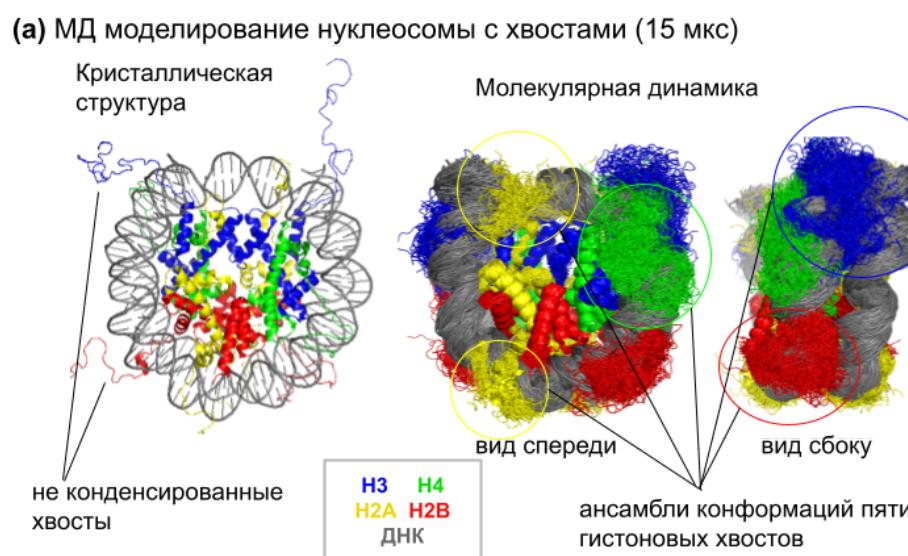


Дополнительные материалы к итоговому отчету

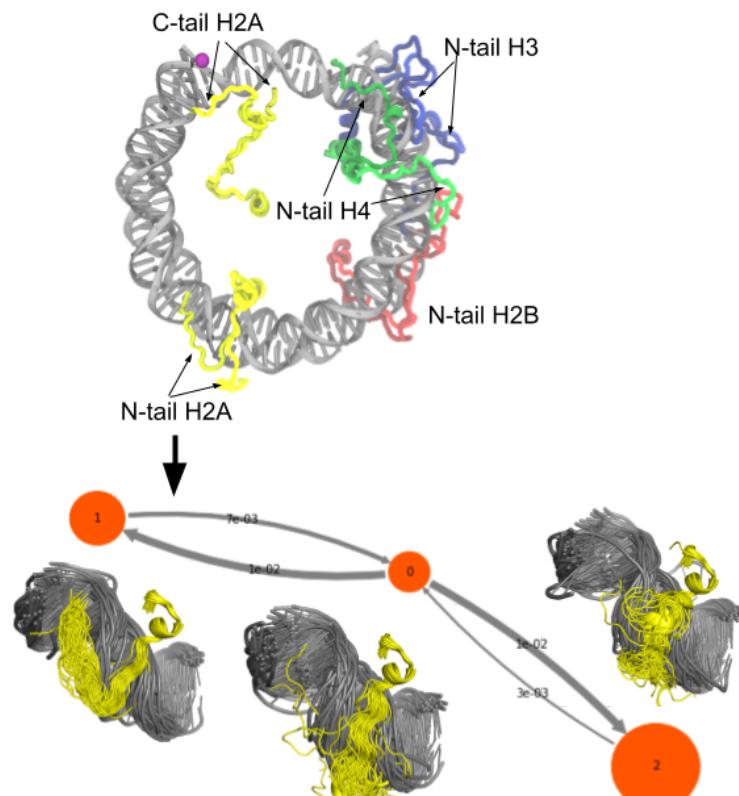
за 2021-2023 годы

по проекту 18-74-10006

“Структурная динамика нуклеосом и их взаимодействий: поиск подходов для диагностики и лечения онкологических заболеваний.”



(б) Кластеры конформаций хвостов и построение модели переходов между кластерами для N-хвоста гистона H2A



(в) МД моделирование нуклеосомы с модификациями

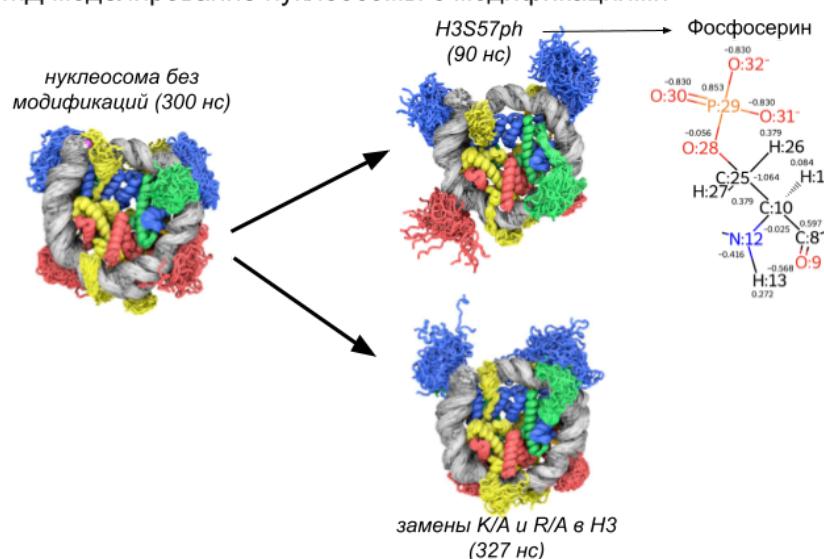
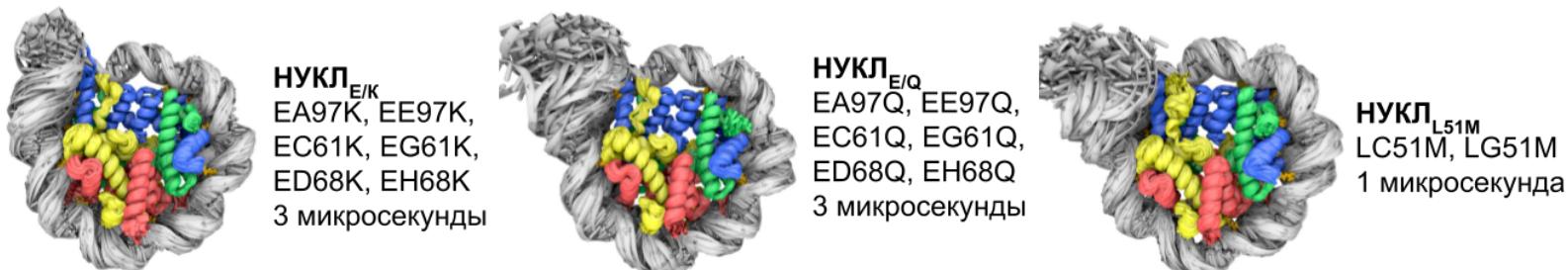
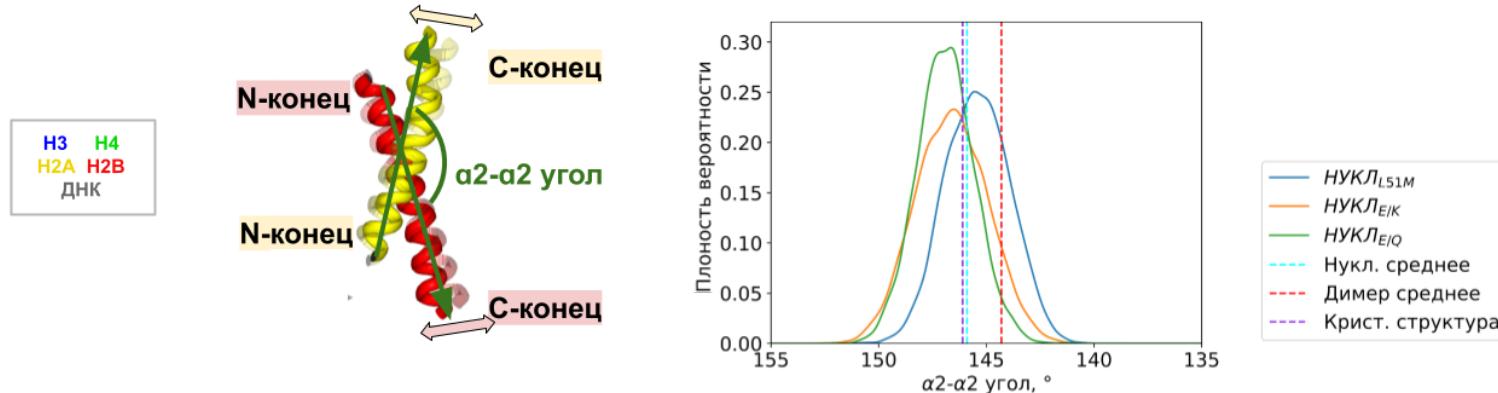


Рис. 1.1.1. Моделирование молекулярной динамики гистоновых хвостов в нуклеосомах и влияния модификаций на динамику нуклеосом. (а) Построение конформационного ансамбля гистоновых хвостов на основе 15-микросекундной траектории МД. (б) Выделенные кластеры конформаций хвостов на основе траектории их пункта (а). Построение модели переходов между подсостояниями хвостов с помощью марковских моделей в программе PyEmma. (в) МД нуклеосом с модификациями. В работе использовались два подхода: моделирование мимиков заряд-экранирующих ПТМ и параметризация нового остатка и интеграция его в силовое поле на примере фосфосерина.

