Электронный научный журнал "Математическое моделирование, компьютерный и натурный эксперимент в естественных науках" http://mathmod.esrae.ru/ URL статьи: mathmod.esrae.ru/30-111

Ссылка для цитирования этой статьи:

Салтыков Ю.В., Ульянов С.С., Колосова А.А., Филонова Н.Н., Федорова В.А. Анализ нуклеотидных последовательностей гена GPCR представителей рода CAPRIPOXVIRUS с помощью спекл-интерферометрии GB-спеклов и вычитания их изображений // Математическое моделирование, компьютерный и натурный эксперимент в естественных науках. 2020. №2

# УДК 532.517.2:539.3 DOI: 10.24411/2541-9269-2020-10202 АНАЛИЗ НУКЛЕОТИДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ГЕНА GPCR ПРЕДСТАВИТЕЛЕЙ РОДА *САРКІРОХVІКUS* С ПОМОЩЬЮ СПЕКЛ-ИНТЕРФЕРОМЕТРИИ GB-СПЕКЛОВ И ВЫЧИТАНИЯ ИХ ИЗОБРАЖЕНИЙ

Салтыков Ю.В.<sup>1</sup>, Ульянов С.С.<sup>2</sup>, Колосова А.А.<sup>3</sup>, Филонова Н.Н.<sup>4</sup>, Федорова В.А.<sup>5</sup>

<sup>1</sup>Федеральный исследовательский центр вирусологии и микробиологии, филиал в Саратове, Россия, Саратов, <u>saltykov3443@mail.ru</u>

<sup>2</sup>Саратовский государственный университет, Россия, Саратов; Федеральный исследовательский центр вирусологии и микробиологии, филиал в Саратове,

Россия, Capatoв, prof.sergey.ulyanov@outlook.com

<sup>3</sup>Федеральный исследовательский центр вирусологии и микробиологии, филиал в Саратове, Россия, Саратов, <u>koloanyuta@yandex.ru</u>

<sup>4</sup>Федеральный исследовательский центр вирусологии и микробиологии, филиал в Саратове, Россия, Саратов, <u>nadejda.filonova@yandex.ru</u>

<sup>5</sup>Федеральный исследовательский центр вирусологии и микробиологии, филиал в Саратове, Россия, Саратов, <u>feodorovav@mail.ru</u>

# ANALYSIS OF NUCLEOTIDE SEQUENCES OF THE GENE GPCR OF THE GENUS CAPRIPOXVIRUS REPRESENTATIVES USING SPECKLE INTERFEROMETRY OF GB SPECKLES AND SUBTRACTING THEIR IMAGES

Saltykov Y.V.<sup>1</sup>, Ulyanov S.S.<sup>2</sup>, Kolosova A.A.<sup>3</sup>, Filonova N.N.<sup>4</sup>, Feodorova V.A.<sup>5</sup> <sup>1</sup>Federal Research Center for Virology and Microbiology, Branch in Saratov, Russia, Saratov, <u>saltykov3443@mail.ru</u>

<sup>2</sup>Saratov State University, Russia, Saratov; Federal Research Center for Virology and Microbiology, Branch in Saratov, Russia, Saratov, <u>prof.sergey.ulyanov@outlook.com</u> <sup>3</sup>Federal Research Center for Virology and Microbiology, Branch in Saratov, Russia, Saratov, koloanyuta@yandex.ru

<sup>4</sup>Federal Research Center for Virology and Microbiology, Branch in Saratov, Russia, Saratov, <u>nadejda.filonova@yandex.ru</u>

## <sup>5</sup>Federal Research Center for Virology and Microbiology, Branch in Saratov, Russia, Saratov, <u>feodorovav@mail.ru</u>

Аннотация. В данной статье рассмотрены алгоритмы преобразования нуклеотидных последовательностей (биодигитализации) гена GPCR представителей рода *Capripoxvirus* в виртуальные оптические спеклы (GB-спеклы). Для различных штаммов вирусов заразного узелкового дерматита крупного рогатого скота и оспы овец получены двумерные распределения интенсивности и фазы GB-спеклов. С целью выявления различий между исходными нуклеотидными последовательностями были изучены интерференционные картины, появляющихся суперпозиций GB-спекл полей и разности изображений GB-спеклов. Показана возможность использования цветных GB-спеклов для выявления различий между сопоставляемыми нуклеотидными последовательностями. Продемонстрирована высокая эффективность предложенных методов для детекции полиморфизма нуклеотидных последовательностей гена GPCR как молекулярной основы внутривидовой дискриминации каприпоксвирусов.

Ключевые слова: нуклеотидные последовательности, спекл-интерферометрия, цветные GB-спеклы, вычитание изображений.

**Abstract.** This article discusses algorithms for converting nucleotide sequences (biodigitalization) of the representatives of the gene GPCR of the genus *Capripoxvirus* into virtual optical gene-based speckles (GB-speckles). For different strains of cattle lumpy skin disease virus (LSDV) and sheeppox virus (SPPV) two-dimensional distributions of the intensity and phase of GB-speckles are obtained. In order to reveal the differences between the initial nucleotide sequences, the interference patterns of the appearing superpositions of GB-speckle fields and the difference in images of GB-speckles were studied. The possibility of using of colour GB-speckles to identify differences between the nucleotide sequences being matched is shown. High efficiency of the proposed methods for detecting polymorphism of nucleotide sequences of the GPCR gene as the molecular basis for intraspecific discrimination of *Capripoxvirus* is successfully demonstrated.

Keywords: nucleotide sequences, speckle-interferometry, colour GB-speckles, image subtraction.

### 1. Введение.

Как известно, нуклеиновые кислоты (ДНК или РНК) хранят и передают генетическую информацию в живых организмах. Структурной и функциональной единицей наследственной информации является ген. Ген представляет собой последовательность нуклеотидов в молекуле нуклеиновых кислот. Молекулы ДНК состоят из четырех типов нуклеотидов. Эти нуклеотиды содержат соответственно четыре азотистых основания, а именно аденин (А), тимин (Т), гуанин (G) и цитозин (С).

