Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Short communication

Biometric analysis and identification of single-stranded RNA sequences recognized by TLR7/8 in the SARS-CoV-2, SARS-CoV, and MERS-CoV genomes

Mario Adán Moreno-Eutimio a, *, Constantino López-Macias b, c, Rodolfo Pastelin-Palacios a

a Facultad de Química, Universidad Nacional Autónoma de México (UNAM), Ciudad Universitaria, CP 04510, Mexico
b Unidad de Investigación Médica en Inmunología, Hospital de Especialidades, Centro Médico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Ciudad de México, CP 06500, Mexico
c Visiting Professor of Immunology, Nuffield Department of Medicine, University of Oxford, UK

Article info

Article history:
Received 17 March 2020
Accepted 23 April 2020
Available online 30 April 2020

Keywords:
SARS-CoV-2
SARS
MERS
TLR7
TLR8

A B S T R A C T

During virus infection, host toll-like receptors (TLRs) can recognize different pathogen-associated molecular patterns and trigger the innate immune response. TLR7/8 can identify the single-stranded RNA (ssRNA) of the virus. This study aimed to search ssRNA sequences recognized by TLR7/8 from the SARS-CoV-2, SARS-CoV, and MERS-CoV whole genomes by a bioinformatic technique. The immunoinformatic approach showed that the SARS-CoV-2 genome has more ssRNA fragments that could be recognized by TLR7/8 than the SARS-CoV genome. These findings suggest innate immune hyperactivation by SARS-CoV-2. This activity is possibly able to provoke a robust proinflammatory response via TLR7/8 recognition and cause acute lung injury.

© 2020 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

The novel coronavirus (SARS-CoV-2) represents a public health emergency of international concern [1]. As of 10th March 2020, the death toll from the novel coronavirus stood at 4,012, with more than 113,702 confirmed cases in China, as well as cases in 109 other countries [2]. Coronaviruses are single, positive-sense RNA viruses belonging to the family Coronaviridae, which includes Middle East respiratory syndrome coronavirus (MERS-CoV) and severe acute respiratory syndrome coronavirus (SARS-CoV) [3].

Viral interactions with the host immune system play a central role in the outcome of infection. In the initial phase, recognition of evolutionarily conserved microbial structures, known as pathogen-associated molecular patterns (PAMPs), is an essential function of the innate immune system [4]. Germ line-encoded pattern recognition receptors (PRRs) are proteins expressed by a variety of cells and are responsible for sensing the presence of PAMPs. Sensing of PAMPs by PRRs markedly upregulates the transcription of genes involved in inflammatory responses [4].

Toll-like receptors (TLRs) belong to a conserved family of innate immune recognition receptors acting as the primary sensors of specific PAMPs expressed by numerous pathogens [5]. The human TLR family comprises 11 members, of which TLR-3, -7, and -8 are essential in recognition of structural components of RNA viruses [6–8]. TLR7 senses single-stranded RNA (ssRNA) oligonucleotides containing guanosine- and uridine-rich sequences from RNA viruses [6]. Recognition occurs in the endosomes of plasmacytoid dendritic cells (DCs) and B cells. TLR8 is phylogenetically and functionally closely related to TLR7 and recognizes ssRNA. It is preferentially expressed in myeloid DCs and monocytes [7].

In this study, we searched ssRNA fragments in the whole genome from SARS-CoV, MERS-CoV, and SARS-CoV-2 (from different geographical origins mainly from Germany [SARS-CoV-2/Germany] and Wuhan [SARS-CoV-2/Wuhan]) by a bioinformatics scanning technique to reveal important TLR7/8 recognition sites in the whole genome of these.

1. Materials and methods

1.1. Bioinformatics analysis of ssRNA sequences

All genomic sequences were collected on 31 January 2020 from GenBank or Gisaid [9]. Those of SARS coronavirus NC_004718.3; Middle East respiratory syndrome coronavirus NC_019843.3;
Crystallographic, biophysical, and cellular data have established ssRNA sequence preferences for TLR7. The UU(U/C) motif fully binds to TLR7, followed by the UU(G/A) motif (moderate binding) [11]. Therefore, we investigated the number of UUU, UUC, UUG, and UUA ssRNA fragments in the whole genome sequences of SARS-CoV-2/Wuhan, SARS-CoV-2/Germany, SARS-CoV and MERS-CoV. The four virus genomes analyzed showed similar amounts of U and G nucleotides, and the full lengths of each genome are comparable (Table 1). The SARS-CoV-2/Wuhan genome presented 708 UUU ssRNA fragments, which represents 17.4% more fragments than the SARS-CoV genome (1.15-fold change) and 8.1% more than the SARS-CoV-2/Germany genome (1.17-fold change) and 5.3% (UUG; 1.05-fold change) more ssRNA fragments in the SARS-CoV genome (1.15-fold change) (Table 1 and Fig. 1).

