



Международная научная конференция
«Молекулярная генетика и биоинформатика»
7 апреля 2026 г., Минск

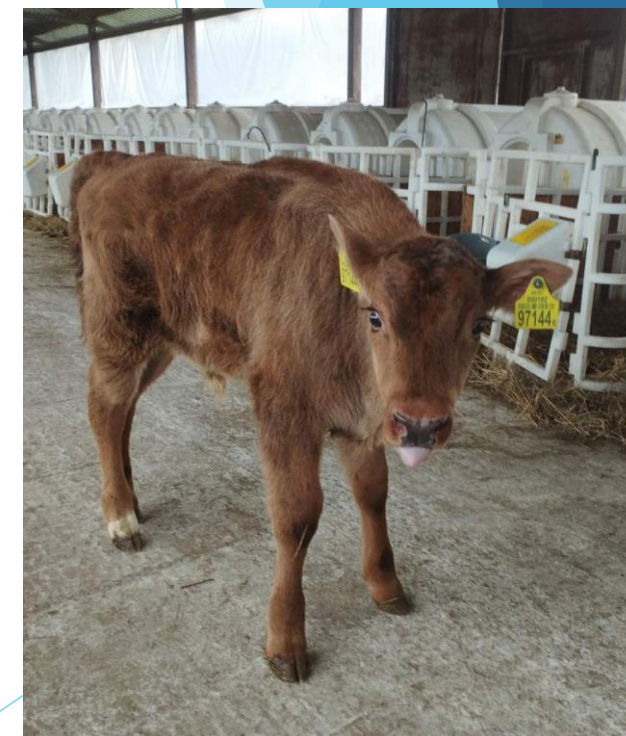
Биоинформатический анализ геномов видов
Bos taurus и *Bison bonasus*
для поиска дифференцирующих SNV

Кипень Вячеслав

к.б.н., доц., e-mail: v.kipen@igc.by

#Текущее состояние дел

- ▶ Методика установления видовой принадлежности животных (млекопитающих) Республики Беларусь, находящихся в сфере интереса, на основе технологии ДНК-баркодирования / А. С. Парфёнова [и др.] ; НПЦ Гос. ком. судеб. экспертиз Респ. Беларусь. - Минск, 2024. - 32 с.



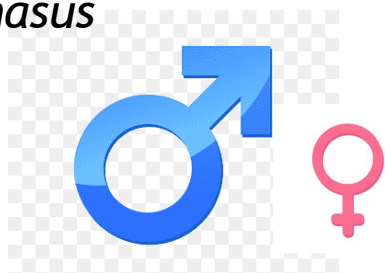
«Корова родила зубренка! Узнали подробности уникального для Беларуси случая»: [сайт]. – [Беларусь], 2025.
– URL: <https://www.sb.by/articles/ukhazhival-zubr-za-korovoy-.html> (дата обращения: 05.02.2026)

#Задача

- ▶ Выявление SNV с высоким дифференцирующим потенциалом для определения принадлежности образцов к видам *Bos taurus* или *Bison bonasus* на основании биоинформатического анализа геномов



Bison bonasus



Bos taurus



#Методология

▶ Образцы

- ▶ *Bison bonasus* - 28 геномов
- ▶ *Bos taurus* - 2 514 геномов
- ▶ 5 069 SNV из BovineSNV50 v3.0

▶ Определение генома in silico

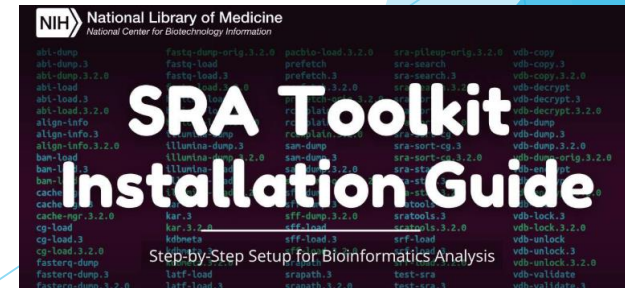
- ▶ Sequence Read Archive (SRA, NCBI)
- ▶ SRA-Toolkit v.2.11
- ▶ GENIS

▶ Статистический анализ данных

- ▶ ROC-анализ (Receiver Operating Characteristic analysis)
- ▶ MDR v.3.0.2
- ▶ Cohen's Kappa
- ▶ Jupyter Notebook (AUC, 95% ДИ)