Нуклеотидную последовательность можно определить, используя специальную процедуру секвенирования [1, 2], которая позволяет представить первичную структуру макромолекулы в виде линейной последовательности мономеров в текстовом формате.

Как было показано ранее [3-9], преобразование генетических данных в компьютерные голограммы или представление последовательности нуклеотидов в виде спекл-структуры позволит, как значительно

усовершенствовать, так и создать инструменты современной биоинформатики и, в перспективе, методы лабораторной диагностики инфекционных и неинфекционных болезней человека и животных [2].

Ранее авторами данной статьи последовательности нуклеотидов гена omp1 бактерии Chlamydia trachomatis (геновары D, E, F, G, J и K) и бактерии C. psittaci были успешно конвертированы в двумерные спекл-поля. В работах [4-6] был введен специальный термин GB-спекл-структуры (gene-based speckles) для определения принципиально нового класса спекл-полей. GB-спекл-поля обладают уникальными статистическими свойствами, которые были частично исследованы в работе [5]. Как было показано в статье [4], использование таких методов спекл-оптики, как спекл-коррелометрия, вычитание изображений и спекл-интерферометрия, позволяет определить наличие природных мутаций в сравниваемых штаммах даже в случае минимальных различий всего в один нуклеотид (SNP, single nucleotide polymorphism). При этом показано, что появление некоторых типов мутаций (в частности, делеций [2]) ведет к формированию полос в интерференционной картине при использовании метода спекл-интерферометрии [4]. В работе [6] проведена оптимизация алгоритма кодирования нуклеотидных последовательностей бактерии C. trachomatis в двумерные GB-спекл-поля, показано, что алгоритм, описанный в [4], близок к оптимальному. В статьях [7, 8] метод виртуальной спекл-интерферометрии фазового сдвига (4 bucket technique) был использован для исследования полиморфизма у двух вариантов omp1 гена C. trachomatis (а именно штаммов E/Bour (E1 sub-type) и E/IU-4 2 0755u4 (E2 subtype)). Предложенный метод был высокоточной внутривидовой применен для дискриминации успешно нуклеотидных последовательностей гена omp1 C. trachomatis всех известных субтипов, несущих генетические мутации в виде одиночных SNP или их комбинации. Не менее информативным оказалось использование GB-спеклов нуклеотидных последовательностей анализа генов. кодирующих лля продукцию сериновых протеаз, белков семейства Omptin, энтеробактерий возбудителей таких актуальных инфекций, как сальмонеллезы, иерсиниозы, шигеллезы и эшерихиозы [9].

В данной статье с помощью GB-спеклов сравнивали последовательности гена хемокинового рецептора, связанного с G-белком каприпоксвирусов (от англ. G-Protein-Coupled Chemokine Receptor Gene, GPCR) – вируса заразного узелкового дерматита (ЗУД) крупного рогатого скота (КРС) и оспы овец. Упомянутые последовательности существенно различаются между собой, но, однако, имеют общие мотивы. Под мотивом (motif) в молекулярной биологии понимается характерная относительно короткая последовательность нуклеотидов в нуклеиновых кислотах или аминокислот в полипептидах, слабо меняющаяся в процессе эволюции и имеющая определённую биологическую функцию [10].

2. Методы и материалы.

2.1. Анализируемые генетические последовательности.

Для анализа были выбраны 5 нуклеотидных последовательностей, а именно, нуклеотидные последовательности гена GPCR возбудителя ЗУД КРС, в частности, «вакциноподобные» штаммы типа Neethling (последовательности взяты из базы данных NCBI GenBank): AF409138 Neethling vaccine LW 1959 Neethling-LSD vaccine-OBP (ген № 1). KX764645 (ген N⁰ 2) и «вакциноподобный» штамм типа Neethling (детектирован/обнаружен В Приволжском Федеральном округе РФ): MT129668 Saratov/Nesterovo/2019/n 1 (ген № 3).

Также были исследованы полевой (дикий) штамм: КY02007 SERBIA/Bujanovac/2016 (ген № 4) и нуклеотидная последовательность гена GPCR возбудителя оспы овец: MN072631 Sheeppox virus isolate Turkey vaccine (ген № 5).

Нуклеотидные последовательности полевых И «вакциноподобных» штаммов возбудителя ЗУД КРС были выбраны исходя из результатов молекулярно-филогенетического анализа И кластеризации штаммов. обнаруженных во время вспышек ЗУД КРС в РФ и других странах (Приложение 1, рис. 9-12). Эпидемиологические данные о вспышках ЗУД КРС в мире были получены из Глобальной информационной системы по болезням животных EMPRES Продовольственной и сельскохозяйственной организации использованием фактических географических Объединенных Наций c координат для каждого отдельного случая [11]. Известные вспышки ЗУД КРС на территории РФ были обнаружены в базах данных ФАО [12]. Все репрезентативные последовательности генов GPCR штаммов вируса ЗУД КРС, обнаруженные в разных регионах РФ и представленные в этом исследовании, были получены из базы данных NCBI GenBank [13]. Картографический и пространственно-временной анализ вспышек ЗУД КРС в регионах РФ и других странах мира проводили с использованием Esri ArcGis Desktop 10.6.1 [14].