Table 1

| Genome                        | SARS-CoV | MERS-CoV | SARS-CoV-2/Wuhan | SARS-CoV-2/Germany |
|-------------------------------|----------|----------|------------------|-------------------|
| Total nucleotides (bp)        | 29,751   | 30,119   | 29,899           | 29,782            |
| G% (bp)                       | 19.97 (5940) | 20.56 (6116) | 18.46 (5492)    | 19.71 (5865)      |
| C% (bp)                       | 19.97 (5940) | 20.56 (6116) | 18.46 (5492)    | 19.71 (5865)      |
| U% (bp)                       | 30.73 (9143) | 32.94 (9799) | 32.42 (9593)    | 32.16 (9567)      |
| UUU % (bp)                    | 30.73 (9143) | 32.94 (9799) | 32.42 (9593)    | 32.16 (9567)      |
| UUC % (bp)                    | 19.97 (5940) | 20.56 (6116) | 18.46 (5492)    | 19.71 (5865)      |
| UUG % (bp)                    | 20.80 (6187) | 21.19 (6304) | 19.71 (5865)    | 19.67 (5851)      |
| UUA % (bp)                    | 10.56 (3126) | 10.80 (3242) | 13.24 (4066)    | 13.11 (4045)      |

The number of oligonucleotide fragments recognized by TLR7

**Motif fully bound**

| Oligonucleotide | SARS-CoV | MERS-CoV | SARS-CoV-2/Wuhan | SARS-CoV-2/Germany |
|-----------------|----------|----------|------------------|-------------------|
| UUU             | 603      | 655      | 708               | 706               |
| (UUU)2          | 2        | 4        | 4                | 4                 |
| UUC             | 563      | 581      | 518               | 513               |
| (UUC)2          | 14       | 15       | 3                | 3                 |
| (UUC)3          | 1        | 0        | 3                | 3                 |
| UUA             | 3126     | 3242     | 4066             | 4045              |
| (UUA)2          | 30       | 35       | 25               | 25                |
| (UUA)3          | 0        | 0        | 0                | 0                 |
| UUG             | 791      | 846      | 817              | 817               |
| (UUG)2          | 1        | 1        | 1                | 1                 |
| (UUG)3          | 1        | 0        | 0                | 0                 |

**Motif moderately bound**

| Oligonucleotide | SARS-CoV | MERS-CoV | SARS-CoV-2/Wuhan | SARS-CoV-2/Germany |
|-----------------|----------|----------|------------------|-------------------|
| UUA             | 791      | 846      | 817              | 817               |
| (UUA)2          | 30       | 35       | 25               | 25                |
| (UUA)3          | 0        | 0        | 0                | 0                 |
| UUG             | 796      | 846      | 817              | 817               |
| (UUG)2          | 31       | 22       | 22               | 22                |
| (UUG)3          | 0        | 0        | 0                | 0                 |

3. Discussion

The SARS-CoV-2, SARS-CoV, and MERS-CoV infections show several similarities regarding the clinical presentations, which can vary from asymptomatic infection to severe disease [13]. Additionally, a cytokine-storm has been observed in the rapid course of SARS [14,15] and MERS [16,17], and in the serum of SARS-CoV-2-infected patients, proinflammatory cytokines are upregulated [16]. It was inferred that an overactive innate immune response could contribute to virus-induced immune pathology resulting in acute lung injury in the infected patients. Therefore, in this study, based on the essential structural feature of the PAMP-PRR complex, specifically the interaction of TLR7/TLR8 with ssRNA [11,12], we searched ssRNA fragments with a pathogenic molecular pattern from the SARS-CoV-2, MERS-CoV and SARS-CoV whole genomes.
Our bioinformatic analysis showed that the SARS-CoV-2 genome contains a large number of fragments that could be recognized by TLR7/8, and it even contains more fragments than the SARS-CoV genome. This result suggests the ability to induce a rapid type I interferon response [7]. The production of type I IFNs (primarily alpha IFN [IFN-α] and IFN-β) plays a central role in the induction of antiviral responses. Type I IFNs influence protein synthesis, growth regulation, and apoptosis and enhance the maturation of DCs, the cytotoxicity of natural killer cells, and the differentiation of virus-specific cytotoxic T lymphocytes [19]. However, innate immune hyperactivation could be involved in the dysregulation of a series of proinflammatory cytokines during viral infection [20].