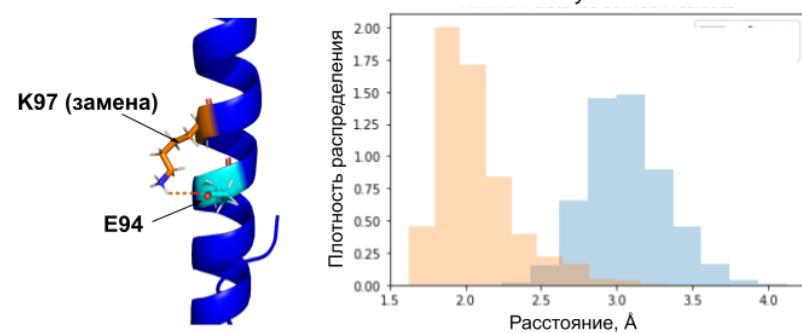
(а) Обзор МД траекторий нуклеосом с заменами, соответствующими онкомутиациям



(б) Сравнение распределений α_2 - α_2 угла H2A-H2B димера в нуклеосомах с заменами и канонических нуклеосомах и димерах



(в) Изменение взаимодействий боковых цепей в α_2 -спирали H3 в системе НУКЛ_{E/K} в сравнении с канонической нуклеосомой



(г) Изменение интерфейса взаимодействия α_2 -спиралей H2A и H2B в системе НУКЛ_{L51M}

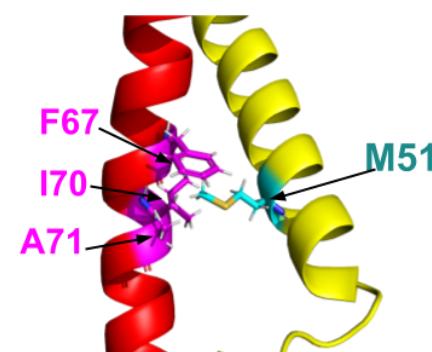


Рис. 1.2.1. Моделирование МД нуклеосом с заменами, соответствующими онкомутиациям, и изменения функционально важных мод пластиности гистонового кора нуклеосом. (а) Обзор траекторий МД систем нуклеосом с заменами в виде наложения кадров траектории. (б) Определение α_2 - α_2 угла димера гистонов H2A-H2B как характеристики динамики изгиба димера и его изменение при введении замен. (в-г) Молекулярные детали изменений, происходящих при введении замен.

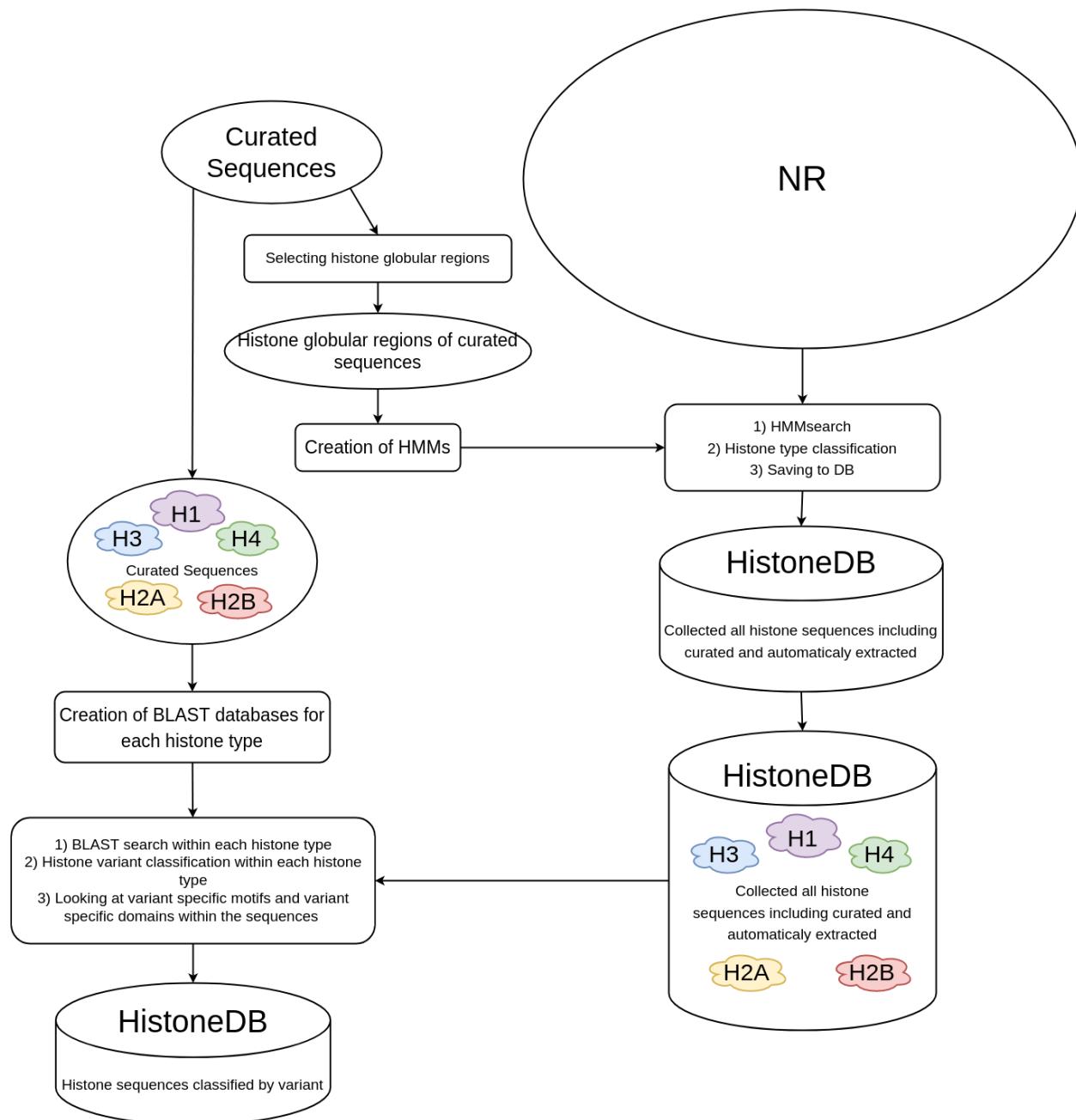


Рис. 2.1.1 Схема алгоритма классификации белков гистонов.

