

Заключение диссертационного совета МГУ.015.10
по диссертации на соискание ученой степени кандидата наук

Решение диссертационного совета от 05.03.2025 года № 4

О присуждении Пензару Дмитрию Дмитриевичу, гражданину РФ, ученой степени кандидата биологических наук.

Диссертация «Вычислительное предсказание эффектов мутаций в регуляторных районах генов» по специальности 1.5.8. Математическая биология, биоинформатика принята к защите диссертационным советом 20.01.2025, протокол № 2.

Соискатель Пензар Дмитрий Дмитриевич, 1996 года рождения, в 2023 окончил аспирантуру Федерального государственного бюджетного учреждения науки Института общей генетики им. Н.И. Вавилова Российской академии наук (ИОГен РАН), срок обучения – 1.10.2019-30.09.2023 года.

Соискатель работал во время подготовки диссертации младшим научным сотрудником лаборатории системной биологии и вычислительной генетики Федерального государственного бюджетного учреждения науки Института общей генетики им. Н.И. Вавилова Российской академии наук (ИОГен РАН) с 17.05.2019 по 09.01.2024. С 30.09.2019 по настоящее время работает по совместительству преподавателем факультета биоинженерии и биоинформатики Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова».

Диссертация выполнена в лаборатории системной биологии и вычислительной генетики Федерального государственного бюджетного учреждения науки Института общей генетики им. Н.И. Вавилова Российской академии наук (ИОГен РАН).

Научный руководитель:

Кулаковский Иван Владимирович, доктор биологических наук, ведущий 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).

Официальные оппоненты:

1. Храмеева Екатерина Евгеньевна, доктор биологических наук, доцент Центра молекулярной и клеточной биологии Автономной некоммерческой образовательной организации высшего образования «Сколковский институт науки и технологий»,

2. Уткин Лев Владимирович, доктор технических наук, профессор, главный научный сотрудник, профессор Высшей школы технологий искусственного интеллекта Института компьютерных наук и кибербезопасности Федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский политехнический университет Петра Великого»,

3. Орлов Юрий Львович, доктор биологических наук, профессор РАН, профессор кафедры информационных технологий и обработки медицинских данных Центра цифровой медицины Института цифрового биодизайна и моделирования живых систем Федерального государственного автономного образовательного учреждения высшего образования Первый Московский государственный медицинский университет имени И.М.Сеченова Министерства здравоохранения Российской Федерации (Сеченовский Университет)

дали положительные отзывы на диссертацию.

Выбор официальных оппонентов обосновывался их высокой компетентностью в области биоинформатики, регуляторной геномики и машинного обучения, а также наличием большого количества публикаций в ведущих российских и зарубежных рецензируемых научных изданиях по тематике диссертации соискателя.

Соискатель имеет 16 опубликованных работ, в том числе по теме диссертации 6 работ, из них 6 статей, опубликованных в рецензируемых научных изданиях, рекомендованных для защиты в диссертационном совете МГУ по специальности 1.5.8. Математическая биология, биоинформатика.

Основные публикации по теме диссертации (в скобках приведен объем публикации и вклад автора в печатных листах):

1. Agarwal V., Inoue F., Schubach M., Penzar D., Martin B.K., Dash P.M., Keukeleire P., Zhang Z., Sohota A., Zhao J., Georgakopoulos-Soares I., Noble W.S., Yardımcı G.G., Kulakovskiy I.V., Kircher M., Shendure J., Ahituv N. Massively parallel characterization of transcriptional regulatory elements // Nature.— Vol. 639, P. 411-420. doi: 10.1038/s41586-024-08430-9. JIF (для WoS) = 50.5, (2.75/0.25)
2. Rafi A.M., Nogina D., Penzar D., Lee D., Lee D., Kim N., Kim S., Kim D., Shin Y., Kwak I.-Y., Meshcheryakov G., Lando A., Zinkevich A., Kim B.-C., Lee J., Kang T., Vaishnav E.D., Yadollahpour P., Random Promoter DREAM Challenge Consortium, Kim S., Albrecht J., Regev A., Gong W., Kulakovskiy I.V., Meyer P., de Boer C.G. A