GENIS



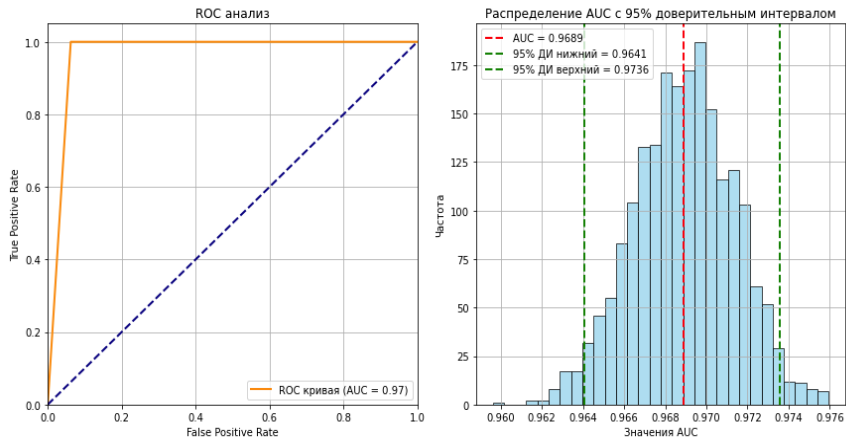
#Результат (1)

- ▶ Отбор наиболее информативных SNV состоял из двух этапов. На первом этапе были сформированы две тестовые выборки:
 - ▶ **Выборка № 1** – европейский зубр (n=28, BioSample ID - SAMEA114863253, SAMEA115388352, SAMEA115388353, SAMEA115388354, SAMEA115388355, SAMEA115388356, SAMEA115388357, SAMEA115675323, SAMEA5039932, SAMN05000810, SAMN05000811, SAMN05000812, SAMN05950802, SAMN05950803, SAMN05950804, SAMN05950805, SAMN05950806, SAMN05950807, SAMN05950808, SAMN08323717, SAMN08323719, SAMN08323720, SAMN08323722, SAMN08323723, SAMN08323724, SAMN08323725, SAMN08323726, SAMN08323727);
 - ▶ **Выборка № 2** – домашний бык (n=30, BioSample ID - SAMEA113983368, SAMEA113983376, SAMEA113983377, SAMEA113983384, SAMEA113983386, SAMEA113983387, SAMEA113983391, SAMEA113983392, SAMEA113983394, SAMEA113983396, SAMEA113983397, SAMEA113983398, SAMEA113983400, SAMEA113983401, SAMEA113983402, SAMEA113983403, SAMEA113983405, SAMEA113983408, SAMEA113983409, SAMEA113983410, SAMEA113983419, SAMEA113983421, SAMEA113983447, SAMEA113983451, SAMEA113983472, SAMEA113983477, SAMEA113983494, SAMEA113983498, SAMEA120322203, SAMEA120322218), - для которых определены генотипы *in silico* с использованием программы GENIS.
- ▶ Для первичной оценки информативности SNV (в сумме для 58 образцов) провели ROC-анализ, который выделил четыре полиморфных варианта со значениями AUC $\geq 0,9$:
rs41590141, rs110457178, rs41609042, rs110712082.

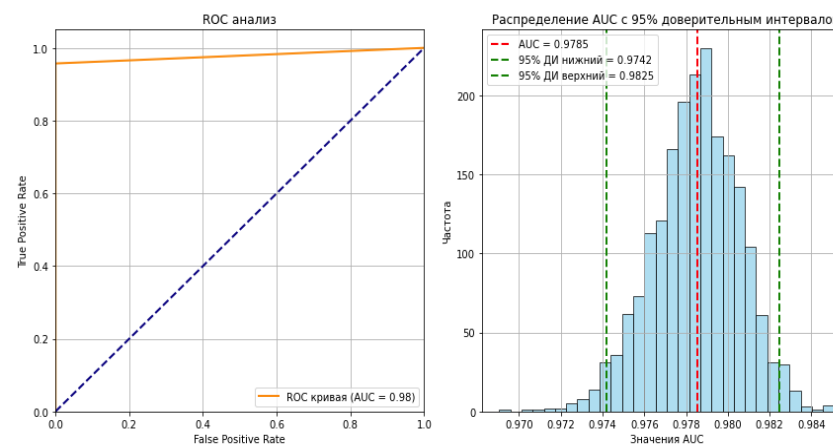
#Результат (2)

- ▶ На втором этапе исследования была расширена выборка с геномами домашнего быка до 2 514 шт., для которых были определены генотипы по четырем SNV, а также проведена итоговая оценка AUC

