

#### RESEARCH ARTICLE

# **REVISED** Discrimination of the SARS–CoV-2 strains using

# of coloured s-LASCA-imaging of GB-

speckles, developed for the gene "S" nucleotide sequences

# [version 4; peer review: 2 approved, 1 approved with

# reservations]

Onega Ulianova<sup>1</sup>, Yury Saltykov<sup>2</sup>, Sergey Ulyanov<sup>1,2</sup>, Sergey Zaytsev<sup>2</sup>, Alexander Ulyanov<sup>1</sup>, Valentina Feodorova<sup>1</sup>

<sup>1</sup>Department of Medical Physics, Saratov State University, Saratov, 410012, Russian Federation

<sup>2</sup>Laboratory for Molecular Biology and NanoBioTechnology, Federal Research Center for Virology and Microbiology, Branch in Saratov, Saratov, 410076, Russian Federation

## First published: 25 Jun 2021, 10:503 https://doi.org/10.12688/f1000research.53214.1 Second version: 06 Sep 2021, 10:503 https://doi.org/10.12688/f1000research.53214.2

Third version: 08 Nov 2021, **10**:503 https://doi.org/10.12688/f1000research.53214.3

Latest published: 22 Jun 2022, 10:503 https://doi.org/10.12688/f1000research.53214.4

#### Abstract

**Background**: A recent bioinformatics technique involves changing nucleotide sequences into 2D speckles. This technique produces speckles called GB-speckles (Gene Based speckles). All classical strategies of speckle-optics, namely speckle-interferometry, subtraction of speckle-images as well as speckle-correlometry have been inferred for processing of GB-speckles. This indicates the considerable improvement in the present tools of bioinformatics. **Methods**: Colour s-LASCA imaging of virtual laser GB-speckles, a new method of high discrimination and typing of pathogenic viruses, has been developed. This method has been adapted to the detecting of natural mutations in nucleotide sequences, related to the spike glycoprotein (coding the gene «S») of SARS-CoV-2 gene as the molecular target.

**Results**: The rate of the colouring images of virtual laser GB-speckles generated by s-LASCA can be described by the specific value of R. If the nucleotide sequences compared utilizing this approach the relevant images are completely identical, then the three components of the resulting colour image will be identical, and therefore the value of R will be equal to zero. However, if there are at least minimal differences in the matched nucleotide sequences, then

#### **Open Peer Review** Approval Status 🗹 ? 🗸 1 2 3 version 4 (revision) 22 Jun 2022 version 3 (revision) view 08 Nov 2021 version 2 ? (revision) view view 06 Sep 2021 ? version 1 25 Jun 2021 view

- Oleg Angelsky D, Chernivtsi National University, Chernivtsi, Ukraine
  Claudia Zenkova, Chernivtsi National University, Chernivtsi, Ukraine
- 2. Alexey Bashkatov, Saratov State University,

the value of R will be positive.

**Conclusion**: The high effectiveness of an application of the colour images of GB-speckles that were generated by s-LASCA-has been demonstrated for discrimination between different variants of the SARS–CoV-2 spike glycoprotein gene.

Keywords LASCA, SARS–CoV-2, GB-speckles, gene

This article is included in the Emerging Diseases and Outbreaks gateway.



This article is included in the Bioinformatics gateway.



This article is included in the Cell & Molecular Biology gateway.



This article is included in the Coronavirus collection.

Corresponding authors: Sergey Ulyanov (prof.sergey.ulyanov@outlook.com), Valentina Feodorova (feodorovav@mail.ru)

Author roles: Ulianova O: Formal Analysis, Writing – Original Draft Preparation; Saltykov Y: Data Curation, Formal Analysis, Investigation, Project Administration; Ulyanov S: Conceptualization, Funding Acquisition, Investigation, Methodology, Resources, Software, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; Zaytsev S: Data Curation, Investigation, Validation; Ulyanov A: Data Curation, Software, Writing – Original Draft Preparation; Feodorova V: Conceptualization, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

Grant information: The reported study was funded by RFBR, project number 20-04-60349.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Copyright:** © 2022 Ulianova O *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Ulianova O, Saltykov Y, Ulyanov S *et al.* Discrimination of the SARS-CoV-2 strains using of coloured s-LASCAimaging of GB-speckles, developed for the gene "S" nucleotide sequences [version 4; peer review: 2 approved, 1 approved with reservations] F1000Research 2022, **10**:503 https://doi.org/10.12688/f1000research.53214.4

First published: 25 Jun 2021, 10:503 https://doi.org/10.12688/f1000research.53214.1

Saratov, Russian Federation

3. **Dmitry A. Zimnyakov**, Yury Gagarin State Technical University of Saratov, Saratov, Russian Federation

Any reports and responses or comments on the article can be found at the end of the article.

#### **REVISED** Amendments from Version 3

Referring to comments from the reviewers, we have made the following changes:

- 1. One new reference (M. Francon' book as the reference # 17) is introduced in the paper.
- 2. We added the reference 16 citation to the "s-LASCA imaging of GB-speckles" section in the text of the paper.
- 3. We provided to the reviewers necessary explanations of the original method that we used to obtain the data presented in our manuscript.
- 4. More details were done in the Legends for the Figures 1-2.
- 5. In the text only several technical revisions were made according to the recommendations of the reviewers.