The recognition of ssRNA fragments by TLR7/8 could depend on the virus replication manner, so that a more robust proinflammatory response should occur after a latent period, except for the host immune cell taking up viral particles. This effect is feasible because a number of immune cells display distinct susceptibility to MERS-CoV [17] and SARS-CoV [21].

Interestingly, in addition to its expression in immunological cells, TLR7/8 is expressed predominantly in the lung, bronchus, breast, rectum, smooth muscle tissue, cerebral cortex, and kidney [22]. Although TLR7 and TLR8 are phylogenetically and structurally related, TLR7- and TLR8-specific agonists trigger different cytokine induction profiles. TLR7-specific agonists generally induce IFN-regulated cytokines, but TLR8-specific agonists lead primarily to the production of proinflammatory cytokines [23]. These factors may provide SARS-CoV-2 with a shortcut to trigger an innate immune response through TLR7/8 and ultimately contribute to the development of immune pathology within the lungs.

Most likely, additional pathogenic molecular patterns can be recognized by other innate immune receptors, such as SARS-CoV spike protein, which was observed to be a TLR2 ligand [24], and damage-associated molecular patterns (DAMPs) that are released in response to tissue damage from cells killed by viruses [25]. All of these factors also probably contribute to the cytokine storm. The interaction between viral PAMPs and PRRs in immune cells plays an essential survival role in the response to viral infections but may be simultaneously responsible for tissue injury associated with severe virus-induced inflammation.

For positive-sense RNA viruses such as coronavirus, it is known that viral genomic ssRNA is recognized by either the endosomal RNA receptors (TLR7/TLR8), and the cytosolic RNA sensors, such as retinoid-inducible gene-1 (RIG1), melanoma differentiation-associated gene 5 (MDA5), laboratory of genetics and physiology 2 (LG2), and cytoplasmic protein kinase R (PKR) [26]. Nevertheless, the consensus definition of cytosolic RNA sensor ligands remains controversial, and the crystal structures with their RNA sequence-specific are not available. For this reason, the cytosolic sensors were not included in the bioinformatic analysis.

In conclusion, the SARS-CoV-2 genome contains more ssRNA fragments that could be recognized by TLR7/8 than the SARS-CoV genome and similar ssRNA fragments than the MERS-CoV genome; possibly making it able to provoke a robust proinflammatory response via TLR7/8 recognition and cause acute lung injury, leading to death. These bioinformatic findings suggest that SARS-CoV-2 plays a crucial role in the overactive innate immune response.
Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 2020;382:727–33.
[2] WHO. Coronavirus disease (COVID-2019) situation report – 50. World Health Organization; 2020.
[3] Masters PS. The molecular biology of coronaviruses. Adv Virus Res 2006;66:193–292.
[4] Medzhitov R, Janeway Jr C. Innate immunity. N Engl J Med 2000;343:338–44.
[5] Imler JL, Hoffmann JA. Toll receptors in innate immunity. Trends Cell Biol 2001;11:304–11.
[6] Diebold SS, Kaisho T, Hemmi H, Akira S, Reis e Sousa C. Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. Science 2004;303:1529–31.
[7] Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data – from vision to reality. Euro Surveill 2017;22.
[8] Maca F, Upton C. Sequence Searcher: a Java tool to perform regular expression and fuzzy searches of multiple DNA and protein sequences. BMC Res Notes 2009;2:14.
[9] Zhang Z, Ohno U, Shibata T, Taoka M, Yamauuchi Y, Sato R, et al. Structural analyses of toll-like receptor 7 reveal detailed RNA sequence specificity and recognition mechanism of agonistic ligands. Cell Rep 2018;25:3371–3378 e5.
[10] Tanji H, Ohno U, Shibata T, Taoka M, Yamauuchi Y, Isobe T, et al. Toll-like receptor 8 senses degradation products of single-stranded RNA. Nat Struct Mol Biol 2015;22:109–15.