## 2.2. Алгоритм преобразования нуклеотидной последовательности в GBспекл-структуру.

На первом этапе последовательность букв из исходной одномерной нуклеотидной последовательности преобразуется в последовательность чисел в соответствии со следующим правилом [4]:

 $A \to 1, C \to 2, G \to 3, T \to 4. (1)$ 

Важно подчеркнуть, что, как было показано в работе [6], взаимосвязь букв и чисел в данном случае не является принципиальной. Иными словами, при кодировании могло быть использовано любое другое правило, например:

C->1, G->2, T->3, A->4. (2) или T->1, A->2, C->3, G->4. (3)

На втором этапе генерируются все возможные комбинации (триады), содержащие лишь три числа из исходного полного набора из всех четырех чисел {1, 2, 3, 4}. В результате формируется полный набор всех триад:

(1 1 1), (1 1 2), (1 1 3), (1 1 4), (1 2 1), (1 2 2), (1 2 3), (1 2 4), (1 3 1), ..., (4 4 4).

Число всех возможных комбинаций из четырех чисел, объединенных в триады, равно 64. Затем, на следующем (третьем) этапе некоторая дискретная

величина h приписывается каждой триаде в соответствии с несложным алгоритмом, описанном в статье [4]. Упомянутый алгоритм был реализован на Matlab R2015a. Величина h является целым положительным числом, варьирующимся в интервале от 1 до 64. При этом каждая триада из исходной нуклеотидной последовательности ассоциируется только с одним значением h. Так, например, комбинация (1 1 1) соответствует величине h=1, (1 1 2) соответствует h = 2, (1 1 3) соответствует h=3, (1 1 4) соответствует h = 4, (1 2 1) соответствует h = 5, (1 2 2) соответствует h = 6 и так далее. Окончательно последняя комбинация (4 4 4) соответствует величине h=64.

На четвертом этапе из одномерного массива h формируется квадратная матрица H<sub>n,m</sub>. Физический смысл сформированной матрицы H состоит в том, что каждый ее элемент представляет собой локальную высоту некоторой виртуальной шероховатой последовательности, соответствующей локальному содержанию анализируемой генетической структуры. Полученные виртуальные шероховатые поверхности будут использованы для моделирования уникальных спекл-структур, соответствующих различным специфическим нуклеотидным последовательностям.

Двумерное спекл-поле, соответствующее каждой конкретной нуклеотидной поверхности, генерируется с использованием дифракции когерентного пучка с профилем квадратного сечения на виртуальной рассеивающей поверхности с микрорельефом, описываемым матрицей H<sub>n,m</sub>. Как уже упоминалось, H<sub>n,m</sub> задает высоты шероховатости поверхности. В каждой точке виртуального диффузора (в плоскости рассеяния пучка) вводится некоторая фазовая модуляция U<sub>n,m</sub>=exp(-2*π*iH<sub>n,m</sub>/64). Поверхность освещается при нормальном падении пучка, фаза в освещающем пучке является постоянной величиной.

3. Сопоставление исходных нуклеотидных последовательностей и их виртуальных образов.

3.1. Сравнение исходных нуклеотидных последовательностей гена GPCR возбудителя ЗУД КРС.

Для выравнивания нуклеотидных последовательностей был применен классический алгоритм Нидлмана-Вунша [15, 16]. Для анализа был использован пакет MATLAB R2015b (в частности, функция seqdotplot).

Графики точечных совпадений для двух последовательностей показаны на рис. 1. Последовательность для штамма AF409138 Neethling vaccine LW 1959 (ген № 1) сопоставляется с последовательностью штамма KX764645 Neethling-LSD vaccine-OBP (ген № 2). Как показывают вычисления, сравниваемые нуклеотидные последовательности являются полностью идентичными (совпадения составляют 100%). Из рис. 1 видно, что диагональная линия в матрице точечных совпадений наблюдается абсолютно ясно.



Рис. 1. Результаты сравнивания двух нуклеотидных последовательностей. Матрица совпадений для последовательностей гена GPCR штамма AF409138 Neethling vaccine LW 1959 и штамма KX764645 Neethling-LSD vaccine-OBP, соответственно

Ha MT129668 рис. 2a последовательность для штамма Saratov/Nesterovo/2019/n\_1 (ген № 3) сравнивается с последовательностью нуклеотидной последовательности гена GPCR возбудителя оспы овец MN072631 Sheeppox virus isolate Turkey vaccine (ген № 5). Очевидно, что диагональная линия в этом случае также наблюдается совершенно отчетливо. Это существенное сходство между сравниваемыми указывает на последовательностями. Однако уровень сходства существенно меньше, чем для предыдущих двух последовательностей и составляет лишь 91 %, см. рис. 26.



Рис. 2а. Результаты прямого сравнивания двух нуклеотидных последовательностей гена GPCR. Матрица совпадений для соответствующих последовательностей штаммов MT129668 Saratov/Nesterovo/2019/n\_1 и возбудителя оспы овец MN072631 Sheeppox virus isolate Turkey vaccine



Рис. 26. Множественное выравнивание нуклеотидных последовательностей гена GPCR штаммов MT129668 Saratov/Nesterovo/2019/n\_1 и MN072631 Sheeppox virus isolate Turkey vaccine с применением программы Multalin (<u>http://multalin.toulouse.inra.fr/multalin/</u>)

## 3.2. Генерирование GB-спеклов.

Процедура перекодирования (оцифровки) исходной нуклеотидной последовательности в GB-спекл-структуру на примере гена GPCR возбудителя вируса ЗУД КY02007 SERBIA/Bujanovac/2016 (ген № 4) приведена ниже.