Any further responses from the reviewers can be found at the end of the article

#### Introduction

As it is well known, if laser light diffracts on random objects, then laser speckles are formed.<sup>1-3</sup> Recently, the possibility of transforming a nucleotide sequence into a pattern of 2D speckles had been demonstrated.<sup>4-9</sup> This new type of speckle pattern has been called "GB-speckles" (gene-based speckles).<sup>5,7,9</sup> Changes within in the structure of the GB-speckles can reflect even negligible changes in the nucleotide sequence, caused by inartificial mutations. This allows detection of single-nucleotide polymorphisms (SNP) using virtual GB-speckles with outstanding precision. In addition, it offers unlimited potential of improving the diagnosis' accuracy by increasing the Fourier transform area.<sup>10</sup>

Essential advancement in the area of GB-speckles has been reported in previous years. According to previously published reports,<sup>4-9,11</sup> implementation of speckle-optics methods, like speckle-interferometry and subtraction of speckle-images as well as speckle-correlometry for processing of GB-speckles, provides considerable progress in the current bioinformatics toolbox. This can become crucial to significantly improve existing routine methods of laboratory diagnostics of infectious diseases. GB-speckles as a technique opens the door to the new horizons in digital biology.<sup>12,13</sup>

Recently, model GB-speckle patterns of nucleotide sequences of the *omp1* genes for two different of *Chlamydia* spp., such as *Chlamydia trachomatis* and *Chlamydia psittaci* of at least six genovars (D, E, F, G, J and K) have been composed.<sup>4,5</sup> Probability density functions and correlation properties of spatial intensity fluctuations for the relevant GB-speckle patterns have been studied.<sup>5-7</sup> As it has been shown in previous studies,<sup>4-7,9</sup> the presence of inartificial mutations in analysed strains, including single SNP cases, can be easily defined using methods of speckle-optics.<sup>4-7,9</sup> More recently, the encoding algorithm's optimization for nucleotide sequences of *C. trachomatis* into two-dimensional GB-speckle pattern had been carried out;<sup>4,6</sup> and speckle-interferometric technique may give rise the ultra-fast optical processors of DNA sequences.<sup>4</sup> This is ensured by the development of the exclusive system of interferential fringes which are generated by the model interference pattern led by the existence of any type of mutations. Additionally, the method of virtual phase-shifting speckle-interferometry was reported to be efficacious<sup>11</sup> to investigate of polymorphism of the *C. trachomatis omp1* gene. This approach allowed the detection of the *C. trachomatis omp1* gene with SNPs, including both a single SNP and a combination of several SNPs in the bacterial strains with genetic mutations (11 known subtypes in total) had been developed.<sup>6</sup>

The format of GB-speckles had been successfully applied to transform the nucleotide sequences of the genes expressing the serine proteases, the well-known Omptin family proteins within the *Enterobacteriaceae*. These proteins have been found on the surface of several bacterial agents causing different enteric infections, such as salmonellosis, shigelosis, yersiniosis, and escherichiosis.<sup>7</sup> Further, the phase and the relevant two-dimensional distributions of the intensity of GB-speckles in various strains of viral pathogens, namely of lumpy skin disease virus of cattle, LSDV, and also for sheep-pox virus, SPPV have been obtained.<sup>8</sup> Additionally, interference patterns for generated the specific superposition in the relevant fields of GB-speckle and the certain difference in their images have been successfully investigated to reveal a minimal discrimination between the initial viral nucleotide sequences.

A new bioinformatics approach has been proposed very recently:<sup>14</sup> GB-speckles processing via an *s*-LASCA technique (from the spatial Laser Speckle Contrast Analysis) application. As it had been demonstrated, it is possible to extend affectability of the proposed approach comparing to current bioinformatics strategies<sup>15</sup> using *s*-LASCA imaging in the GB-speckles' processing. It had been shown in Ref. 16, that the GB-speckles' generation combined with *s*-LASCA imaging method are very effective to analyze nucleotide polymorphism in several genes of *C. trachomatis*.

This paper is devoted to development of advantageously new technique: the coloured *s*-LASCA imaging of GB-speckles. Such a technique is an improved version of previously suggested "greyscale" *s*-LASCA imaging that was recently

developed especially for GB-speckles. Nucleotide sequences for some target genes SARS–CoV-2 have been successfully processed using coloured *s*-*LASCA*-imaging. Natural mutations in the comparing genes have been reliably and accurately detected.

#### Methods

Nucleotide sequences under consideration.

Seven nucleotide sequences of spike glycoprotein of SARS-CoV-2, namely:

the gene#1. hCoV-19/cat/USA/TX-TAMU-078/2020 (Accession ID: EPI ISL 699509),

the gene#2. hCoV-19/cat/Russia/RII-LEN-22246S/2021 (Accession ID: EPI ISL 811147),

the gene#3. hCoV-19/cat/Greece/2K/2020 (Accession ID: EPI ISL 717979),

the gene#4. hCoV-19/Wuhan/WIV04/2019 (Accession ID: EPI ISL 402124),

the gene#5. hCoV-19/England/QEUH-B11766/2020 (Accession ID: EPI ISL 642476),

the gene#6. hCoV19/South Africa/KRISP-EC-K005299/2020 (Accession ID: EPI ISL 678597),

the gene#7. hCoV-19/Russia/MOS-CRIE-13604226/2020 (Accession ID: EPI ISL 754198).

have been compared on the base of analysis of GB-speckles. The official reference sequences were taken from the GISAID database.