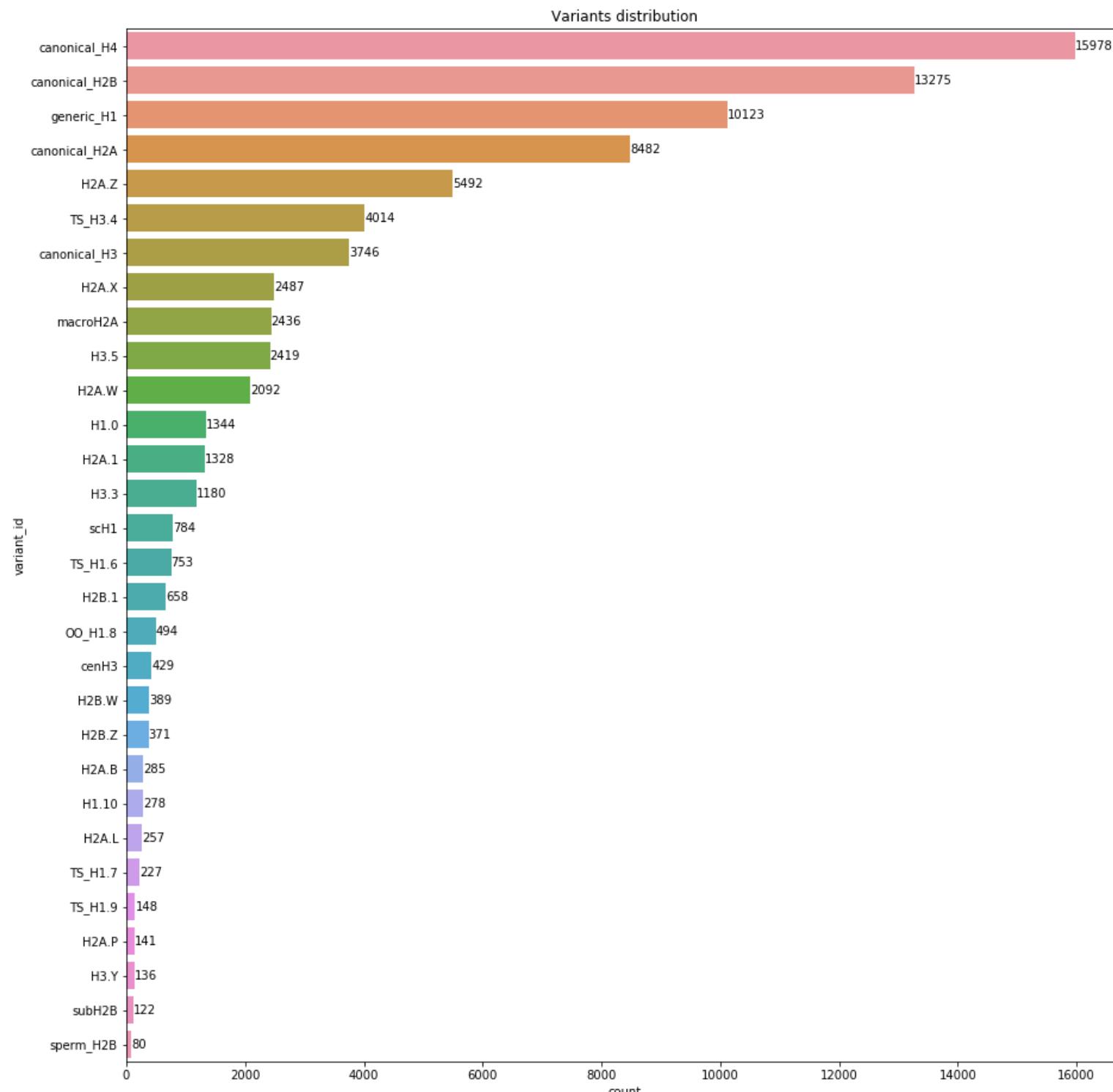


Рис. 2.1.2 Статистическое распределение последовательностей по гистоновым вариантам в старом алгоритме HistoneDB 2.0

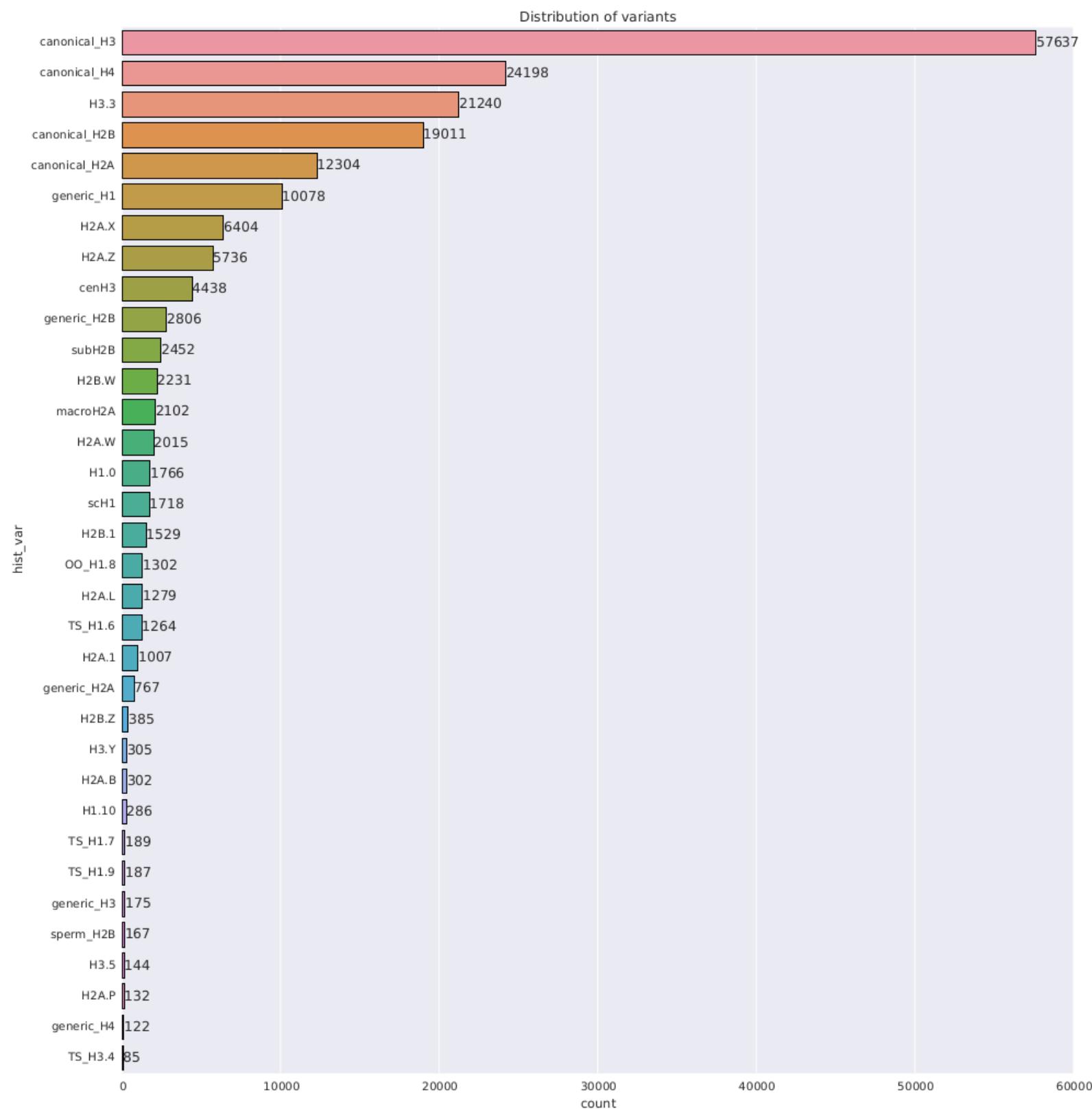


Рис. 2.1.3 Статистическое распределение последовательностей по гистоновым вариантам в обновленном алгоритме HistoneDB 3.0

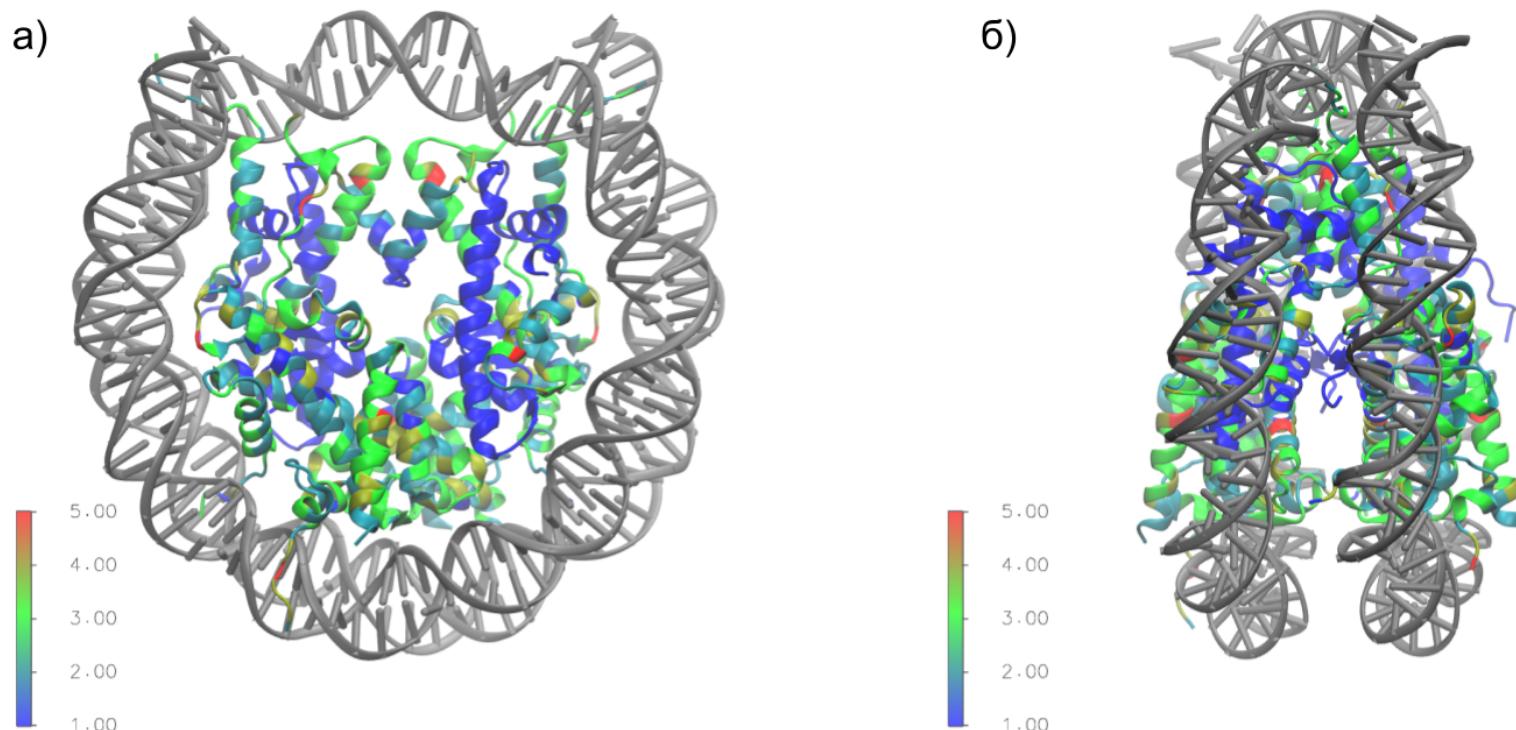


Рис. 2.2.1. Количественная оценка различных вариаций гистонов человека в структуре нуклеосомы. Цветовая шкала - число наблюдаемых вариаций аминокислотных замен (без учета делеций).