Исходная нуклеотидная последовательность выглядит следующим образом:

## 

После конвертирования последовательности букв в последовательность чисел в соответствии с алгоритмом, описываемым формулой (1), нуклеотидная последовательность приобретает следующий вид:

3 4 1 3 2 1 3 4 1 1 4 1 4 4 1 2 2 1 2 4 1 4 1 3 2 4 1 2 4 1 2 1 1 4 4 1 4 4 1 3 4 1 2 1 1 4 4 4 2 1 1 2 1 1 1 4 2 1 1 1 1 4 1 1 4 3 4 4 1 2 1 1 2 3 2 2 4 4 2 1 1 2 4 4 1 4 3 1 1 1 1 4 1 2 1 1 2 1 1 2 3 1 4 1 4 2 4 1 1 4 4 1 4 1 2 1 1 2 2 3 2 1 4 1 4 1 1 4 1 2 1 1 2 4 4 1 4 4 1 4 1 3 2 3 1 4 3 1 4 4 1 4 3 1 4 3 1 4 4 1 4 3 1 1 3 4 3 1 3 2 1 4 1 3 4 2 3 1 4 1 4 2 2 212144343143143343434331412412113444433124314412 3 4 3 4 4 1 1 2 4 3 4 4 2 4 4 2 3 4 1 1 1 4 1 4 1 1 3 1 4 1 1 1 1 1 2 1 1 4 1 2 1 3 3 1 4 1 4 3 4 4 4 4 3 2 4 4 1 1 4 4 4 3 1 2 1 2 4 3 4 2 4 3 1 4 4 4 1 1 4 4 4 4 2 3 4 3 4 4 3 3 4 3 4 4 4 2 2 4 4 4 4 1 1 4 4 4 1 4 1 2 3 1 4 1 3 4 1 4 2 3 2 4 1 1 1 2 1 1 4 3 3 1 3 4 4 4 1 3 3 1 3 1 4 4 3 4 4 4 3 4 3 4 1 1 1 4 4 4 1 1 1 3 2 4 1 4 3 4 4 4 4 1 2 4 4 4 3 4 4 3 4 4 3 4 4 4 4 4 1 2 1 1 4 1 3 2 1 4 3 4 2 1 4 4 4 1 4 1 1 2 1 4 4 3 1 4 3 1 3 4 1 4 4 3 1 4 1 3 1 4 1 2 2 4 1 3 2 4 3 4 1 3 4 4 2 1 2 2 2 1 3 4 1 1 1 1 4 2 1 1 4 3 2 2 3 1 4 1 1 3 3 1 2 1 1 1 1 231414331144341244134143343433444331443424211241 4 4 3 1 1 4 2 2 4 4 4 2 2 1 1 4 1 1 4 3 4 4 1 4 4 4 4 4 4 4 3 1 1 1 2 1 1 1 1 1 1 1 3 4 1 4 1 4 3 3 1 1 4 1 1 2 3 4 1 4 4 3 4 2 1 4 3 4 1 4 4 4 4 4 4 1 4 1 1 2 3 1 4 1 1 4 3 2 1 1 1 1 1 4 4 4 4 4 1 1 1 1 1 2 1 4 2 1 2 1 1 1 2 1 1 1 3 1 1 4 1 1 3 1 1 1 3 2 2 1 4 1 1 1 3 1 4 3 3 4 3 44443144344142434421341443444412422214441343 1 4 3 4 1 4 3 3 2 3 4 4 1 1 3 1 4 4 4 3 4 2 1 1 2 2 4 4 3 2 1 3 4 4 2 1 4 3 4 1 3 2 4 3 1 1  $1\ 4\ 4\ 3\ 4\ 3\ 4\ 2\ 4\ 1\ 4\ 3\ 4\ 2\ 1\ 4\ 4\ 4\ 1\ 4\ 3\ 2\ 3\ 4$ 4 4 4 3 4 1 3 4 1 3 1 3 1 1 4 4 4 1 2 4 1 1 1 1 1 3 2 4 4 4 4 1 2 3 1 4 4 3 2 3 4 1 3 2 1 2 2 1 3 4 1 3 4 3 2 4 3 3 4 1 3 4 1 4 4 1 3 2 1 4 4 3 3 1 4 1 1

Преобразования остальных текстовых последовательностей в числовой формат представлено в Приложении 2.

В результате дифракции когерентного пучка на фазовом экране (с высотами виртуальных неоднородностей, представленных выше) с квадратным сечением формируется GB-спекл-структура, рис. За. Двумерное распределение фазы GB-спекл-структуры показано на рис. Зб.



Рис. За. Характерный вид двумерного распределения интенсивности GB-спеклструктуры, полученной для гена GPCR возбудителя вируса ЗУД штамма КY02007 SERBIA/Bujanovac/2016



Рис. 3б. Характерный вид двумерного распределения фазы GB-спекл-структуры для гена GPCR возбудителя вируса ЗУД штамма КY02007 SERBIA/Bujanovac/2016

На рис. 4а,б и рис. 5а,б представлены GB-спекл структуры, полученные для «вакциноподобного» штамма типа Neethling (детектирован в Приволжском Федеральном округе РФ), и последовательности гена GPCR возбудителя оспы овец MN072631 Sheeppox virus isolate Turkey vaccine.



Рис. 4а. Характерный вид двумерного распределения интенсивности GB-спеклструктуры, полученной для нуклеотидной последовательности гена GPCR «вакциноподобного» штамма типа Neethling (обнаружен в Приволжском Федеральном округе РФ), MT129668 Saratov/Nesterovo/2019/n\_1



Рис. 4б. Характерный вид двумерного распределения фазы GB-спекл-структуры для нуклеотидной последовательности гена GPCR «вакциноподобного» штамма типа Neethling (обнаружен в Приволжском Федеральном округе РФ), MT129668 Saratov/Nesterovo/2019/n\_1



Рис. 5а. Характерный вид двумерного распределения интенсивности GB-спеклструктуры, полученной для последовательности гена GPCR возбудителя оспы овец MN072631 Sheeppox virus isolate Turkey vaccine



Рис. 5б. Характерный вид двумерного распределения фазы GB-спекл-структуры для последовательности гена GPCR возбудителя оспы овец MN072631 Sheeppox virus isolate Turkey vaccine

Сравнивая спекл-поля, изображенные на рис. 4а и рис. 5а, можно сделать заключение о существенном сходстве в структурах между GB-спеклами, полученных для различных генов. Прежде всего, сходство касается наличия зеркального компонента во всех сопоставляемых спекл-полях. Однако, как показывает корреляционный анализ, величина коэффициента кросс-корреляции не столь велика и составляет 0.812 для рис. 3а и рис. 4а, 0.83 для рис. 4а и рис. 5а и 0.764 для рис. 3а и рис. 5а. Корреляция между фазовыми структурами GB-

спеклов (рис. 36, 46 и рис. 56) полностью отсутствует. Но, так или иначе, снижение коэффициента кросс-корреляции с 1 до уровня примерно 0.8 свидетельствует о возможности выявления полиморфизма у генетических последовательностей с помощью кросс-корреляционного анализа GB-спеклов.