Algorithm for the total conversion of a nucleotide sequence to a colour GB speckle structure, processed by *s*-LASCA imaging technique

#### Initial processing of nucleotide sequence

First, the sequence of the letters derived from the original one-dimensional nucleotide sequence was converted into the sequence of numbers in accordance with the following rule:<sup>4</sup>

$$A \to 1; \ C \to 2; \ G \to 3; \ T \to 4. \tag{1}$$

It is critical to emphasize that the specific relationship between the letters and numbers in this case is not critical as used earlier;<sup>6</sup> thus, other rules could have been applied to the encoding, for instance:

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

or

$$T \to 1; A \to 2; C \to 3; G \to 4. \tag{3}$$

Next, all possible triad combination are generated. As a result, a complete set of all triads is formed:

$$(1\ 1\ 1), (1\ 1\ 2), (1\ 1\ 3), (1\ 1\ 4), (1\ 2\ 1), (1\ 2\ 2), (1\ 2\ 3), (1\ 2\ 4), (1\ 3\ 1), \dots, (4\ 4\ 4).$$
 (4)

The number of all possible combinations of four numbers combined in triads is 64.

Then, a discrete magnitude, h, is allotted to each triad in accordance with the simple algorithm described previously.<sup>4</sup> This algorithm was implemented in Matlab R2015a (**RRID:SCR\_001622**); an open access alternative is Julia. The value of h is a positive integer, varying in the range from 1 to 64. In this case, each triad from the original nucleotide sequence is associated with only one h value. So, for example, the combination (1 1 1) conforms to the value h = 1, (1 1 2) corresponds to h = 2, (1 1 3) conforms to h = 3, (1 1 4) conforms to h = 4, (1 2 1) conforms to h = 5, (1 2 2) conforms to h = 6, and so on. Finally, the latest combination (4 4 4) conforms to the value h = 64. Finally, a square matrix  $H_{n,m}$  was formed by a one-dimensional array h. The physical significance of the shaped matrix  $H_{n,m}$  is that each of its elements represents the local height of some virtual rough surface corresponding to the local content of the analyzed genetic construction. The resulting

virtual rough surfaces could be used to model original speckle structures corresponding to diverse particular nucleotide sequences.

The two-dimensional speckle patterns that corresponded to each specific sequence was generated with the use the diffraction of a coherent beam with a square cross-section profile on a virtual scattering surface with a microrelief described by the matrix  $H_{n,m}$ . At each point of the virtual diffuser (in the beam scattering plane), some phase modulation  $U_{n,m} = \exp(-2\pi j H_{n,m}/64)$  is introduced (*j* is an imaginary unit). The surface is illuminated at the normal incidence of the beam; the phase in the illuminating beam was a constant value.

It is assumed that speckles are formed in the far diffraction zone and described in the Fraunhofer approximation. In this case, the expression for the amplitude of the scattered field is the Fourier transform of the field in the diffraction plane, evaluated at frequencies spaces<sup>10</sup>

$$F_x = X_o/(z * \lambda), \quad F_y = Y_o/(z * \lambda), \tag{5}$$

where  $X_o$  and  $Y_o$  are the coordinates in the observation plane, z is the distance between the scattering plane and the observation plane,  $\lambda$  is the wavelength. The illuminating radiation is completely monochromatic, thus,  $\lambda = \text{const.}$  In this situation, the structure of speckles does not depend on the wavelength and z. Only the sizes of GB-speckles depend on these values, the average size of which is determined by the ratio:

$$d \sim 3 * z * \lambda/a \tag{6}$$

where *a* is the size of the illuminated fragment of virtual surface.<sup>17</sup> It is important to emphasize that the ratio  $\lambda/a$  characterizes the diffraction angular divergence of a laser beam in the far field, and the product of this divergence angle by the light traveled distance *z* is equal to the lateral size of the beam. Thus, it can be seen that the diameter of the undisturbed laser beam (namely, this value is on the right side of the expression (6)) and the average speckle size are approximately equal to each other in any observation plane. In other words, when the parameters *z* and  $\lambda$  change, a proportional change in the size of all speckles occurs synchronously. At the same time, the structure of speckle-patterns in all observation planes are completely similar, only their scale changes from plane to plane, but not the shape of the speckles or their location in the speckle pattern.

#### Generating GB-speckles

The procedure for transcoding the original nucleotide sequence into a GB-speckle structure using the example of the hCoV-19/cat/USA/TX-TAMU-078/2020/2020-07-29 gene (the gene #1) is shown below.