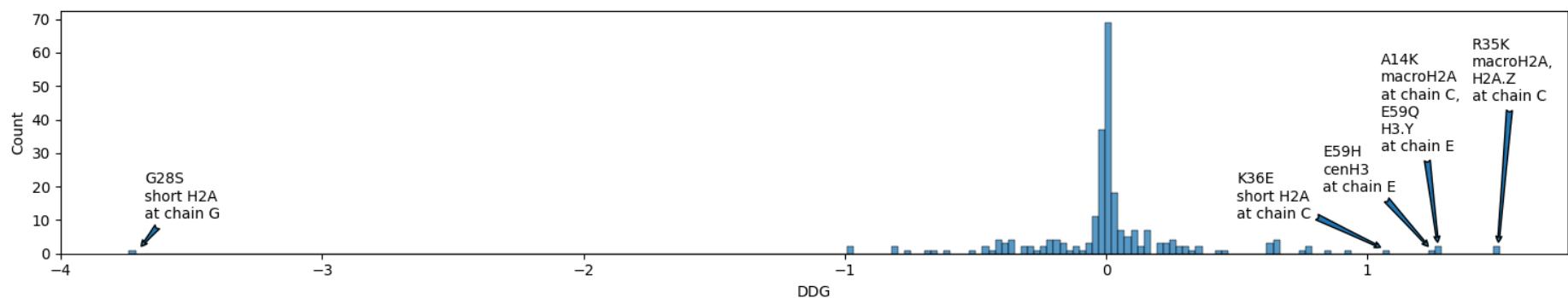


Рис. 2.2.2 Распределение параметра DDG (Delta Delta G), рассчитанного для вариаций в аминокислотных последовательностях гистонов человека.

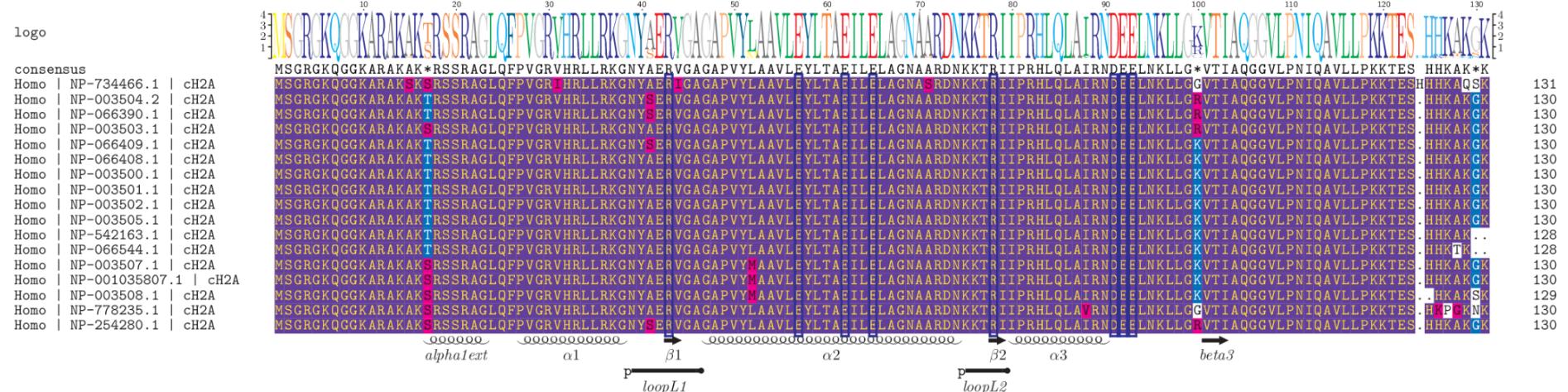


Рис. 2.2.3 Множественное выравнивание аминокислотных последовательностей канонических гистонов H2A. Фиолетовым цветом выделены аминокислоты, встречающиеся в данной позиции выравнивания более 80%, синим - более 50%. Розовым цветом отмечены похожие замены.

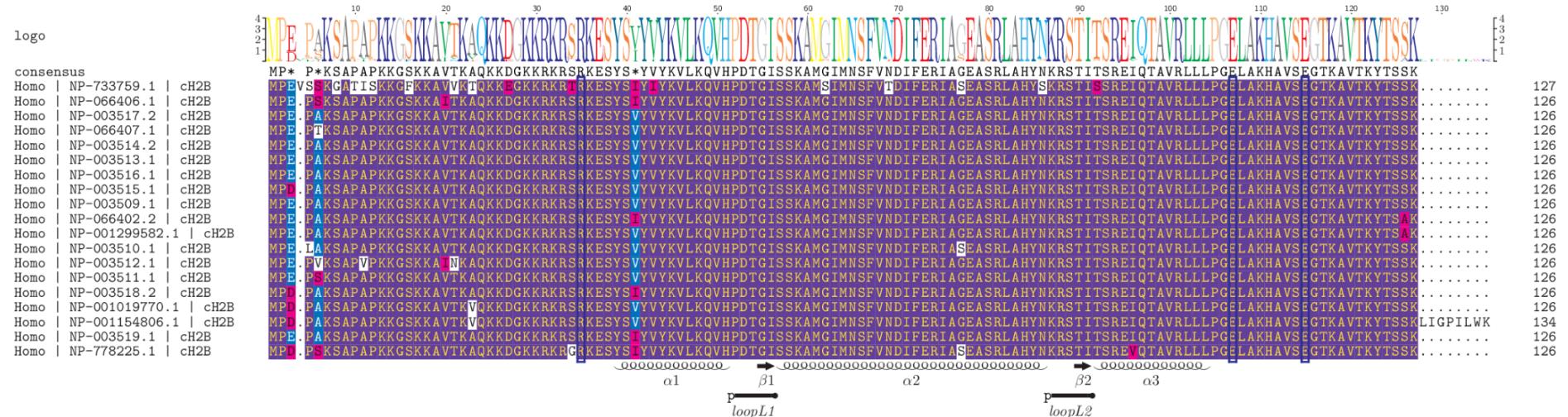


Рис. 2.2.4 Множественное выравнивание аминокислотных последовательностей канонических гистонов H2B. Фиолетовым цветом выделены аминокислоты, встречающиеся в данной позиции выравнивания более 80%, синим - более 50%. Розовым цветом отмечены похожие замены.

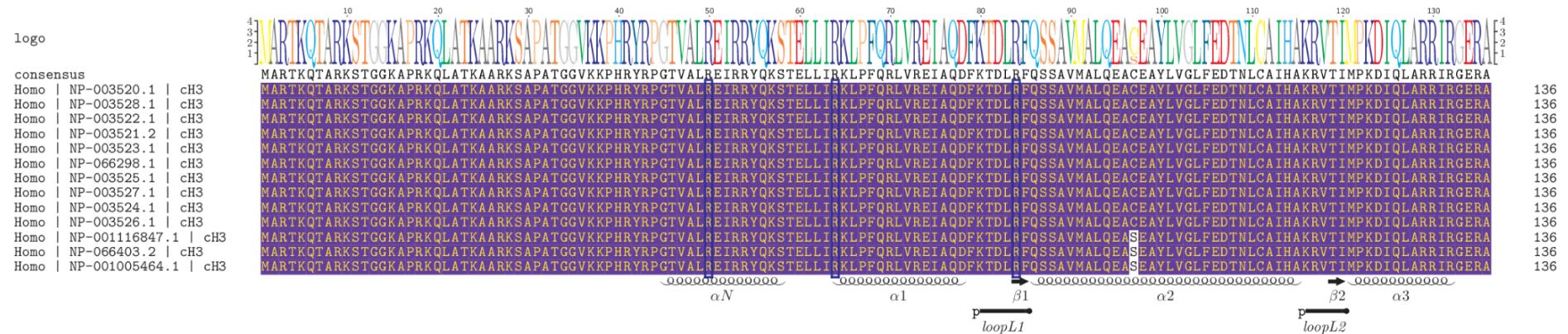


Рис. 2.2.5 Множественное выравнивание аминокислотных последовательностей канонических гистонов H3. Фиолетовым цветом выделены аминокислоты, встречающиеся в данной позиции выравнивания более 80%, синим - более 50%. Розовым цветом отмечены похожие замены.