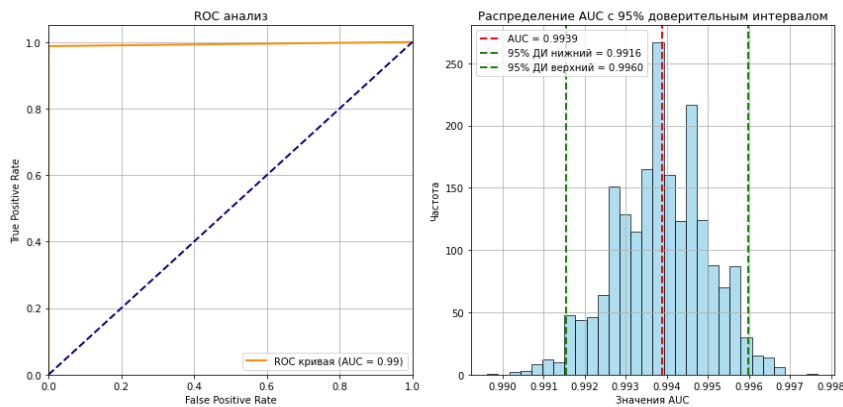
rs41590141



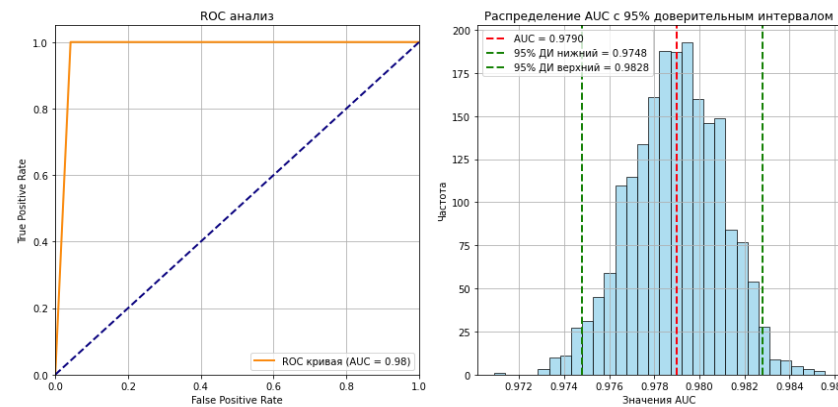
rs110457178



rs41609042



rs110712082



#Результат (3)

- ▶ Далее провели анализ многофакторного сокращения размерности (Multifactor Dimensionality Reduction, MDR) и определили четыре модели с высокими значениями показателей: сбалансированная точность (adj. Balanced accuracy), чувствительность (Sensitivity), специфичность (Specificity), воспроизводимость (CV Consistency), коэффициент Каппа (Cohen's Kappa):
 - ▶ модель № 1 «rs41590141 / rs110457178 / rs41609042» (суммарное количество образцов в модели - 2231): 0.9983, 1.0, 0.9986, 10/10, 0.9355 соответственно;
 - ▶ модель № 2 «rs41590141 / rs110457178 / rs110712082» (суммарное количество образцов в модели - 2260): 0.9959, 1.0, 0.9919, 10/10, 0.7462 соответственно;
 - ▶ модель № 3 «rs41590141 / rs41609042 / rs110712082» (суммарное количество образцов в модели - 2197): 0.9984, 1.0, 0.9982, 10/10, 0.9191 соответственно;
 - ▶ модель № 4 «rs110457178 / rs41609042 / rs110712082» (суммарное количество образцов в модели - 2218): 0.9974, 1.0, 0.9977, 10/10, 0.9008 соответственно.
- ▶ Модель №1 (rs41590141 / rs110457178 / rs41609042) характеризуется наилучшими показателями коэффициента Каппа (Cohen's Kappa) - **93,55%**, именно для этой модели далее дано развернутое описание.

Cohen's Kappa показывает, насколько согласие превышает уровень, ожидаемый чисто случайно. Значения Cohen's Kappa >0,81 принято интерпретировать как почти идеальное согласие.