The original nucleotide sequence is as follows:

ATGTTTGTTTTCTTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAACTCAATT ACCCCCTGCATACACTAATTCTTTCACACGTGGTGTTTATTACCCTGACAAAGTTTTCAGATCCTCAGTT TTACATTCAACTCAGGACTTGTTCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATACATGTCTC TGGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGTGTTTATTTTGCTTCC ACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTA CTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTT GGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAA TAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTCAAA AATCTTAGGGAATTTGTGTTTAAGAATATTGATGGTTATTTTAAAATATATTCTAAGCACACGCCTATTA ATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAA CATCACTAGGTTTCAAACTTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTT GGACAGCTGGTGCTGCAGCTTATTATGTGGGTTATCTTCAACCTAGGACTTTTCTATTAAAATATAATGA AAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCAGAAGCAAAGTGTACGTTGAA TAGATTTCCTAATATTACAAACTTGTGCCCTTTTGGTGAAGTTTTTAACGCCACCAGATTTGCATCTGTTT ATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCCTATATAATTCCGCATCATT TTCCACTTTTAAGTGTTATGGAGTGTCTCCTACTAAATTAAATGATCTCTGCTTTACTAATGTCTATGCA GATTCATTTGTAATTAGAGGTGATGAAGTCAGACAAATCGCTCCAGGGCAAACTGGAAAGATTGCTGA TTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAATTCTAACAATCTTGATTCT AAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAGAGA GATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTAC 

CTTTCTTTTGAACTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAA ACAAATGTGTCAATTTCAACTTCAATGGTTTAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGT TTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGATCCACAGACAC TTGAGATTCTTGACATTACACCATGTTCTTTTGGTGGTGTCAGTGTTATAACACCAGGAACAAATACTT CTAACCAGGTTGCTGTTCTTTATCAGGGTGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCA ACTTACTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGTGCAGGCTGTTTAATA GGGGCTGAACATGTCAACAACTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTAT CAGACTCAGACTAATTCTCCTCGGCGGGCACGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATG TCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGCCATACCCACAAATTTTACTATTA GTGTTACCACAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTG GTGATTCAACTGAATGCAGCAATCTTTTGTTGCAATATGGCAGTTTTTGTACACAATTAAACCGTGCTTT AAACACCACCAATTAAAGATTTTGGTGGTTTTAATTTTTCACAAATATTACCAGATCCATCAAAAACCAA GCAAGAGGTCATTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAAC AATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACCTCATTTGTGCACAAAAGTTTAACGGCCTTAC TGTTTTGCCACCTTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATC ACTTCTGGTTGGACCTTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATAGG AGTGCTATTGGCAAAATTCAAGACTCACTTTCTTCCACAGCAAGTGCACTTGGAAAACTTCAAGATGTG GTCAACCAAAATGCACAAGCTTTAAACACGCTTGTTAAACAACTTAGCTCCAATTTTGGTGCAATTTCA AGTGTTTTAAATGATATCCTTTCACGTCTTGACAAAGTTGAGGCTGAAGTGCAAATTGATAGGTTGATC TTCTGCTAATCTTGCTGCTACTAAAATGTCAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTGT GGAAAGGGCTATCATCTTATGTCCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTT ATGTCCCTGCACAAGAAAAGAACTTCACAACTGCTCCTGCCATTTGTCATGGAAAAGCACACTTTC CTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTGTAACACAAAGGAATTTTTATGAACCAC AAATCATTACTACAGACAACACATTTGTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACA CAGTTTATGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTTAAGAATC ATACATCACCAGATGTTGATTTAGGTGACATCTCTGGCATTAATGCTTCAGTTGTAAACATTCAAAAAG AAATTGACCGCCTCAATGAGGTTGCCAAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACTTGGAA AGTATGAGCAGTATATAAAATGGCCATGGTACATTTGGCTAGGTTTTATAGCTGGCTTGATTGCCATAG TAATGGTGACAATTATGCTTTGCTGTATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTG GATCCTGCTGCAAATTTGATGAAGACGACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACATTACACA TAA (7)

After converting a sequence of letters into a sequence of numbers in accordance with the algorithm described by rule (1) described previously, the nucleotide sequence takes the following form:

422124313113424112141141131332433144444334124124441314423113122213422241244144344114112324124 443124224334314424424421334433121324334324321324414414343334414244211224133124444241441111414 331142414211124424112444131342211221121311424144344131444224114144121112443432224444334311344 324433114424112114244314424113344334334114414114412243414131443444133113424114242111224444313 224344324144214321314211244124224124433234344414424121334424114344444211121234321332434441141

333324311214342112112421414313434312141222144334321334141432324134414213124213124114424224233 3144211243114321321142444434432114143321344444341212114411122343244411243311413243444311211312 2213142214211112211321131334214441443113142412444442112111343121244321314324332442142111211414 334314432244334314144324324131312242144434321211113444112332244124344443221224443242121314311 143144324211412124424321243441323334121142124424334433122444334321334324321441211141221444324 144332111144211312421244424422121321134321244331111244211314343342112211114321211324441112123 244344111211244132422114444334321144421134344441114314142244421234244312111344313324311343211 42132122421433431342442443214343124414342224321211311113112442121124324224322144434214314331 44211331331344131411141444411311421412142122131434431444133431214242433214411432442134431112141111433221433412144433241334444141324332443144322141341143343121144143244432434143122134432 43413443424211333243443442443433142243243243111444314311312312424313221343242111331342111441214 412121411 (8)

As a result of diffraction of coherent beam on the phase screen (the virtual heights of the irregularities presented in the table (6)) with a square cross-section is formed GB-speckle-structure of two-dimensional intensity distribution, see Figure 1a. Two-dimensional phase distribution GB, the speckle structure is shown in Figure 1b.

