение**, подписанное кандидатом биологических наук Алексеем Дмитриевичем Неверовым, старшим научным сотрудником, руководителем научной группы биоинформатики ФБУН ЦНИИ Эпидемиологии Роспотребнадзора, и утвержденное академиком РАН, профессором Василием Геннадьевичем Акимкиным, директором ФБУН ЦНИИ Эпидемиологии Роспотребнадзора. В заключении содержится **ряд замечаний и вопросов:**

- На рисунке 14 патогенные аллели человека возникают независимо в достаточно далеких группах живых организмов, например, такие аллели ни разу не возникали у млекопитающих. На глаз, это не выглядит слишком убедительно, особенно по сравнению с примерами, представленными на рисунках 15 – 17. Существует ли какое-либо существенное систематическое различие, например, по степени радикальности замен или степени тяжести, вызванного у человека дефекта между примерами, представленными на рис. 14 и рис. 15-17?

- Какая доля АВП будет определена как аминокислоты в сайтах под направленным отбором с помощью DEPS [MBE 25(9): 1809-1824] или в сайтах под «переключающимся» отбором с помощью TOGGLE [PLoS Pathogens 4(12): e1000242] из пакета программ HYPHY?

- На рисунке 22в есть некоторое количество красных точек, расположенных существенно ниже пунктирной диагонали. Эти точки, по-видимому, соответствуют аминокислотам, приспособленность которых в экспериментах DMS существенно отличается от приспособленности in-vivo. Есть ли среди этих сайтов/аминокислот известные адаптации к культурам клеток?



- Стр. 35. «По нашим результатам, отношение П/Д достигало 1.7-2.5 на филогенетических расстояниях меньше 0.1, быстро падая до 1.0 с увеличением расстояния (Рисунок 8, Приложение 2)» - на больших филогенетических расстояниях на рисунке 8 для многих белков, например, COX1, CYTB, ND1 и др., П/Д убывает до значений 0.6 – 0.7, примерно с той же скоростью. Не очень понятно, о чем идет речь в утверждении автора.

- Стр. 44. «В целом, скорость изменения ОПАЛов в наших данных сопоставима со скоростью аминокислотных замен (Рисунок 8), что согласуется с [61]» - как из рисунка 8 следует, что скорость изменения ОПАЛов сопоставима со скоростью аминокислотных замен?

- Стр. 65. «Аминокислоту считали «потенциально значимой» (проксимальной или дистальной), если минимальная достижимая значимость (то есть, доля нулевых подвыборок, попавших в крайнюю точку нулевого распределения) была меньше  $\alpha$ . Мы определяли ожидаемую долю ложноположительных результатов (false discovery rate, FDR) как  $\text{argmin}((\alpha \times \text{PotVar}) / \text{VarFit}, 1)$ , где  $\alpha$  – уровень значимости, PotVar – число потенциально значимых аминокислот для данного  $\alpha$ , VarFit – число значимых аминокислот для данного  $\alpha$ » - определение FDR написано непонятно. Хотелось подчеркнуть, что PotVar – не должен зависеть от альфа, если под альфа понимать пороговое значение p-value.  $\text{FDR} = E[V/R]$  – где V/R - доля ложноположительных тестов относительно всех тестов объявленных положительными или значимыми. Тогда  $\text{FDR} \leq P_k * m/k$ , где m – число аминокислот, для которых применялся тест, k – число тестов, которые объявлены значимыми,  $P_k$  – p-value для теста k ( $k < m$ ), тогда выполняется  $E[V] \leq P_k * m$ .

Несмотря на замечания, в отзыве содержится высокая оценка научного уровня выполненной работы и отмечено «Диссертационная работа Галины Викторовны Клинк на тему: «Расположение аминокислотных замен на эволюционном дереве как показатель изменчивости однопозиционного адаптивного ландшафта», представленная на соискание ученой степени кандидата биологических наук, является законченной научно-квалификационной работой, в которой содержится решение научной задачи – определение влияния изменения однопозиционных адаптивных ландшафтов на эволюцию белков, имеющей важное значение для эволюционной биологии и биоинформатики».

**Соискатель имеет 6 статей в российских и международных рецензируемых журналах, включенных в перечень ВАК, из них 2 статьи по теме диссертации, общим объемом 30 страниц. Кроме того, соискателем опубликовано 5 тезисов**



по теме диссертации в материалах российских и международных конференций.