#Результат (4)

Предсказанный класс	Количество образцов	Генотип rs41609042 / rs41590141 / rs110457178 /	Распространенность особей с данным генотипом [95% ДИ], %
Истинный класс – <i>Bos taurus</i> (n=2 209)			
<i>Bos taurus</i>	1592	GG/AA/CC	72.07 [70.25-73.85]
<i>Bos taurus</i>	208	GG/AT/CC	9.42 [8.42-10,45]
<i>Bos taurus</i>	92	GG/AA/AC	4.16 [3.47-4.90]
<i>Bos taurus</i>	57	GG/TT/CC	2.58 [1.97-3.23]
<i>Bos taurus</i>	36	GG/AT/AC	1.63 [1.14-2.17]
<i>Bos taurus</i>	26	AG/AA/CC	1.18 [0.77-1.00]
<i>Bos taurus</i>	21	AG/AT/AC	0.95 [0.60-1.34]
<i>Bos taurus</i>	17	GG/AT/AA	0.77 [0.47-1.11]
<i>Bos taurus</i>	17	GG/TT/AA	0.77 [0.47-1.11]
<i>Bos taurus</i>	17	AG/TT/AC	0.77 [0.47-1.11]
<i>Bos taurus</i>	15	GG/AA/AA	0.68 [0.40-1.01]
<i>Bos taurus</i>	15	GG/TT/AC	0.68 [0.40-1.01]
<i>Bos taurus</i>	15	AG/AT/CC	0.68 [0.40-1.01]
<i>Bos taurus</i>	14	AG/AT/AA	0.63 [0.36-0.95]
<i>Bos taurus</i>	11	AG/AA/AA	0.50 [0.27-0.77]
<i>Bos taurus</i>	11	AG/AA/AC	0.50 [0.27-0.77]
<i>Bos taurus</i>	10	AG/TT/AA	0.45 [0.23-0.72]
<i>Bos taurus</i>	в сумме 23	AA/AT/AA AG/TT/CC AA/TT/AC AA/AT/CC AA/AA/AC	1.04 [0.78-1.34]
Истинный класс – <i>Bison bonasus</i> (n=22)			
<i>Bison bonasus</i>	22	AA/TT/AA	100 [99.8-100]

85,65%

- *Bos taurus* - 0,54% [0,27%-0,88%]
- *Bison bonasus* были классированы корректно во всех четырех моделях

#Результат (5)

- ▶ Дополнительно определены SNV, расположенные в непосредственной близости от вышеописанных, и также обладающие высоким дифференцирующим потенциалом (расположены в порядке возрастания значений хромосомной позиции; жирным шрифтом указаны SNV из модели №1, описанной выше):
 - ▶ **на 4 хромосоме (Chr 4 (AC_000161.1))**
 - ▶ Chr.4:77221218C>G / AGGTGTTTCA[G/C]AAAGAGAGAA
 - ▶ Chr.4:77221599C>A / GTACCCAGTC[A/C]AACTACCAAT
 - ▶ **на 12 хромосоме (Chr 12 (AC_000169.1))**
 - ▶ Chr.12: 35188725A>G (rs432511880) / AAGTTTCAGT[A/G]TCTCTAAATA
 - ▶ Chr.12:35188752C>T (rs524291159) / TATTGGCACA[C/T]TGATCTGTGC
 - ▶ Chr.12:35188789G>A / AGCAGCAGA[A/G]CCAAGCATGA
 - ▶ Chr.12:35188873A>T / TGGTGACCCC[T/A]GCATTAGATG
 - ▶ **на 16 хромосоме (Chr 16 (AC_000173.1))**
 - ▶ Chr.16:12694162C>T (rs41609039) / GAATGATGGA[T/C]GGAGGTTCAT
 - ▶ Chr.16:12694323C>T (rs41609041) / ACCAGAAGCA[C/T]CATAAAAAAT
 - ▶ Chr.16:12694545C>T / CTACCGGCAA[C/T]GAGAАCTTTT



#Заключение

Среди всего многообразия SNV - более 5 тыс., вовлеченных в биоинформатический анализ, наиболее перспективными для молекулярно-генетических исследований могут быть полиморфные варианты (Bos_taurus_UMD_3.1.1 (GCF_000003055.6)):

- ▶ rs41590141 (Chr.4:77221374A>T),
- ▶ rs110457178 (Chr.12:35189028A>C),
- ▶ rs41609042 (Chr.16:12694384A>G).

В соответствии с действующим законодательством Республики Беларусь, уничтожение зубра, являющегося видом, находящимся под угрозой исчезновения и включенным в **Красную книгу** страны, влечет за собой особую правовую ответственность.

Поскольку зубры находятся под особой охраной государства, их уничтожение квалифицируется не только как причинение ущерба имуществу, но и как деяние, направленное против природных ресурсов. В данном случае применяются нормы, предусматривающие более строгое наказание.

Дополнительно к уголовной ответственности может быть применена гражданско-правовая ответственность в виде возмещения причиненного ущерба, размер которого определяется с учетом особой ценности уничтоженного животного и его значимости для сохранения вида.



Кипень Вячеслав

v.kipen@igc.by

+375259030765

Благодарю за внимание