Important to emphasize, that experimental studies were not carried out in this work, only computer modeling. The scheme for calculating GB-speckles during radiation diffraction on a virtual scattering surface is described in detail in the work.<sup>9</sup>

#### s-LASCA imaging of GB-speckles

*s-LASCA* strategy has been connected for handling of GB-speckles. The strategy of *s-LASCA* is based on the examination of an individual realization of static speckles.<sup>3,16</sup> In this case, the whole realization of the speckle field is divided into square zones; typically, each counting  $5 \times 5$  or  $7 \times 7$  pixels.



Figure 1. GB-speckles, generated for genes #1, #2 and #3: (a) Speckle pattern for 2D intensity distribution, (b) Speckle pattern for 2D phase distribution.

For each zone, the contrast of GB-speckles was calculated using the simplest formula:

$$C = \frac{\sigma_I}{\langle I \rangle}$$

where I was the varying intensity of GB-speckles, changing from point to point;  $\sigma_I$  was the standard deviation of the intensity of fluctuations. After the contrast C is calculated in each point, LASCA image is developed. Here, the size of subarea for the local contrast calculating was  $2 \times 2$  pixels. As it has been demonstrated<sup>14</sup> this size of subarea is close to optimal.

#### Coloured GB speckles

To generate three two-dimensional implementations of GB speckles built for different genetic sequences, it is necessary to construct a colour image, where each colour component (red, green, and blue) has its own GB speckle structure. When all three speckle structures were totally indistinguishable, the colour images look grey-scale. If the colour components differ from each other, then, as a result, colouring will appear in the image.

In Figure 2a, the coloured speckle-pattern for intensity distribution is presented (the red component obtained for the nucleotide sequence derived from gene #1, the green component corresponded to the nucleotide sequence of gene #2, and blue component was the relevant to gene #3 nucleotide sequence, respectively).

Figure 2a, demonstrates the differences in the initial nucleotide sequences, a slight staining appears in the colour specklepattern structure for a two-dimensional intensity distribution.

In Figure 2b, the coloured speckle-pattern for phase distribution is shown for such nucleotide sequences as: (i) the red component for gene #1, greenfor gene #2, blue for gene #3.

It is quite obvious that in the case under consideration, there is a pronounced colouring over the whole image for the field of GB-speckle.

Thus, the obtained colour image for the intensity and phase of GB speckles is a reliable diagnostic sign of the presence of polymorphism.

#### A novel detection technique based on the s-LASCA images with coloured GB-speckles

Once an s-LASCA image is obtained for each of the three components of the matched genetic sequence, the final colour image can be constructed. An example of such an image is shown in Figure 3a.

(a)



Figure 2. Coloured GB speckles, generated for genes #1, #2 and #3: (a) Coloured speckle-pattern for 2D intensity distribution, (b) Coloured speckle-pattern for 2D phase distribution.



Figure 3a. Coloured *s-LASCA* images of GB-speckles, generated for three SARS–CoV-2 genes (the gene#1, the gene#2 and the gene#3).



Figure 3b. Coloured *s-LASCA* images of GB-speckles, generated for other three SARS–CoV-2 genes (the gene#4, the gene#5 and the gene#6).

#### **Results and discussion**

It is obvious that the image shown in Figure 3a in comparison with the image in Figure 2a has a more pronounced colouring over the entire field of view, but is characterized by a higher contrast. From a quantitative point of view, the degree of colouring can be described by the value

$$R = \frac{1}{N \times M} \sum_{i=1}^{N \times M} \sqrt{\frac{1}{3} \times \frac{(Ir_i - Io_i)^2 + (Ig_i - Io_i)^2 + (Ib_i - Io_i)^2}{(Io_i)^2}}$$

where  $Ir_i$ ,  $Ig_i$  and  $Ib_i$  are values of intensity for the red, green, and blue components in each pixel,

$$Io_i = \frac{Ir_i + Ig_i + Ib_i}{3}$$



Figure 3c. Coloured *s-LASCA* images of GB-speckles, generated for more three SARS–CoV-2 genes (the gene#5, the gene#6 and the gene#7).

is the average intensity value in each pixel, *i* is the pixel number, *M* and *N* are the number of rows and columns of the analyzed image,  $N \times M$  is the total number of pixels in the image.

Obviously, if the nucleotide sequences compared using *s*-LASCA imaging of GB-speckles are completely identical, then the three components of the resulting colour image will be identical, and therefore the value of R will be equal to zero. However, if there are at least minimal differences in the compared nucleotide sequences, then the value of R will take a positive value. Thus, the value of R calculated for the Figure 3a is 0.1 (gene#1, gene#2 and gene#3 are compared).

In Figure 3b, comparison of new SARS–CoV-2 genes: hCoV-19/Wuhan/WIV04/2019/2019-12-30 (gene#4), hCoV-19/ England/QEUH-B11766/2020/2020-11-02 (gene#5) and hCoV19/South Africa/KRISP-EC-K005299/2020/2020-11-19 (gene#6) is presented. *R* calculated for Figure 3b equals to 0.596.

The physical meaning of the introduced parameter R is that this parameter characterizes the degree of coloring of the picture (GB-speckle- pattern). The bioinformatic (molecular biology) value of R is that it takes positive values, even in the case of the appearance of a one SNP in the analyzed nucleotide sequences. Thus, the minimum natural mutations of the virus can be determined using the parameter R.