В следующих разделах будет рассмотрена возможность выявления полиморфизма с помощью методов вычитания изображений, спекл-интерферометрии и построения цветных спекл-структур.

#### 3.3. Вычитание GB-спеклов.

На рис. ба представлено вычитание GB-спеклов, построенных для генов № 3 и № 4, на рис. бб – для генов № 3 и № 5, а на рис. 6в, соответственно, для генов № 4 и № 5.



Рис. 6а. Результат вычитания GB-спеклов, построенных для нуклеотидных последовательностей гена GPCR штаммов MT129668 Saratov/Nesterovo/2019/n\_1 и KY02007 SERBIA/Bujanovac/2016



Рис. 6б. Результат вычитания GB-спеклов, построенных для нуклеотидных последовательностей гена GPCR штаммов MT129668 Saratov/Nesterovo/2019/n\_1 и MN072631 Sheeppox virus isolate Turkey vaccine



Рис. 6в. Результат вычитания GB-спеклов, построенных для нуклеотидных последовательностей гена GPCR штаммов KY02007 SERBIA/Bujanovac/2016 и MN072631 Sheeppox virus isolate Turkey vaccine

Следует отметить, что в данной статье интенсивность на всех рисунках перенормирована в интервал от 0 до 255 для лучшей визуализации GB-спеклструктур. Как видно из рисунков ба, бб и бв, наличие даже минимальных различий в сопоставляемых нуклеотидных последовательностях немедленно приводит к появлению существенной неоднородности в результирующем изображении GB-спеклов. При этом контраст этих изображений чрезвычайно высок и составляет порядка 10<sup>7</sup>. Это означает, что техника вычитания изображений так же может быть эффективно использована для выявления минимальных различий сравниваемых нуклеотидных последовательностей.

## 3.4. Спекл-интерферометрия GB-спеклов.

Как было показано ранее, при интерференции двух GB-спекл-полей в результирующей интерференционной картине появляются интерференционные полосы, как правило, модулированные спеклами. Иногда амплитудная модуляция бывает исчезающе малой, но формирующиеся интерференционные полосы имеют изгибы.

На рис. 7 показана интерференционная картина, образующаяся при суперпозиции GB-спеклов, сгенерированных для генов № 3 и № 4 (рис. 7а), сгенерированных для генов № 3 и № 5 (рис. 7б), и, соответственно, для генов № 4 и № 5 (рис. 7в).



Рис. 7а. Интерференционная картина, образующаяся при суперпозиции GB-спеклов, сгенерированных для нуклеотидных последовательностей гена GPCR штаммов MT129668 Saratov/Nesterovo/2019/n\_1 и KY02007 SERBIA/Bujanovac/2016



Рис. 76. Интерференционная картина, образующаяся при суперпозиции GB-спеклов, сгенерированных для нуклеотидных последовательностей гена GPCR штаммов MT129668 Saratov/Nesterovo/2019/n\_1 и MN072631 Sheeppox virus isolate Turkey vaccine



Рис. 7в. Интерференционная картина, образующаяся при суперпозиции GB-спеклов, сгенерированных для нуклеотидных последовательностей гена GPCR штаммов KY02007 SERBIA/Bujanovac/2016 и MN072631 Sheeppox virus isolate Turkey vaccine

Как видно из рисунков 7а, 7б и 7в, в интерференционной картине наблюдаются прообразы полос, ширина которых превышает размер спеклов, но эти полосы малозаметны. Таким образом, спекл-интерферометрия виртуальных GB-спеклов малоинформативна с точки зрения выявления полиморфизма генетических последовательностей.

### 3.5. Цветные GB-спеклы.

Если сгенерировать три двумерные реализации GB-спеклов, построенные для различных генетических последовательностей, то можно сконструировать цветное изображение, где в качестве каждого цветового компонента (красный, зеленый и синий) поступает своя GB-спекл-структура. Если все три спекл-структуры идентичны, то цветовое изображение будет являться таковым лишь формально, а выглядеть оно будет как серое. Если же цветовые компоненты будут различаться между собой, то в результате в изображении будет появляться окрашивание.

На рис. 8а представлена цветная спекл-структура (красный компонент получен для нуклеотидной последовательности гена № 3, зеленый компонент – для нуклеотидной последовательности гена № 4, синий – для нуклеотидной последовательности гена № 5).



Рис. 8а. Цветная спекл-структура для нуклеотидных последовательностей гена GPCR штаммов MT129668 Saratov/Nesterovo/2019/n\_1, KY02007 SERBIA/Bujanovac/2016 и MN072631 Sheeppox virus isolate Turkey vaccine

Из рис. 8а видно, что при наличии различий в исходных нуклеотидных последовательностях в цветной спекл-структуре для двумерного распределения интенсивности появляется незначительное окрашивание.

На рис. 8б показана цветная спекл-структура для фазы GB-спеклов (попрежнему красный компонент получен для нуклеотидной последовательности гена № 3, зеленый компонент – для нуклеотидной последовательности гена № 4, синий – для нуклеотидной последовательности гена № 5).



Рис. 8а. Цветная спекл-структура для фазы GB-спеклов нуклеотидных последовательностей гена GPCR штаммов MT129668 Saratov/Nesterovo/2019/n\_1, KY02007 SERBIA/Bujanovac/2016 и MN072631 Sheeppox virus isolate Turkey vaccine

Совершенно очевидно, что в рассматриваемом случае наблюдается ярко выраженное окрашивание по всему полю изображения.