- community effort to optimize sequence-based deep learning models of gene regulation. // Nature Biotechnology – 2024. doi: 10.1038/s41587-024-02414-w. JIF (для WoS) = 33.1 (1.5/0.30)
3. Penzar D., Nogina D., Noskova E., Zinkevich A., Meshcheryakov G., Lando A., Rafi A.M., de Boer C., Kulakovskiy I.V. LegNet: a best-in-class deep learning model for short DNA regulatory regions // Bioinformatics.– 2023.– Vol. 39, № 8. doi: 10.1093/bioinformatics/btad457. JIF (для WoS) = 4.4 (0.95/0.45)
 4. Abramov S., Boytsov A., Bykova D., Penzar D., Yevshin I., Kolmykov S.K., Fridman M.V., Favorov A.V., Vorontsov I.E., Baulin E., Kolpakov F., Makeev V.J., Kulakovskiy I.V. Landscape of allele-specific transcription factor binding in the human genome // Nature Communications – 2021.– Vol. 12, № 1.– P. 2751. doi: 10.1038/s41467-021-23007-0. JIF (для WoS) = 14.7 (1.20/0.20)
 5. Ambrosini G., Vorontsov I., Penzar D., Groux R., Fornes O., Nikolaeva D.D., Ballester B., Grau J., Grosse I., Makeev V., Kulakovskiy I., Bucher P. Insights gained from a comprehensive all-against-all transcription factor binding motif benchmarking study // Genome Biology – Springer Science and Business Media LLC, 2020.– Vol. 21, № 1.– P. 114. doi: 10.1186/s13059-020-01996-3. JIF (для WoS) = 10.1, (1.12/0.15)
 6. Penzar D., Zinkevich A.O., Vorontsov I.E., Sitnik V.V., Favorov A.V., Makeev V.J., Kulakovskiy I.V. What Do Neighbors Tell About You: The Local Context of Cis-Regulatory Modules Complicates Prediction of Regulatory Variants // Frontiers in Genetics– 2019.– Vol. 10.– P. 1078. doi: 10.3389/fgene.2019.01078. JIF (для WoS) = 2.8, (0.70/0.40)

На автореферат диссертации поступило 3 дополнительных отзыва, все положительные.

Диссертационный совет отмечает, что представленная диссертация на соискание ученой степени кандидата биологических наук является научно-квалификационной работой, в которой на основании выполненных автором исследований содержится решение задач, имеющих значение для развития математической биологии, биоинформатики, в частности регуляторной геномики и вычислительной генетики.

Диссертация представляет собой самостоятельное законченное исследование, обладающее внутренним единством. Положения, выносимые на защиту, содержат новые научные результаты и свидетельствуют о личном вкладе автора в науку:

1. Данные параллельных репортерных экспериментов на основе мутагенеза насыщающей ПЦР в значительной степени отражают локальные зависимости в геномных сигналах. Это приводит к неоправданному завышению качества модели в случае использования простых традиционных разбиений доступной для обучения и тестирования моделей выборки данных. В задаче предсказания результатов насыщающего мутагенеза это приводит к невозможности использовать традиционные подходы для оценки реальной точности моделей.
2. Модель на основе случайного леса с использованием геномных признаков достигает приемлемого качества в задаче предсказания аллель-специфичного связывания факторов транскрипции в отдельных хорошо изученных типах клеток.
3. Сверточная нейронная сеть LegNet, обученная на задаче предсказания активности регуляторных последовательностей и их влияния на экспрессию репортерных генов, демонстрирует наилучшее качество среди всех моделей в независимом исследовании на промоторах дрожжей. LegNet переносится на другие типы данных, и на результатах МПРЭ, полученных в клетках человека, превосходит по точности предсказаний имеющиеся альтернативы.
4. Архитектура LegNet при помощи подхода “холодная диффузия” генерирует реалистичные регуляторные последовательности с заданной активностью.

На заседании 05.03.2025 года диссертационный совет принял решение присудить Пензару Дмитрию Дмитриевичу ученую степень кандидата биологических наук.

При проведении тайного голосования диссертационный совет в количестве 12 человек, из них докторов наук по специальности 1.5.8. Математическая биология, биоинформатика - 6, участвовавших в заседании, из 16 человек, входящих в состав совета, проголосовали: «за» – 12, «против» – 0, «недействительных бюллетеней» – 0.

Председатель совета, д.х.н., проф.

Швядас В. К.

Ученый секретарь совета, к.х.н.

Шаповалова И. В.

06.03.2025