Finally, three SARS–CoV-2 genes are reflected in Figure 3c (hCoV-19/England/QEUH-B11766/2020/2020-11-02 (gene#5), hCoV19/South Africa/KRISP-EC-K005299/2020/2020-11-19 (gene#6) and hCoV-19/Russia/MOS-CRIE-13604226/2020/2020-11-09 (gene#7). Again, R equals to 0.596 for this case.

It is important to note that the value of R calculated for Figure 2a and Figure 2b (coloured bare GB-speckle) equals to 0.049 and 0.026, respectively. This means that the value of R at least in two times higher for GB speckles, processed by *s*-LASCA imaging technique.

Evidently, R is positive for all images in Figures 3; so, R is an important diagnostic feature when detecting the presence of SNPs in SARS–CoV-2 genes. This is the main result of this paper.

#### Conclusion

A fundamentally new bioinformatics technique for reliable detection of single SNPs is proposed. The new method is based on the applying of the *s*-*LASCA* 'imaging technique' generating original GB-speckles. It is established that even one SNP can be reliably detected. It has been demonstrated that suggested technique is very effective tool for discrimination between different variants of the SARS–CoV-2 spike glycoprotein gene.

#### Data availability

Underlying data

GISAID Gene: hCoV-19/cat/USA/TX-TAMU-078/2020. Accession number EPI ISL 699509;

GISAID Gene: hCoV-19/cat/Russia/RII-LEN-22246S/2021. Accession number EPI ISL 811147;

GISAID Gene: hCoV-19/cat/Greece/2K/2020. Accession number EPI ISL 717979;

GISAID Gene: hCoV-19/Wuhan/WIV04/2019. Accession number EPI ISL 402124;

GISAID Gene: hCoV-19/England/QEUH-B11766/2020. Accession number EPI ISL 642476;

GISAID Gene: hCoV19/South Africa/KRISP-EC-K005299/2020. Accession number EPI ISL 678597;

GISAID Gene: hCoV-19/Russia/MOS-CRIE-13604226/2020. Accession number EPI ISL 754198.

Sequences are available after registration at the GISAID public database.

#### References

- Dainty C: Laser speckle and related phenomena. Topics in Applied Physics. 1984; 9.
  PubMed Abstract
- Aizu Y, et al.: Biospeckle method for retinal blood flow analysis: flexible correlation measurements. Optical Diagnostics of Biological Fluids and Advanced Techniques in Analytical Cytology. SPIE Press; 1997; 91-102.
- Briers JD: Laser Doppler, speckle and related techniques for blood perfusion mapping and imaging. *Physiol Meas.* 2001; 22: R35–R66.
  PubMed Abstract | Publisher Full Text
- Ulyanov SS, et al.: Using of methods of speckle optics for Chlamydia trachomatis typing. Proc. SPIE. 2017; 10336: 103360D.
- Ulyanov SS, et al.: 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; 15.
  Publisher Full Text
- Feodorova VA, et al.: Optimization of algorithm of coding of genetic information of Chlamydia. Proc. SPIE. 2018; 10716: 107160Q.
  Publisher Full Text
- Ulyanov SS, et al.: Study of statistical characteristics of GB-speckles, forming at scattering of light on virtual structures of nucleotide gene sequences of Enterobacteria. *Izv. Saratov Univ.* (N. S.), Ser. Physics. 2018; 18: 123–137.
  Publisher Full Text
- 8. Saltykov YV, et al.: Analysis of nucleotide sequences of the gene GPCR of the genus Capripoxvirus representatives using speckle

interferometry of GB speckles and subtracting their images. *Math. mod., comp. and natural exp. in natural sci.* 2020: 2.

- Ulianova OV, et al.: Speckle-interferometry and specklecorrelometry of GB-speckles. Front Biosci (Landmark Ed). 2019; 24: 700–711.
  PubMed Abstract
- 10. Goodman JW: *Introduction to fourier optics*. New York: McGraw Hill Companies; 1988.
- Feodorova VA, et al.: Application of virtual phase-shifting speckleinterferometry for detection of polymorphism in the Chlamydia trachomatis omp1 gene. Proc. SPIE. 2018; 10716: 107160M. Publisher Full Text
- Sleator RD: Digital biology. A new era has begun. Bioengineered. 2012; 3: 311–312.
  PubMed Abstract | Publisher Full Text | Free Full Text
- Freddolino PL, Tavazoie S: The Dawn of Virtual Cell Biology. Cell. 2012; 150: 248–250.
- PubMed Abstract | Publisher Full Text | Free Full Text
- Ulianova OV, et al.: LASCA-imaging of gene-based speckles: application to detection of the gene polymorphism in a bacterial model. Laser Physics Letters. 2020; 17. Publisher Full Text
- 15. Lesk AM: Introduction to Bioinformatics. Oxford University Press; 2002.
- 16. Ulianova OV, *et al.*: **Could LASCA-imaging of GB-speckles be applied for a high discrimination and typing of pathogenic bacteria?** *PloS One.* 2021; **16**.
- 17. Francon M: *La granularite laser (speckle) et ses applications en optique.* Masson, Paris, New York, Barcelone, Milan; 1978.