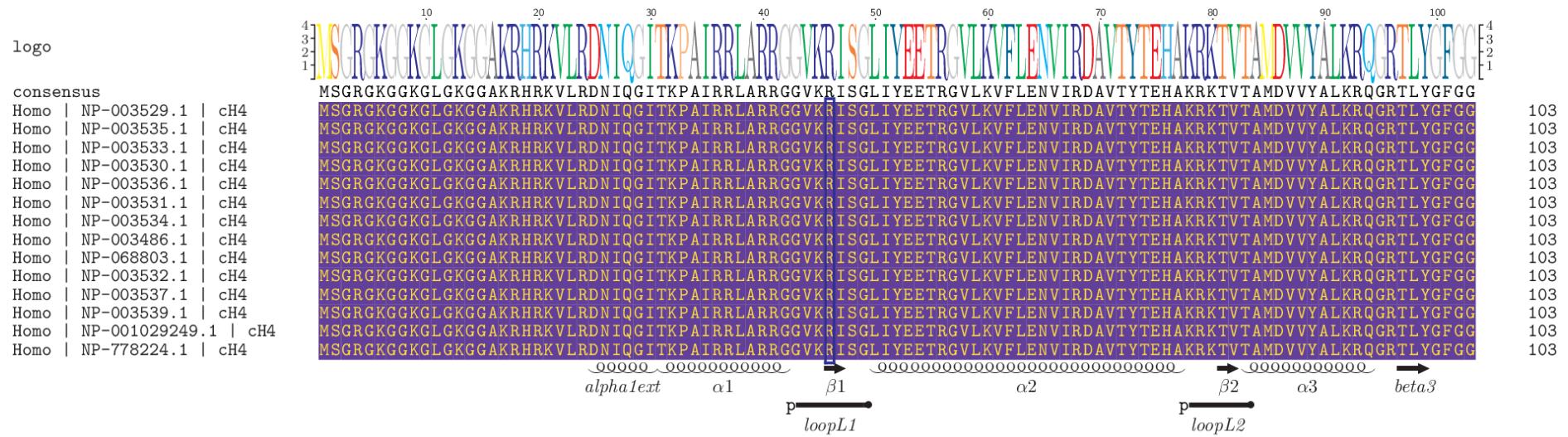


Рис. 2.2.6 Множественное выравнивание аминокислотных последовательностей канонических гистонов H4. Фиолетовым цветом выделены аминокислоты, встречающиеся в данной позиции выравнивания более 80%, синим - более 50%. Розовым цветом отмечены похожие замены.

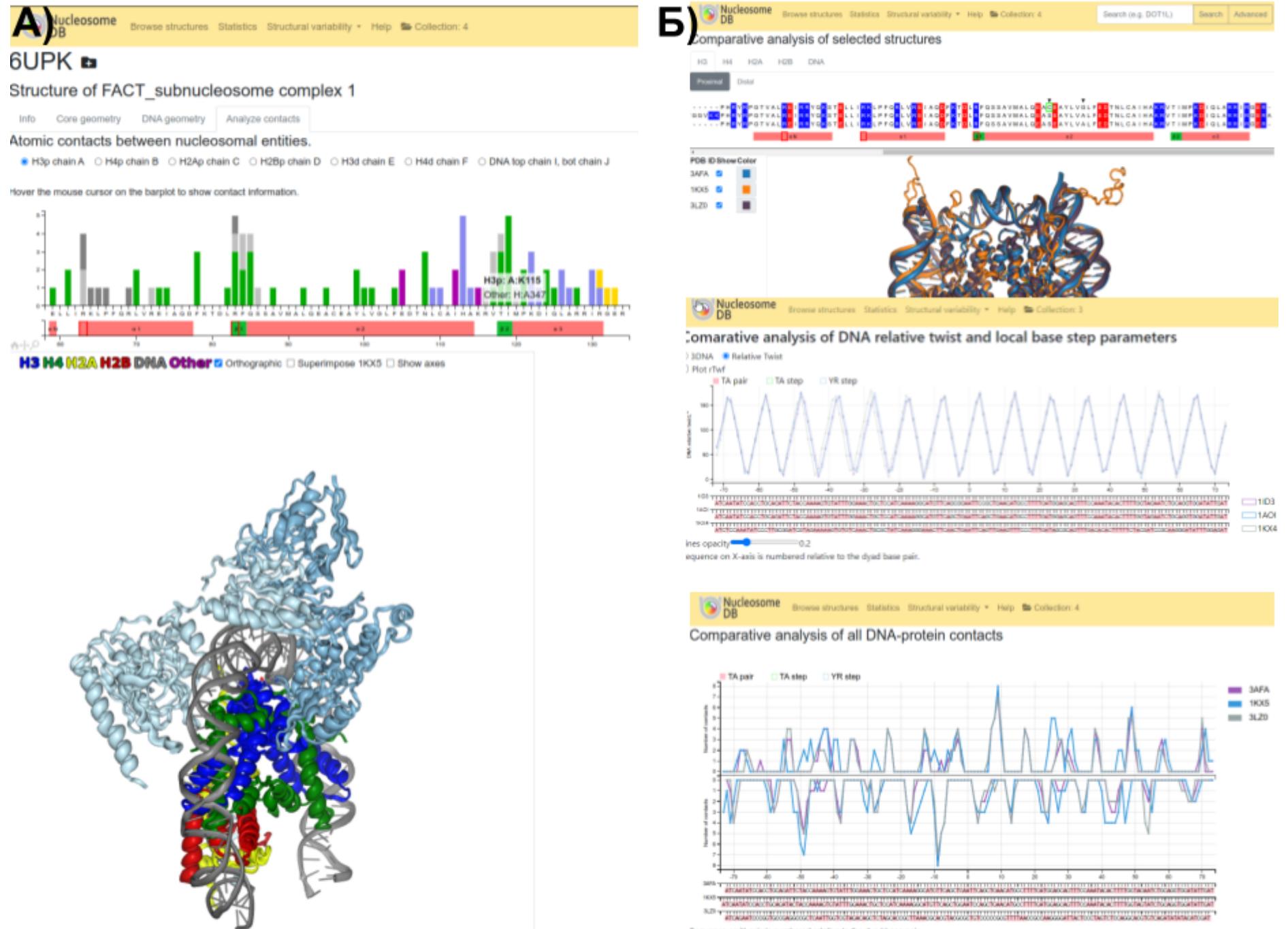


Рисунок 3.1. А) Инструментарий по аннотации контактов нуклеосом с интеракторами реализованный в сервисе NuclDB. Сервис позволяет выбирать контакты из профиля взаимодействий и визуализировать их на трехмерной структуре (пример <https://nucldb.intbio.org/structure/6UPK>, вкладка Analyze contacts) Б) Примеры возможности сравнительного анализа последовательностей, структур, геометрических характеристик и взаимодействий для разных структур нуклеосом.