**Статьи по теме диссертации:**

1) **Klink G. V., Bazykin G. A.** Parallel evolution of metazoan mitochondrial proteins //Genome biology and evolution. – 2017. - P. 1341-1350 (импакт-фактор журнала: 3.726)

2) **Klink G. V., Golovin A. V., Bazykin G. A.** Substitutions into amino acids that are pathogenic in human mitochondrial proteins are more frequent in lineages closely related to human than in distant lineages //PeerJ. – 2017. – Vol. 5. – P. e4143 (импакт-фактор журнала: 2.353)

**Вклад диссертанта в опубликованные работы по теме диссертации состоит в непосредственном планировании исследований, участии в формулировках задач, теоретической разработке и практической реализации методов, обработке и анализе данных.** В обеих журнальных публикациях диссертант является первым автором.

Полученные результаты могут быть использованы при чтении спецкурсов на биологических и биоинформатических факультетах ВУЗов, на курсах повышения квалификации биологов. Также, полученные автором результаты и отработанные методические подходы могут служить базисом для проведения последующих исследований по изучению роли изменчивости адаптивных ландшафтов в эволюции белков. Исследования такого рода проводятся в ведущих лабораториях по всему миру.

**На диссертацию поступило 2 отзыва на автореферат, все отзывы положительные.** В отзывах указывается, что не вызывает сомнений актуальность темы рассматриваемой диссертации, а также научная и практическая ценность работы. Диссертация является законченной квалификационной работой, выполненной автором самостоятельно на высоком научном уровне.

Выбор официальных оппонентов обосновывается тем, что за последние годы ими было опубликовано большое количество научных работ по сравнительной и эволюционной геномике, в том числе по темам, относящимся к тематике диссертации.

**В отзыве оппонента Владимира Александровича Трифонова, доктора биологических наук, заведующего лабораторией сравнительной геномики в Институте молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, в целом давшего работе высокую оценку и рекомендовавшего диссертацию к защите, отмечены недочеты:**



- 1) На стр. 5 «Число потомков, оставляемых особью – ее приспособленность» - приспособленность может быть определена шире и включать не только индивидуальную, но и совокупную приспособленность, обусловленную групповым отбором.
- 2) Возможно, текст бы сильно упростило введение понятия «аминокислотного сайта», или «варианта аминокислотного остатка в сайте», тогда фразы вроде «поиска аминокислот, приспособленность которых неодинакова у разных видов» или «наблюдать патогенную для человека аминокислоту» звучали бы менее двусмысленно.
- 3) Стр. 8 «Кластеризация на филогении» - лучше говорить про кластеризацию на филогенетическом древе.
- 4) Стр. 18. «В связи с развитием методов высокопроизводительного секвенирования, такие данные сейчас можно найти для некоторых белков, например, белков клеточных органелл». Стоило бы уточнить, что в данной работе исследованы только белки органелл, закодированные в митохондриальном геноме, так как некоторые белки органелл кодируются в ядре. Кроме того, выбор этих белков можно обосновать не только их широкой доступностью для большого числа организмов (ядерных уже тоже немало), но и отсутствием паралогов в результате дупликаций генов и геномов.
- 5) Опечатки в видовых названиях: на стр. 21 «*Aequorea 21erevisi*» - должно быть «*A. victoria*», на стр. 22 «*Saccharomyces 22erevisiae*» - должно быть «*S. cerevisiae*».
- 6) Стр. 32 «хордовых, нехордовых и грибов.» Нехордовые не являются монофилетическим таксоном. Не очень понятно, зачем для парафилетической группы строить матрицу характерных замен.
- 7) Стр. 39 «это означает, что наш результат не вызван структурой генетического кода», тут, скорее всего, идет речь о предпочтении кодонов, а не о структуре генетического кода.
- 8) В главе 2 даже в названии главы имеется противопоставление двух таксонов - Metazoa и Opisthokonta, чего следует избегать, поскольку первый таксон входит в состав второго. Кроме того, латинские названия таксонов высокого порядка не следует выделять курсивом.
- 9) Для глав 2 и 3 хотелось бы добавить в обсуждение, как полученные модели варибельности ОПАЛ могут быть перенесены на ядерные белки, ведь они составляют значительно большую долю функционального генома эукариот.



10) В главе 3 автор справедливо замечает, что мутации митохондриальных белков могут иметь разную степень патогенности, но стоило бы отметить, что эта патогенность может развиваться с возрастом и таким образом не сильно влиять на приспособленность, а также может быть по-разному выражена в зависимости от контекста ядерного генома.

11) Предложение в выводе 3 «Список таких аминокислот для поверхностного белка из подтипов А, В, С ВИЧ-1 представлен в виде веб-сервера» несет информацию о полезном прикладном результате работы, но не является научным выводом.