# **Open Peer Review**

# Current Peer Review Status: 💙 ? 🗸

Version 3

Reviewer Report 16 June 2022

#### https://doi.org/10.5256/f1000research.78849.r119557

© **2022 Zimnyakov D.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



### Dmitry A. Zimnyakov

Department of Physics, Yury Gagarin State Technical University of Saratov, Saratov, Russian Federation

I am very grateful for the chance to take part in the review of the article for your respected journal. The current manuscript is devoted to demonstrating the application of the s-LASKA method for discrimination of different variants of the SARS–CoV-2 spike gene. Fortunately, the authors consider this approach as very attractive and available for precise diagnostics of the infection with the use of a new generation of medical devices. In fact, this can be recognized as the new direction for bioscience.

In my point of view, this is very interesting and important research that contains several new findings. The article is clearly written, conclusions correspond to data obtained, and statistical treatment is appropriate. However, the paper needs minor revision before it can be indexed. The following are my concerns about the article:

- 1. p. 3: 'A new bioinformatics approach has been proposed very recently: 14 GB-speckles processing *via* an *s-LASCA* technique...' via should not be italic;
- 2. p. 7-8: the legends for Figures 1-2 contain no indication for which genes the GB-speckles were generated, as well as it is pointed in the legends for Figures 3a, 3b, and 3c. It is critical for the extended auditorium of readers to clearly understand what is demonstrated in Figures 1-2. These points need a correction.

I have no additional proposals for edits. The revised version of the article can be recommended for indexing with no further reviewing.

# Is the work clearly and accurately presented and does it cite the current literature? Yes

### Is the study design appropriate and is the work technically sound?

Yes

## Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?  $\ensuremath{\mathsf{Yes}}$ 

Are the conclusions drawn adequately supported by the results? Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: laser physics, biomedical science

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 17 Jun 2022

**Valentina Feodorova**, Federal Research Center for Virology and Microbiology, Branch in Saratov, Saratov, Russian Federation

Dear Prof Dmitry A. Zimnyakov, we highly appreciate your review for our manuscript. All your recommendations were accepted. All necessary corrections have been done in the text of the paper. Thank you very much once more. The authors.

Competing Interests: No competing interests were disclosed.

## Version 2

Reviewer Report 20 September 2021

https://doi.org/10.5256/f1000research.76647.r93566

© **2021 Angelsky O et al.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Oleg Angelsky 问

Department of Correlation Optics, Chernivtsi National University, Chernivtsi, Ukraine **Claudia Zenkova** 

Chernivtsi National University, Chernivtsi, Ukraine

The paper can de accepted for indexing in the present new form.

# Is the work clearly and accurately presented and does it cite the current literature? $\ensuremath{\mathsf{Yes}}$

Is the study design appropriate and is the work technically sound? Yes

Are sufficient details of methods and analysis provided to allow replication by others?  $\ensuremath{\mathsf{Yes}}$ 

If applicable, is the statistical analysis and its interpretation appropriate?  $\ensuremath{\mathsf{Yes}}$ 

Are all the source data underlying the results available to ensure full reproducibility?  $\ensuremath{\mathsf{Yes}}$ 

Are the conclusions drawn adequately supported by the results?  $\ensuremath{\mathsf{Yes}}$ 

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Singular and Correlation Optics

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 14 September 2021

#### https://doi.org/10.5256/f1000research.76647.r94278

© **2021 Bashkatov A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ? Alexey Bashkatov

Department of Optics and Biophotonics, Saratov State University, Saratov, Russian Federation

The article is devoted to an extremely relevant and very interesting topic – an application of GBspeckles, in particular, processed by the s-LASCA imaging method, in relation to the discrimination of SARS–CoV-2 strains. GB-speckle is a new word in the field of bioinformatics. In the future, these optical virtual speckle structures can be effectively used as an alternative to classical bioinformatics methods. The usage of GB-speckles can be considered as a replacement for traditional computer methods of sequencing of any nucleotide sequences. As it follows from the analysis of the literature, the authors of this article have successfully used these methods, based on the processing of virtual speckles, to the analysis of natural mutations in the comparing genes.

Also, this article presents a fundamentally new method, based on the using of colored GBspeckles, processed by s-LASCA-imaging technique. The new method is based on the applying of the *s-LASCA* 'imaging technique' generating original GB-speckles. The article demonstrates that colored s-LASCA-imaging can be successfully used for the diagnosis and differentiation of various strains of SARS-CoV-2.

The utilization of GB-speckles in diagnostics is an absolutely new direction, which has been developed and is being successfully promoted by the authors of this article. I have no questions about the application of speckle interferometry methods in molecular biology, reflected in this work. I only have some minor criticisms, regarding bioinformatics:

- The authors need to check carefully the reference(s) and link(s) for the bioinformatic tools used.
- After the providing of clear the answer to this question, the reviewed article can be accepted for indexing without any additional changes and without the additional cycles of reviewing.

Is the work clearly and accurately presented and does it cite the current literature? Partly

Is the study design appropriate and is the work technically sound?  $\ensuremath{\mathsf{Yes}}$ 

Are sufficient details of methods and analysis provided to allow replication by others?  $\ensuremath{\mathsf{Yes}}$ 

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?  $\ensuremath{\mathsf{Yes}}$ 

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 24 Sep 2021

**Valentina Feodorova**, Federal Research Center for Virology and Microbiology, Branch in Saratov, Saratov, Russian Federation

The authors are extremely grateful to the respected reviewer Alexey Bashkatov for a detailed analysis of our article. In accordance with the recommendations of the reviewer, all references and links are checked, thank you.