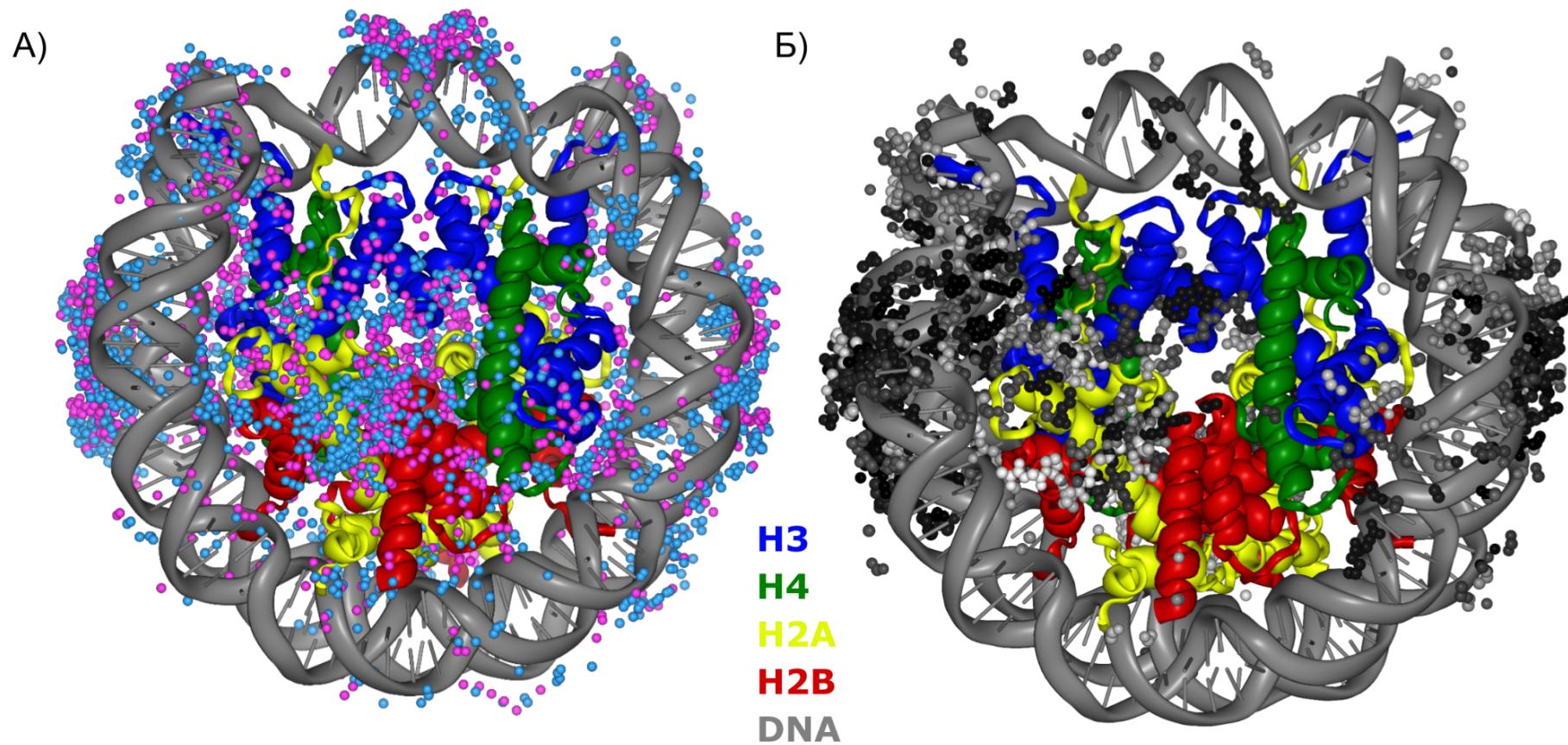


Рис. 3.2. Визуализация трехмерной статистической модели взаимодействия интеракторов с нуклеосомой. А) частота встречаемости заряженных атомов интеракторов вблизи поверхности нуклеосомы. Розовым показаны ориентально-заряженные атомы, голубым - положительно. Б) Визуализация вариативности аминокислот интеракторов, находящихся вблизи нуклеосомы. Цветовая шкала соответствует вариации энтропии Шеннона от малой (черный цвет) до большой (белый цвет).

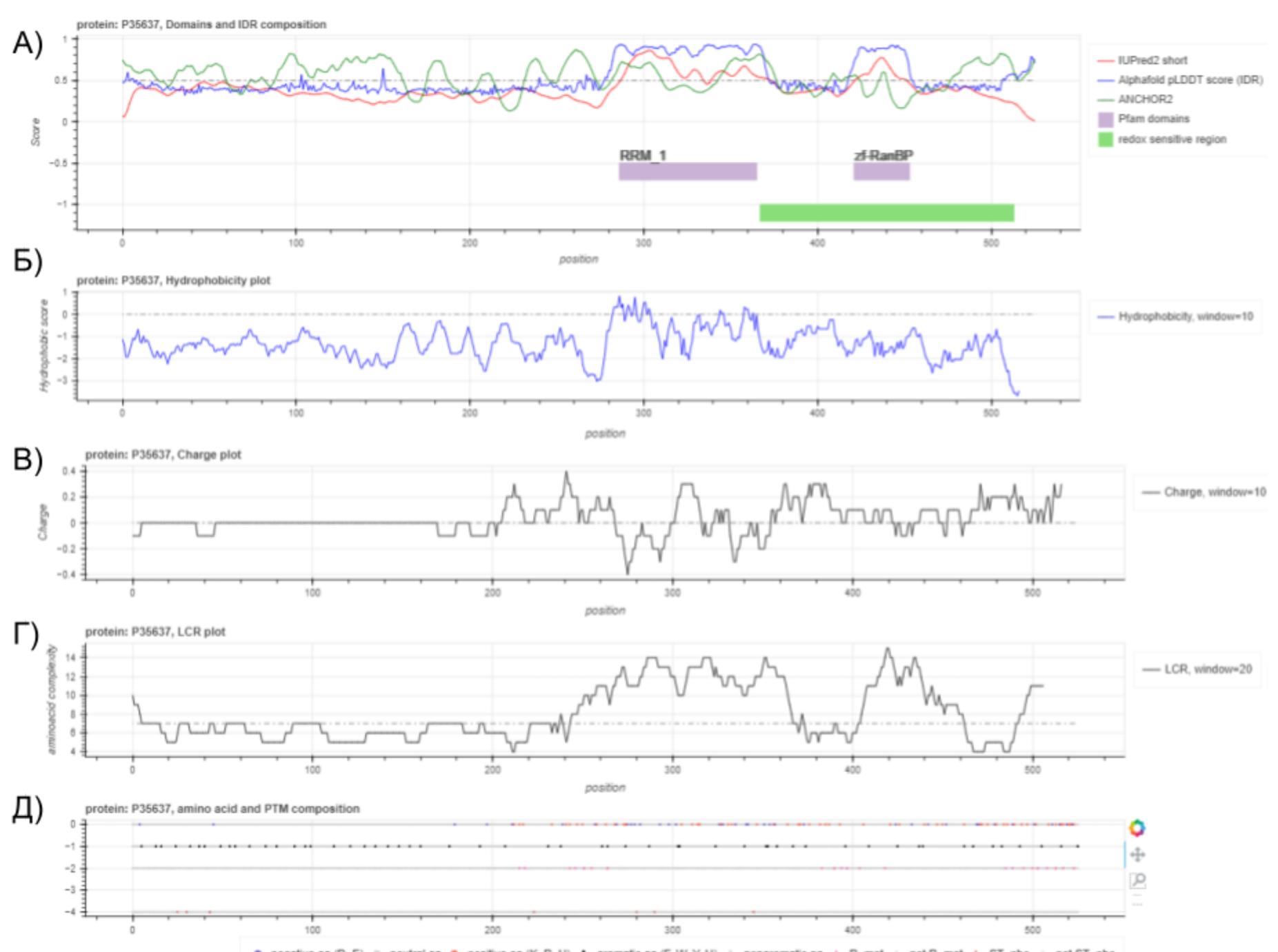


Рис. 4.1.1. Пример интерактивного графика для белка P35637 (FUS_HUMAN), А) доменная архитектура и неупорядоченные регионы, Б) индекс гидрофобности (усреднение по окну 10 а.о.) , В) общий заряд (усреднение по окну 10 а.о.), Г) регионы низкой сложности, Д) разметка заряженных и ароматических а.о., а также сайтов ПТМ.