Таким образом, полученное цветное изображение для фазы GB-спеклов является надежным диагностическим признаком появления полиморфизмов генетических последовательностей.

### 4. Заключение.

В данной работе продемонстрированы перспективы таких оптических методов как кросс-корреляционный анализ и спекл-интерферометрия GB-спеклов, вычитание их изображений, и применение цветовых GB-спеклов в задачах детекции полиморфизмов нуклеотидных последовательностей

таргетного гена GPCR двух представителей рода *Capripoxvirus*, возбудителей ЗУД КРС и оспы овец.

Показано, что наибольшей информативностью отличается диагностика, основанная на применении цветового анализа фазовой структуры GB-спеклов. Высокой точностью диагностики обладают так же кросс-корреляционный анализ GB-спеклов и вычитание изображений. Информативность методов спекл-интерферометрии несколько ниже, чем остальных предложенных оптических методов.

### Литература

- 1. Sintchenko V., Roper M.P. Pathogen genome bioinformatics // Methods in Molecular Biology. 2014. Vol. 1168. P. 173-193.
- Lesk A.M. Introduction to bioinformatics. Oxford: Oxford University Press, 2002. 314 p.
- 3. Madi A., Friedman Y., Roth D., Regev T., Bransburg-Zabary S., Jacob E.B. Genome holography: Deciphering function-form motifs from gene expression data // PLoS One. 2008. Vol. 3. No 7. 114 p.
- 4. Ulyanov S.S., Zaytsev S.S., Ulianova O.V., Saltykov Y.V., Feodorova V.A. Using of methods of speckle optics for Chlamydia trachomatis typing // Proc. SPIE. 2017. Vol. 10336. No 103360D.
- 5. Ulyanov S.S., Ulianova O.V., Zaytsev S.S., Saltykov Y.V., Feodorova V.A. Statistics on gene-based laser speckles with a small number of scatterers: implications for the detection of polymorphism in the *Chlamydia trachomatis omp1* gene // Laser Physics Letters. 2018. Vol. 15. No 4. P. 1-6.
- 6. Feodorova V.A., Ulyanov S.S., Zaytsev S.S., Saltykov Y.V., Ulianova O.V. Optimization of algorithm of coding of genetic information of *Chlamydia* // Proc. SPIE. 2018. Vol. 10716. No 107160Q.
- Feodorova V.A., Saltykov Y.V., Zaytsev S.S., Ulyanov S.S., Ulianova O.V. Application of virtual phase-shifting speckle-interferometry for detection of polymorphism in the Chlamydia trachomatis omp1 gene // Proc. SPIE. 2018. Vol. 10716. No 107160M.
- Ulianova O.V., Ulyanov A.S., Zaytsev S.S., Saltykov Y.V., Ulyanov S.S., Feodorova V.A. LASCA-imaging of GB-speckles: Application for detection of the gene polymorphism in bacterial model // Laser Physics Letters. 2020. Vol. 17. No 6.
- 9. Ульянов С.С., Ульянова О.В., Зайцев С.С., Хижнякова М.А., Салтыков Ю.В., Филонова Н.Н., Субботина И.А., Ляпина А.М., Федорова В.А. Исследование статистических характеристик оптических GB-спеклов, формирующихся при рассеянии света на виртуальных структурах нуклеотидных последовательностей генов энтеробактерий // Изв. Сарат. ун-та. Нов. сер. Сер. Физика. 2018. Т. 18, № 2. С. 123-137.
- 10. D'haeseleer P. What are DNA sequence motifs? // Nature Biotechnology. 2006. Vol. 24. No 4. P. 423-425.

- 11. Summary of outbreaks of in. URL: <u>https://www.oie.int/wahis\_2/public/wahid.php/</u> <u>Diseaseinformation/Immsummary/listoutbreak</u>.
- 12. EMPRES-i Global Animal Disease Information System. URL: <u>http://empres-i.fao.org/</u>.
- 13. GenBank Overview. URL: <u>https://www.ncbi.nlm.nih.gov/genbank/</u>.
- 14. Esri: GIS Mapping Software. URL: <u>https://www.esri.com</u>.
- 15. Needleman S.B., Wunsch C.D. A general method applicable to the search for similarities in the amino acid sequence of two proteins // Journal of Molecular Biology. 1970. Vol. 48. No 3. P. 443-453.
- 16. Сетубал Ж., Мейданис Ж. Введение в вычислительную молекулярную биологию. Москва, Ижевск: Регулярная и хаотическая динамика, Ижевский институт компьютерных исследований, 2007. 420 с.

#### Приложение 1





0.005

Рис. 9 – Филогенетический анализ на основе гена GPCR полевых, «вакциноподобных» штаммов вируса ЗУД и изолята «МТ129668 Saratov/Nesterovo/2019/n 1»

## Результаты картографического анализа.

Вспышки ЗУД КРС, зарегистрированные во всем мире с 2006 года, представлены на рис. 10.



Рис. 10 – Вспышки ЗУД КРС, зарегистрированные во всем мире с 2006 года

До 1999 года, варианты вируса ЗУД КРС «вакциноподобного» Neethlingтипа выявляли только в Южной Африке, тогда как полевые штаммы вируса ЗУД КРС были обнаружены в основном (кроме Хорватии в 2016 г.) в других странах, включая РФ (рис. 11).



Рис. 11 – Данные о странах, где были обнаружены полевые и «вакциноподобные» штаммы вируса ЗУД КРС

Начиная с 2017 года, в различных регионах РФ, включая Саратовскую область, были одновременно выявлены как полевые, так и «вакциноподобные» варианты вируса ЗУД КРС (рис. 12).