*Competing Interests:* No competing interests were disclosed.

## Version 1

Reviewer Report 16 July 2021

#### https://doi.org/10.5256/f1000research.56573.r88324

© **2021 Angelsky O.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ? Oleg Angelsky 🗓

Department of Correlation Optics, Chernivtsi National University, Chernivtsi, Ukraine

The paper is devoted to the investigations of virtual Gene Based speckles for the determination the difference between SARS-Cov-2 spike glycoprotein genes. As it was noted by the authors the given method can use the classical approaches of speckle-optics. But there are some questions:

- 1. Any optical speckle is the result of the interference of waves scattered by inhomogeneities, and to obtain a speckle pattern, the wavelength is decisive. Thus, the question arises about the concrete wavelengths used in the study and their relationships with pseudo inhomogeneities. It would be interesting to know about the coherence of the sources used the modelling.
- 2. Please present an experiment setup, and the results of experimental modelling.
- 3. It would be interesting to conduct a complete analysis of the entered parameter R, to represent its physical, biochemical meaning. The virus is constantly mutating, how is it possible to estimate the degree of virus mutation using this parameter.

The paper can be indexed after reworking.

# Is the work clearly and accurately presented and does it cite the current literature? Partly

## Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others? Yes

If applicable, is the statistical analysis and its interpretation appropriate?  $\ensuremath{\mathsf{Yes}}$ 

Are all the source data underlying the results available to ensure full reproducibility?  $\ensuremath{\mathsf{Yes}}$ 

Are the conclusions drawn adequately supported by the results? Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Singular and Correlation Optics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 30 Jul 2021

**Valentina Feodorova**, Federal Research Center for Virology and Microbiology, Branch in Saratov, Saratov, Russian Federation

Reply to the comments of reviewer Prof. Oleg Angelsky:

- Any optical speckle is the result of the interference of waves scattered by inhomogeneities, and to obtain a speckle pattern, the wavelength is decisive. Thus, the question arises about the concrete wavelengths used in the study and their relationships with pseudo inhomogeneities. It would be interesting to know about the coherence of the sources used the modelling.
- **Response**: This is really reasonable and very important question. Fragment below is added to the new version of the text:

It is assumed that speckles are formed in the far diffraction zone and described in the Fraunhofer approximation. In this case, the expression for the amplitude of the scattered field is the Fourier transform of the field in the diffraction plane, evaluated at frequencies spaces

Fx=Xo/( z\* λ), Fy=Yo/( z\* λ),

(\*)

where Xo and Yo are the coordinates in the observation plane, z is the distance between the scattering plane and the observation plane,  $\lambda$  is the wavelength. The illuminating radiation is completely monochromatic, thus,  $\lambda$  =const.

Reference to Goodman, J. W. Introduction to fourier optics. *McGraw Hill Companies, New York* (1988).

In this situation, the structure of speckles does not depend on the wavelength and z. Only the sizes of GB-speckles depend on these values, the average size of which is determined by the ratio:

d~3\*z\* λ /a

where a is the size of the illuminated fragment of virtual surface. It is important to emphasize that the ratio  $\lambda$ /a characterizes the diffraction angular divergence of a laser beam in the far field, and the product of this divergence angle by the light traveled distance z is equal to the lateral size of the beam.

New reference: M. Francon. La granularite laser (speckle) et ses applications en optique. Masson, Paris, New York, Barcelone, Milan 1978.

Thus, it can be seen that the diameter of the undisturbed laser beam (namely, this value is on the right side of the expression (\*\*)) and the average speckle size are approximately equal to each other in any observation plane.

In other words, when the parameters z and  $\lambda$  change, a proportional change in the size of all speckles occurs synchronously. At the same time, the structure of speckle-patterns in all observation planes are completely similar, only their scale changes from plane to plane, but not the shape of the speckles or their location in the speckle pattern.

- Please present an experiment setup, and the results of experimental modelling.
- Response: Experimental studies were not carried out in this work, only computer modeling. The scheme for calculating GB speckles during radiation diffraction on a virtual scattering surface is described in detail in the work: Ulianova OV, et al.: Speckle-interferometry and speckle-correlometry of GB-speckles. Frontiers in Bioscience-Landmark 2019; 24, 700-711. This fact is mentioned in the new version of this article.
- It would be interesting to conduct a complete analysis of the entered parameter R, to represent its physical, biochemical meaning. The virus is constantly mutating, how is it possible to estimate the degree of virus mutation using this parameter.
- Response: The physical meaning of the introduced parameter R is that this parameter characterizes the degree of coloring of the picture (GB-speckle- pattern). The bioinformatic (molecular biology) value of R is that it takes positive values, even in the case of the appearance of a one SNP in the analyzed nucleotide sequences. Thus, the minimum natural mutations of the virus can be determined using the parameter R. This circumstance is also noted in the Conclusions of new version of this article.

The authors are extremely grateful to the respected reviewer, whose critical comments have significantly improved the quality of this article.

*Competing Interests:* No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000 Research