Несмотря на сделанные замечания, в отзыве говорится, что диссертационная работа Клинк Галины Викторовны «выполнена на высоком научном и методическом уровне, соответствует критериям, предъявляемым к кандидатским диссертациям, и соответствует специальности 03.01.09 – математическая биология, биоинформатика. Автореферат достаточно полно отражает объем и содержание работы. По актуальности, новизне, теоретической и практической значимости, обоснованности научных положений и качеству научных публикаций представленная работа соответствует критериям пп. 9-14 «Положения о присуждении ученых степеней», утвержденного Правительством РФ, а автор диссертации, Клинк Галина Викторовна, заслуживает присуждения ученой степени кандидата биологических наук по специальности **03.01.09 – математическая биология, биоинформатика**».

**В отзыве оппонента Инны Сергеевны Поволоцкой**, кандидата биологических наук, заведующей лабораторией клинической геномики и биоинформатики в Обособленном структурном подразделении "Научно-исследовательский клинический институт педиатрии имени академика Ю.Е.Вельтищева" ФГАОУ ВО РНИМУ им.Н.И.Пирогова, в целом дана высокая оценка работе, но **имеется ряд замечаний и вопросов:**

- 1) В главе 1 не ясна роль в анализе выравниваний митохондриальных белков из статьи Breen, et al, 2012 и необходимость их использования.
- 2) В главе 1 (Рисунок 9) наблюдается небольшой избыток числа параллельных к числу дивергентных замен на близких филогенетических расстояниях в данных симуляции, что, безусловно, подозрительно. Не очевидно, были ли данные симуляции также скорректированы относительно структуры генетического кода.
- 3) В главе 2 в анализе митохондриальных патогенных вариантов присутствует существенный изъян. Тогда как роль эпистаза и изменчивости ландшафта приспособленности в эволюции белковых последовательностей, безусловно,



значительна, нельзя преуменьшать также роль постоянного отбора. Выводы базируются на тех позициях в вырывании, в которых присутствуют хотя бы по одной параллельной и дивергентной замене, тем самым производится допущение, что патогенная аминокислота, наблюдаемая у человека, должна иметь достаточно высокий фитнес хотя бы в одном из других видов, что, вероятнее всего, неправда для подавляющего числа патогенных вариантов с менделевским типом наследования, о чем свидетельствует также число исследуемых патогенных вариантов (6), тогда как по состоянию на сентябрь 2020 года в курируемой базе патогенных вариантов MITOMAP присутствует 40 патогенных аминокислот с минимум 2 ссылками, все из которых связаны с менделевскими наследственными заболеваниями. Минимум два из описанных в главе 2 шести патогенных вариантов связаны же с многофакторными заболеваниями.

Встает вопрос о верности определения патогенности исследуемых вариантов, и на настоящий момент описанные в главе 2 варианты MT-ND1 V113A, MT-CO3 V91A и MT-ND6 I33V находятся в базе данных доброкачественных полиморфизмов MITOMAP, а также классифицированы доброкачественными в открытой базе данных патогенных вариантов clinvar, что согласуется с рекомендациями по определению патогенности вариантов ACMG.

Таким образом, весьма сомнительными представляются существование выборки истинно патогенных вариантов, удовлетворяющих заданным условиям (наличие минимум одной параллельной и одной дивергентной замены), и соответственно, выводы о динамике ландшафта приспособленности этого типа вариантов.

**В отзыве отмечено, что** «указанные недостатки не снижают ценности полученных результатов, и не носят принципиального характера. Работа выполнена на высоком научном уровне с использованием современных методов и подходов» **и что** «автор заслуживает присуждения искомой степени кандидата биологических наук по специальности 30.01.09 – «математическая биология, биоинформатика»».

**В отзыве на автореферат С.В.Денисова, к.б.н., научного сотрудника Манчестерского университета, отмечено, что** «Автореферат хорошо написан, задачи четко сформулированы, изложение происходит последовательно и логично».

**В целом дана положительная оценка, но есть вопросы и замечания:**

- 1) Из рис.3в следует, что замены на патогенные (у человека) аллели располагаются ближе к человеку на дереве, чем ожидается. Этот эффект усиливается, если рассматривать только подтверждённые патогенные аллели



(которые, кажется, должны быть еще более патогенными и следовательно неадаптивными). Почему так происходит?

2) На том же рис. 3в случайные замены (которые надо полагать использовались как контроль) происходят чаще, чем ожидается. Можно ли это как-то интерпретировать?

3) Почему только на рис. 3а (но не на 3б и 3в) показан результат симуляции?