Рис.12 – Данные о регионах РФ, где были обнаружены полевые и «вакциноподобные» штаммы вируса ЗУД КРС

#### Приложение 2

# Преобразования текстовых последовательностей в числовой формат (биодигитализация).

Исходная нуклеотидная последовательность гена GPCR «вакциноподобного» штамма типа Neethling (детектирован в Приволжском Федеральном округе РФ) MT129668 Saratov/Nesterovo/2019/n\_1 (gene4), выглядит следующим образом:

ATGAATTATACTCTTAGCACAGTTAGTAGCGCCACCATGTATAATAGTAG CAGTAATATTACCACTATAGCTACTACAATTATTAGCACAATTCTCAGTAC AATTTCAACAAATCAAAATAATGTTACAACGCCTTCAACTTATGAAAATA CGACAACAATATCTAATTATACAACCGCATATAATACAACTTATTATAGC GATGATTATGATGATTATGAAGTGAATATAGTCGATATCCCACATTGTGA TGATGGTGTGGATACTACAAGTTTTGGACTGATTACTTTATATTCGACTAT ATTCTTTCTTGGATTATTTGGAAAATATAATTGTGTTAACTGTTCTTCGTAA ATATAAGATAAAAAACAATACAGGATATATTTTTGCTTAATCTGACATTGT CTGATTTAATTTTCGTGTTGGTGTTTCCTTTTAATTTATACGATAGTATCGC TAAACAATGGAGTTTAGGAGATTGTTTGTGTAAATTCAAAGCTATGTTTT ACTTTGTTGGTTTTTACAATAGTATGTCATTTATAACATTGATGAGTATCG ATAGATACCTAGCTGTAGTTCACCCAGTAAAATCAATGCCGATAAGGACA AAACGATATGGAATTGTACTTAGTATGGTGGTTTGGATTGTTTCAACTATT GAATCCTTTCCAATAATGTTATTTATGAAACAAAAAAGTATATGGAAT AACGTATTGTCATGTATTTTATAACGATAATGCAAAAATTTGGAAAATTATT TATAAATTTTGAAATAAACATATTTGGAATGATTATACCGCTAATTATTTT GCTATATTGTTACTATAAAAATCTTAAAATACTTTAAAAAACCTCGCAAACAA AGAATAAGAAAGCCATAAAGATGGTGTTTTTGATTGTTATCTGTTCAGTA TTGTTTTACTCCCATTTAGTGTAACTGTATTTGTTTCATCGTTGTATTTGT TAAATGTTTTTAGTGGATGTACGGCATTACGATTTGTCAACCTTGCAGTTC ATGTAGCTGAAATTGTGTCTCTATGTCATTGTTTTATTAATCCACTAATTT ATGCGTTTTGTAGTAGAGAATTTACTAAAAAGCTTTTACGATTGCGTAGC ACTAGTAGTGCTGGTAGTATTAGCATTGGATAA

После конвертирования последовательности букв в последовательность чисел в соответствии с алгоритмом (1), нуклеотидная последовательность приобретает следующий вид:

1 1 4 1 2 1 3 3 1 4 1 4 1 4 4 4 4 4 3 2 4 4 1 1 4 2 4 3 1 2 1 4 4 3 4 2 4 3 1 4 4 4 1 1 4 4 2444344334444412114134143421444141121443143134 1 4 2 3 1 4 1 3 1 4 1 2 2 4 1 3 2 4 3 4 1 3 4 4 2 1 2 2 2 1 3 4 1 1 1 1 4 2 1 1 4 3 2 2 3 1 4 1 1 3 3 1 2 1 1 1 1 2 3 1 4 1 4 3 3 1 1 4 4 3 4 1 2 4 4 1 3 4 1 4 3 3 4 3 3 4 4 4 3 3 1 4 4344421124144311422444221141143441444414311121  $1\,1\,1\,1\,1\,3\,4\,1\,4\,1\,4\,3\,3\,1\,1\,4\,1\,1\,2\,3\,4\,1\,4\,4\,3\,4\,2\,1\,4\,3\,4\,1\,4\,4\,4\,4\,4\,4\,4\,1\,4\,1\,1\,2\,3\,1\,4\,1$ 4 3 3 1 1 4 3 1 4 4 1 4 1 2 2 3 2 4 1 1 4 4 1 4 4 4 4 3 2 4 1 4 1 4 4 3 4 4 1 2 4 1 4 1 1 1 1424411141244411111224232111211131141131113221 4 1 1 1 3 1 4 3 3 4 3 4 4 4 4 4 3 1 4 4 3 4 4 1 4 2 4 3 4 4 2 1 3 4 1 4 4 3 4 4 4 4 1 2 4 2 2 2 1 4 4 4 1 3 4 3 4 1 1 2 4 3 4 1 4 4 4 3 4 4 4 2 1 4 2 3 4 4 3 4 1 4 4 4 3 4 4 1 1 1 4 3 4 4 4 4 1 3 4 3 3 1 4 3 4 1 2 3 3 2 1 4 4 1 2 3 1 4 4 4 3 4 2 1 1 2 2 4 4 3 2 1 3 4 4 2 1 4 3 4 1 3 2 4 3 1 1 1 4 4 3 4 3 4 2 4 2 4 1 4 3 4 2 1 4 4 3 4 4 4 4 1 4 4 1 1 4 2 2 1 2 4 4 4 3 2 3 4 1 3 2 1 2 4 1 3 4 1 3 4 3 2 4 3 3 4 1 3 4 1 4 4 1 3 2 1 4 4 3 3 1 4 1 1

Исходная нуклеотидная последовательность гена GPCR возбудителя оспы овец MN072631 Sheeppox virus isolate Turkey vaccine (gene7), выглядит следующим образом:

ATGAATTATACTCTTAGAACAGTTAGTAGCAGTAATATTACCACTATAGC TACTACAATTATTAGTACAATTCTCAGTAGAATTTCAACAAATAAAAATA ATGTTACAACGCCTTCAACTTATGAAAATACGACAGCAATATCTAATTAT AAAACAGCATATAATATAACCTATTACAGCGATGATTATGATGATTATGA AAATATAATTGTGTTAACTGTTCTTCGTAAATATAAGATAAAAACGATAC AGGATATATTTTGCTTAATTTGACATTGTCTGATTTAATTTTCGTTTTGGT GTTTCCTTTTAATTTATACAATAGTATAGCTAAACAATGGAGTTTGGGAG ATTGTTTGTGTAAATTCAAAGCTATGTTTTACTTTGTTGGTTTTTACAATA GCATGTCATTTATAACATTGATGAGTATTGATAGATACCTAGCTATAGTTT ACCCAGTAAAATCAATAACGATAAGGACAAAACAATACGGAATTGTACT TAGTATGGTGGTTTGGATTGTCTCAATTATTGAATCCTTTCCAATAATGTT ATTTTATGAAACAAAAAAAGTATATGGAATAACACATTGTCATGTATTTT ATAACGATAATGCAAAAATTTGGAAATTATTTATAAATTTTGAAAATAAAC ATATTTGGAATGATTATACCGCTAATTATTTTGCTATATTGTTATTATAAA ATCTTAAATACTTTAAAAACATCACAAACAAAGAATAAGAAAGCCATAA AGATGGTGTTTTTAATTGTTATCTGTTCAGTATTGTTTTTACTCCCATTTAG TGTAACTGTATTTGTTTCATCGTTGTATTTGTTAAATGTTTTAGTGGATGT ATTGCATTAAGATTTGTCAACCTTGCAGTTCATGTAACTGAAATTGTGTCT CTATGTCATTGTTTTATCAACCCACTAATTTATGCGTTTTGTAGTAGAGAA

## TTTACTAAAAAGCTTTTACAATTGCGTAGCACCAGTAGTGTTGGTAGTATT AGCATTGGATAA

После конвертирования последовательности букв в последовательность чисел в соответствии с алгоритмом (1), вышеприведенная нуклеотидная последовательность приобретает следующий вид:

4 1 3 2 4 1 2 4 1 2 1 1 4 4 1 4 4 1 3 4 1 2 1 1 4 4 2 4 2 1 3 4 1 3 1 1 4 4 4 2 1 1 2 1 1 1 4 1 1 1 1 1 4 1 1 4 3 4 4 1 2 1 1 2 3 2 2 4 4 2 1 1 2 4 4 1 4 3 1 1 1 1 4 1 2 3 1 2 1 3 2 1 1414241144141111213214141141411224144121323143  $1\ 4\ 4\ 1\ 4\ 3\ 1\ 4\ 3\ 1\ 4\ 3\ 1\ 4\ 3\ 1\ 1\ 3\ 4\ 3\ 1\ 1\ 2\ 1\ 4\ 1\ 3\ 4\ 2\ 3\ 1\ 4\ 1\ 4\ 2\ 2\ 2\ 1\ 2\ 1\ 2\ 4\ 3\ 4\ 3\ 1\ 1$ 4314334343414124121134444331244144124441414441  $1\ 2\ 4\ 1\ 4\ 1\ 4\ 4\ 2\ 4\ 4\ 3\ 3\ 1\ 4\ 4\ 4\ 4\ 3\ 3\ 1\ 1\ 1\ 4\ 1\ 4\ 4\ 3\ 4\ 3\ 4\ 4\ 1\ 1\ 2\ 4\ 3$ 4424423411141411314111123141213314121331414144444324 11444141211413414132411121143313444333131443444 3 4 3 4 1 1 1 4 4 2 1 1 1 3 2 4 1 4 3 4 4 4 1 2 4 4 4 3 4 4 3 3 4 4 4 4 1 2 1 1 4 1 3 2 1 4 3 4 2 1 4 4 4 1 4 1 1 2 1 4 4 3 1 4 3 1 3 4 1 4 4 3 1 4 1 3 1 4 1 2 2 4 1 3 2 4 1 4 1 3444122213411114211411231411331211112114123311 4 4 3 4 1 2 4 4 1 3 4 1 4 3 3 4 3 3 4 4 4 3 3 1 4 4 3 4 2 4 2 1 1 4 4 1 4 4 3 1 1 4 2 2 4 4 4 2 2 1 1 4 1 1 4 3 4 4 1 4 4 4 4 4 4 4 3 1 1 1 2 1 1 1 1 1 1 1 3 4 1 4 1 4 3 3 1 1 4 1 1 2 1 4 4 1 4 1 1 1 4 4 4 4 3 1 1 1 4 1 1 1 2 1 4 1 4 4 4 3 3 1 1 4 3 1 4 4 1 4 1 2 2 3 2 4 1 1 4 4212111211131141131113221411131433434444411443 4 4 1 4 2 4 3 4 4 2 1 3 4 1 4 4 3 4 4 4 4 4 1 2 4 2 2 2 1 4 4 4 1 3 4 3 4 1 1 2 4 3 4 1 4 4 43444214234434144434411143444444134331434144321 4 4 1 1 3 1 4 4 4 3 4 2 1 1 2 2 4 4 3 2 1 3 4 4 2 1 4 3 4 1 1 2 4 3 1 1 1 4 4 3 4 3 4 2 4 2 414342144344414211222124114441432344443413413  $1\ 3\ 1\ 1\ 4\ 4\ 4\ 1\ 2\ 4\ 1\ 1\ 1\ 1\ 3\ 2\ 4\ 4\ 4\ 4\ 1\ 2\ 1\ 1\ 4\ 4\ 3\ 2\ 3\ 4\ 1\ 3\ 2\ 1\ 2\ 2\ 1\ 3\ 4\ 1\ 3\ 4\ 3\ 4\ 3\ 4\ 4$ 334134144132144331411