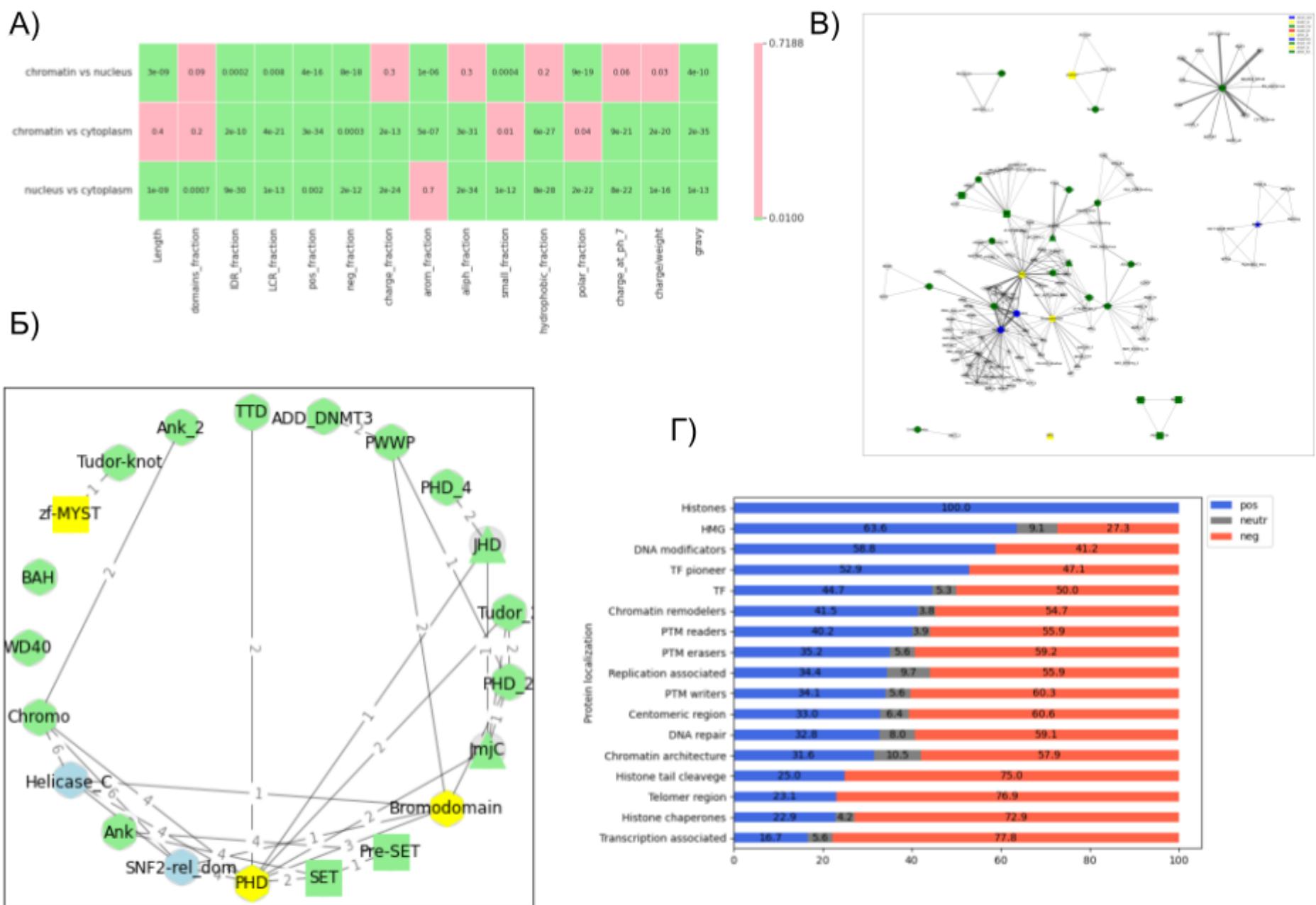
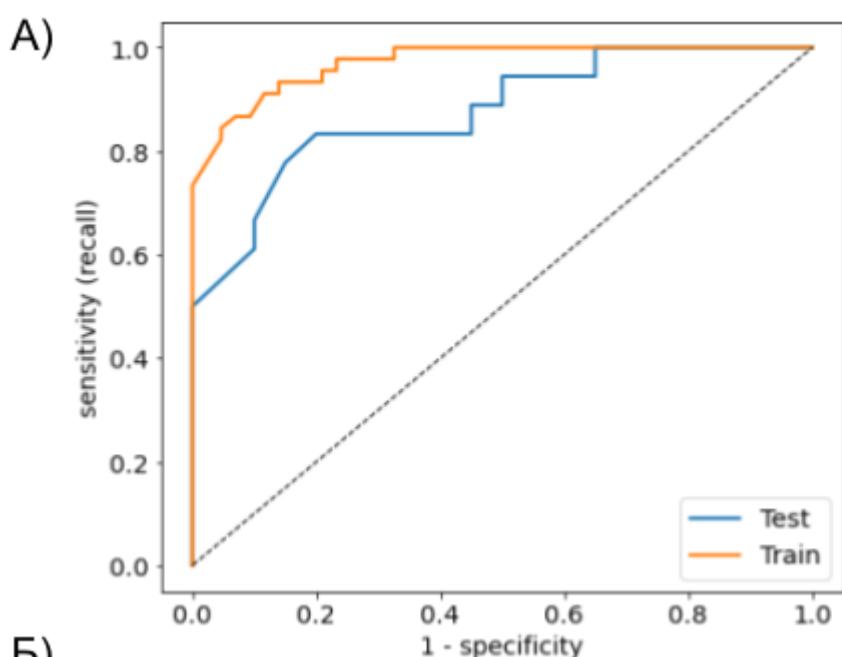
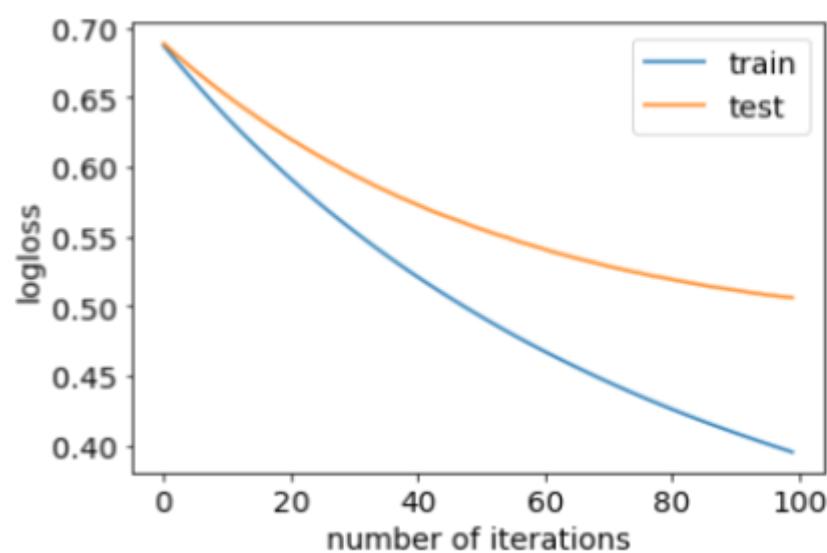


Рис. 4.1.2. Анализ физико-химических свойств и доменной архитектуры белков хроматина человека. А) Сводная таблица парных статистических тестов (критерий Манна-Уитни) по рассматриваемым признакам групп белков хроматина, других белков ядра и цитоплазмы (зеленый - стат.значимое значение p-value с поправкой на множественное тестирование, < 0.009). Б) Ковстречаемость доменов, считающих пост-трансляционные модификации гистонов в белках хроматина человека, значение на ребре графа - количество белков с данными доменами, легенда: зеленый - метильная метка, желтая - ацетильная, синий - хроматиновый ремоделлер, многоугольник - PTM reader, квадрат - PTM writer, треугольник - PTM eraser . В) Граф, построенный из всех доменов, встреченных в белках хроматина вместе с доменами, считающими пост-трансляционные модификации. Г) Фракции положительно (>1), отрицательно (<-1), и нейтрально заряженных белков в различных функциональных классах белков хроматина.



Б)



Б)

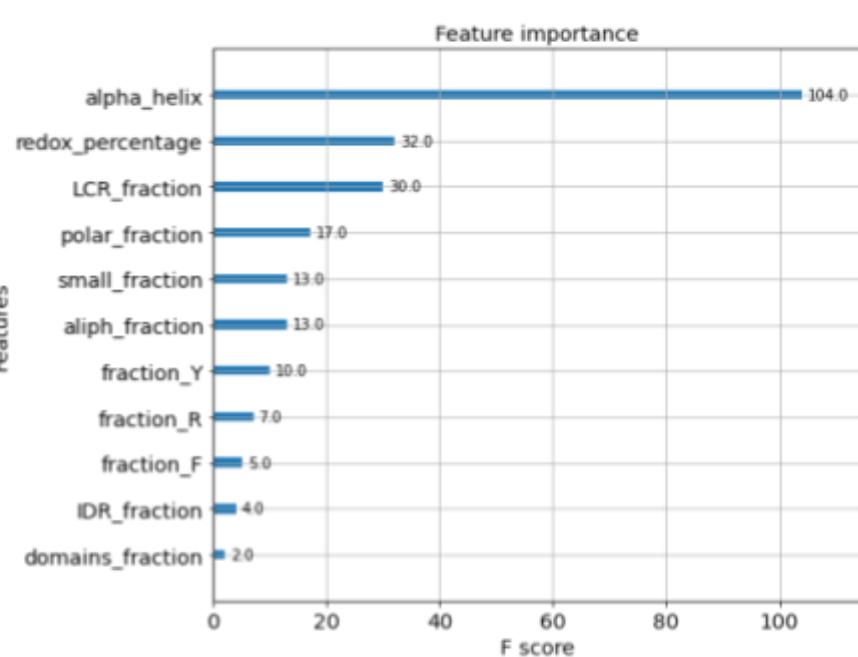


Рис. 4.1.3. Анализ построенного бинарного классификатора для выявления белков, разделяющих жидкые фазы. А) ROC-кривая на тренировочных тестовых данных, Б) функция потерь в зависимости от итерации для тренировочных и тестовых данных, В) график с важными для предсказания признаками.