4) На рис. 1в иллюстрируется идея о том, что “объединение информации из сайтов с разными свойствами также может приводить к избытку гомоплазий на коротких филогенетических расстояниях”. Без дополнительных пояснений, мне не ясно, что именно значат замены  $C>B$ ,  $B>A$ ,  $B>C$ ,  $B>D$ , и как их расположение иллюстрирует высказанную идею?

**В конце указывается, что «Работа выполнена на высоком методическом уровне, результаты не вызывают сомнений, выводы работы обоснованы» и «автор, Галина Викторовна Клинок, несомненно заслуживает присуждения ей ученой степени кандидата биологических наук».**

**В отзыве на автореферат Д.Н.Иванкова, кандидата физико-математических наук, Старшего преподавателя Центра наук о жизни Сколковского института науки и технологий, дана высокая оценка работы, указано, что «Автореферат написан аккуратно, опечаток практически нет» и что Г.В. Клинок «несомненно, заслуживает присвоения ей ученой степени кандидата биологических наук».**

**Диссертационный совет отмечает, что на основании выполненных соискателем исследований:**

**впервые показано** с помощью разработанного диссертантом метода значительное влияние изменчивости однопозиционных адаптивных ландшафтов на эволюцию белков, кодируемых в митохондриях;

**впервые показано, что** в митохондриально-кодируемых белках замены на известные патогенные для человека варианты в среднем возникают на более близких эволюционных расстояниях от человека, чем замены на аминокислоты, не встречающиеся в геномах людей в тех же позициях белков;

**разработан** не имеющий аналогов метод поиска отдельных аминокислотных вариантов, меняющих своё влияние на приспособленность в ходе эволюции. Метод **применён** к поверхностным белкам вируса иммунодефицита человека и вируса



гриппа. Для обоих белков **показано** высокое соответствие результатов метода с экспериментальными данными.

**Теоретическая значимость исследования обоснована тем, что:**

оно вносит вклад в понимание роли изменений однопозиционных адаптивных ландшафтов в эволюции белков, а также является важным шагом в изучении возможностей филогенетических методов в применении к изучению адаптивных ландшафтов. Разработанный метод поиска аминокислот, влияние которых на приспособленность меняется в ходе эволюции, позволяет изучать особенности изменения однопозиционных адаптивных ландшафтов в конкретных позициях белков.

**Значение полученных соискателем результатов исследования для практики подтверждается тем, что:**

Результаты, полученные автором для белков митохондрий, могут быть учтены при разработке методов определения эффектов мутаций в геноме человека. Описанный метод поиска аминокислотных сайтов с изменяющимися однопозиционными адаптивными ландшафтами может применяться при изучении эволюции вирусов для поиска функционально важных аминокислотных сайтов и особенностей однопозиционных адаптивных ландшафтов этих сайтов в разных штаммах. Такие сведения могут быть полезны при разработке противовирусных препаратов.

**Свидетельства достоверности результатов исследования:**

- результаты работы согласуются с существующими представлениями об эволюции белков и дополняют их;
- результаты, полученные методом, изложенным в главе 4, согласуются с экспериментальными данными.

**Личный вклад соискателя состоит в планировании исследований, участии в постановке задач, теоретической разработке и практической реализации методов, анализе данных и написании статей. Результаты, описанные в диссертации, получены автором самостоятельно. В диссертации решены задачи, имеющие значение для развития исследований, направленных на дальнейшее изучение изменчивости однопозиционных адаптивных ландшафтов. По своему содержанию диссертация отвечает паспорту специальности «03.01.09 – математическая биология, биоинформатика».**

Диссертационный совет пришёл к выводу о том, что диссертация представляет собой завершённое научное исследование. По актуальности, новизне, практической значимости диссертация соответствует требованиям, установленным «Положением о порядке присуждения ученых степеней», утверждённым постановлением



Правительства Российской Федерации от 24 сентября 2013 года № 842 (в редакции постановления Правительства Российской Федерации от 21 апреля 2016 г. № 335), предъявляемым к диссертациям на соискание ученой степени кандидата биологических наук.

На заседании 12 октября 2020 года диссертационный совет принял решение присудить Галине Викторовне Клинок ученую степень кандидата биологических наук по специальности 03.01.09 – математическая биология, биоинформатика.

При проведении открытого голосования диссертационный совет в количестве 14 человек, из них 9 докторов наук по специальности рассматриваемой диссертации, участвовавших в заседании, из 21 человека, входящих в состав совета, проголосовали за - 14, против – 0.

Председатель

диссертационного совета Д 002.077.04

д.б.н., профессор

Ученый секретарь

диссертационного совета Д 002.077.04

д.б.н., профессор

12 октября 2020 г.

 М.С. Гельфанд

 Г.И. Рожкова