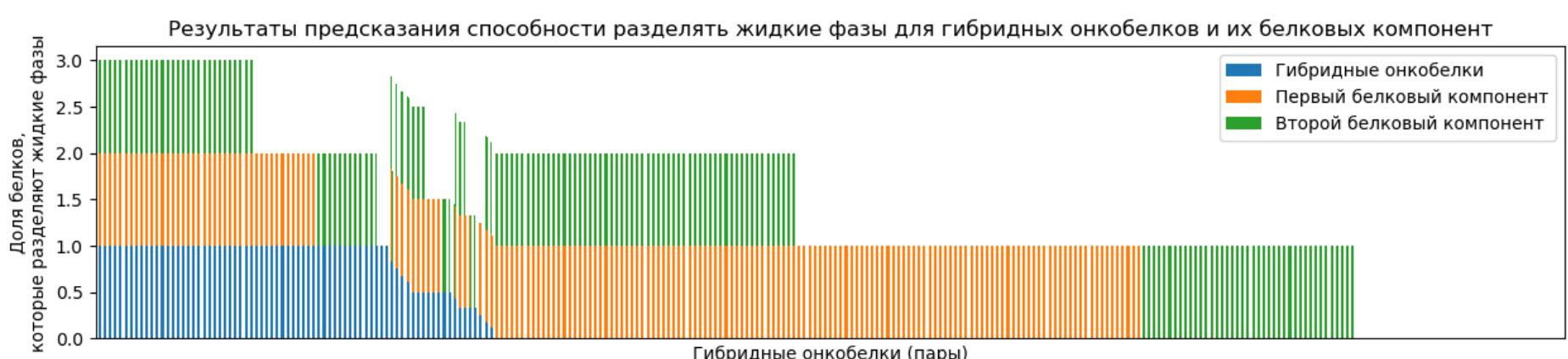


Рисунок 4.2.1. Анализ гибридных онкобелков и их белковых компонентов на предмет разделения фаз жидкость-жидкость. На графике представлены результаты предсказания классификатора машинного обучения на предмет разделения белками жидких фаз. Каждый столбик иллюстрирует пару белков, чьи гибридные белки и белковые компоненты были проанализированы (рыжий и зеленый столбик могут принимать значения 0 или 1).

==== Список цитированной литературы ===

===== Список литературы =====

- Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2: 19–25.
- Armeev GA, Kniazeva AS, Komarova GA, Kirpichnikov MP, Shaytan AK. 2021. Histone dynamics mediate DNA unwrapping and sliding in nucleosomes. *Nat. Commun.* 12.
- Barad BA, Echols N, Wang RY-R, Cheng Y, DiMaio F, Adams PD, Fraser JS. 2015. EMRinger: side chain–directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods* 12: 943–946.
- Bernhofer M, Goldberg T, Wolf S, Ahmed M, Zaugg J, Boden M, Rost B. 2018. NLSdb—major update for database of nuclear localization signals and nuclear export signals. *Nucleic Acids Res.* 46: D503–D508.
- Bhattacharya S, Reddy D, Jani V, Gadewal N, Shah S, Reddy R, Bose K, Sonavane U, Joshi R, Gupta S. 2017. Histone isoform H2A1H promotes attainment of distinct physiological states by altering chromatin dynamics. *Epigenetics Chromatin* 10: 1–19.
- Bowman GD, Poirier MG. 2015. Post-Translational Modifications of Histones That Influence Nucleosome Dynamics. *Chem. Rev.* 115: 2274–2295.

- Buß O, Rudat J, Ochsenreither K. 2018. FoldX as Protein Engineering Tool: Better Than Random Based Approaches? *Comput. Struct. Biotechnol. J.* 16: 25–33.
- Espiritu D, Gribkova AK, Gupta S, Shaytan AK, Panchenko AR. 2021. Molecular Mechanisms of Oncogenesis through the Lens of Nucleosomes and Histones. *J. Phys. Chem. B* 125: 3963–3976.
- Farahi N, Lazar T, Wodak SJ, Tompa P, Pancsa R. 2021. Integration of Data from Liquid–Liquid Phase Separation Databases Highlights Concentration and Dosage Sensitivity of LLPS Drivers. *Int. J. Mol. Sci.* 22: 3017.
- Franklin SG, Zweidler A. 1977. Non-allelic variants of histones 2a, 2b and 3 in mammals. *Nature* 266: 273–275.
- Ginno PA, Burger L, Seebacher J, Iesmantavicius V, Schübeler D. 2018. Cell cycle-resolved chromatin proteomics reveals the extent of mitotic preservation of the genomic regulatory landscape. *Nat. Commun.* 9: 4048.
- Huynh LM, Shinagawa T, Ishii S. 2016. Two Histone Variants TH2A and TH2B Enhance Human Induced Pluripotent Stem Cell Generation. *Stem Cells Dev.* 25: 251–258.
- Jumpur J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*: 1–7.
- Knizanova AS, Armeev GA, Shaytan AK. 2022. H2A-H2B Histone Dimer Plasticity and Its Functional Implications. *Cells* 11: 2837.
- Kustatscher G, Hégarat N, Wills KLH, Furlan C, Bukowski-Wills J-C, Hochegger H, Rappoport J. 2014. Proteomics of a fuzzy organelle: interphase chromatin. *EMBO J.* 33: 648–664.
- Li Q, Peng X, Li Y, Tang W, Zhu J, Huang J, Qi Y, Zhang Z. 2020. LLPSDB: a database of proteins undergoing liquid–liquid phase separation *in vitro*. *Nucleic Acids Res.* 48: D320–D327.
- Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. 2015. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 11: 3696–3713.
- Mészáros B, Erdős G, Szabó B, Schád É, Tantos Á, Abukhairan R, Horváth T, Murvai N, Kovács OP, Kovács M, Tosatto SCE, Tompa P, Dosztányi Z, Pancsa R. 2020. PhaSePro: the database of proteins driving liquid–liquid phase separation. *Nucleic Acids Res.* 48: D360–D367.
- van Mierlo G, Jansen JRG, Wang J, Poser I, van Heeringen SJ, Vermeulen M. 2021. Predicting protein condensate formation using machine learning. *Cell Rep.* 34: 108705.
- Millán-Zambrano G, Burton A, Bannister AJ, Schneider R. 2022. Histone post-translational modifications — cause and consequence of genome function. *Nat. Rev. Genet.*
- Nacev BA, Feng L, Bagert JD, Lemiesz AE, Gao J, Soshnev AA, Kundra R, Schultz N, Muir TW, Allis CD. 2019. The expanding landscape of “oncohistone” mutations in human cancers. *Nature* 567: 473–478.
- Ning W, Guo Y, Lin S, Mei B, Wu Y, Jiang P, Tan X, Zhang W, Chen G, Peng D, Chu L, Xue Y. 2020. DrLLPS: a data resource of liquid–liquid phase separation in eukaryotes. *Nucleic Acids Res.* 48: D288–D295.
- Saar KL, Morgunov AS, Qi R, Arter WE, Krainer G, Lee AA, Knowles TPJ. 2021. Learning the molecular grammar of protein condensates from sequence determinants and embeddings. *Proc. Natl. Acad. Sci.* 118.
- Scherer MK, Trendelkamp-Schroer B, Paul F, Pérez-Hernández G, Hoffmann M, Plattner N, Wehmeyer C, Prinz J-H, Noé F. 2015. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* 11: 5525–5542.
- Shi M, You K, Chen T, Hou C, Liang Z, Liu M, Wang J, Wei T, Qin J, Chen Y, Zhang MQ, Li T. 2021. Quantifying the phase separation property of chromatin-associated proteins under physiological conditions using an anti-1,6-hexanediol index. *Genome Biol.* 22: 229.
- Sun T, Li Q, Xu Y, Zhang Z, Lai L, Pei J. 2020. Prediction of Liquid-Liquid Phase Separation Proteins Using Machine Learning. Rochester, NY: Social Science Research Network.
- Thul PJ, Åkesson L, Wiklund M, Mahdessian D, Geladaki A, Ait Blal H, Alm T, Asplund A, Björk L, Breckels LM, Bäckström A, Danielsson F, Fagerberg L, Fall J, Gatto L, Gnann C, Hober S, Hjelmare M, Johansson F, Lee S, Lindskog C, Mulder J, Mulvey CM, Nilsson P, Oksvold P, Rockberg J, Schutten R, Schwenk JM, Sivertsson Å, Sjöstedt E, Skogs M, Stadler C, Sullivan DP, Tegel H, Winsnes C, Zhang C, Zwahlen M, Mardinoglu A, Pontén F, von Feilitzen K, Lilley KS, Uhlén M, Lundberg E. 2017. A subcellular map of the human proteome. *Science* 356: eaal3321.
- Voevodin VV, Antonov AS, Nikitenko DA, Shvets PA, Sobolev SI, Sidorov IY, Stefanov KS, Voevodin VV, Zhumatiy SA. 2019. Supercomputer Lomonosov-2: Large Scale, Deep Monitoring and Fine Analytics for the User Community. *Supercomput. Front. Innov.* 6: 4–11.
- Wang D, Liu D, Yuchi J, He F, Jiang Y, Cai S, Li J, Xu D. 2020. MusiteDeep: a deep-learning based webserver for protein post-translational modification site prediction and visualization. *Nucleic Acids Res.* 48: W140–W146.
- Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall III WB, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC. 2018. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* 27: 293–315.
- Yoo J, Aksimentiev A. 2018. New tricks for old dogs: improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions. *Phys. Chem. Chem. Phys.* 20: 8432–8449.
- You K, Huang Q, Yu C, Shen B, Sevilla C, Shi M, Hermjakob H, Chen Y, Li T. 2020. PhaSepDB: a database of liquid–liquid phase separation related proteins. *Nucleic Acids Res.* 48: D354–D359.