

## RESEARCH ARTICLE

# Typing of alpha and beta coronaviruses by DNA barcoding of NSP12 gene

Waleed A. Nembr<sup>1</sup>  | Nashwa K. Radwan<sup>2</sup>

<sup>1</sup>Department of Radiation Microbiology, National Center for Radiation Research and Technology (NCRRT), Egyptian Atomic Energy Authority (EAEA), Cairo, Egypt

<sup>2</sup>Department of Health Radiation Research, National Center for Radiation Research and Technology (NCRRT), Egyptian Atomic Energy Authority (EAEA), Cairo, Egypt

**Correspondence**

Waleed A. Nembr, Department of Radiation Microbiology, National Center for Radiation Research and Technology (NCRRT), Egyptian Atomic Energy Authority (EAEA), Cairo, Egypt.  
Email: [wnebr@hotmail.com](mailto:wnebr@hotmail.com)

**Abstract**

Since the spread of the COVID-19 pandemic, the world paid attention to coronaviruses (CoVs) evolution and their diverged lineages because many researches studies supposed that the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is evolutionarily developed from a lineage of bats CoVs. This is due to the ability of some mutant CoVs to transmit from a host to different hosts. For this reason, there are many fears about the pathogenicity of the upcoming variants of CoVs. Thus, it is important to get a rapid and economic technique for typing a wide range of human and animal CoVs species for following up their mutant transmission. Therefore, the present study aims at approaching a simple design of DNA barcoding of a wide range of mammals' CoVs (including alpha and beta CoVs), by universal amplification of a species-specific sequence inside a conserved gene (NSP12) followed by amplicon sequencing. The *in silico* evaluation involved 96 nucleotide sequences of different CoVs (18 alpha CoVs and 78 beta CoVs), and was applied experimentally into the lab on 5 human CoVs isolates; 3 of them belong to beta CoVs (OC43, MERS, and SARS-CoV-2) and 2 are alpha CoVs (229E and NL63). The results indicated that the designed universal primers are able to amplify 332 bp of a taxonomic region inside the NSP12 coding sequence that facilitates the identification and classification of mammals' CoVs upon the resulting phylogenetic tree.

**KEYWORDS**

coronaviruses, DNA bar code, molecular typing, taxonomy

## 1 | INTRODUCTION

Coronaviruses (CoVs) are a group of RNA viruses that can infect a wide range of mammals and birds. The coronavirus particle consists of a single-stranded positive-sense RNA up to 32 kbp encapsulated by nucleocapsid protein (N) which is embedded into three outer canonical structural proteins namely, envelop (E), membrane (M), and spike glycoprotein (S).<sup>1</sup>

CoVs belong to *Orthornavirae*, subdivide into four genera; two of them infect mammals (alphacoronaviruses and beta coronaviruses), while gammacoronaviruses and delta coronaviruses infect birds. Only two strains of alpha CoVs are able to infect the human respiratory tract and cause mild symptoms of the common cold (human coronavirus NL63 and human coronavirus 229E).

Beta CoVs are subdivided into five subgenera according to the International Committee on Taxonomy of Viruses (ICTV); *Embecovirus* (lineage A) *Sarbecovirus*, (lineage B) *Merbecovirus*, (lineage C) *Nobecovirus*, (lineage D), and *Hibecovirus*. There are five strains of beta CoVs associated with human respiratory tract infection (human coronavirus [hCoVs]) causing symptoms ranging from mild to severe and may lead to death. Two of them belong to *Embecovirus*, known as HKU1 and OC43. Another two strains belong to *Sarbecovirus*, known as SARS and SARS-CoV-2, and the last one belongs to *Merbecovirus*, known as Middle East respiratory syndrome coronavirus (MERS). There are common symptoms among hCoVs include common cold, fever, sore throat, and pneumonia (which may result from direct viral infection or by secondary bacterial infection).<sup>2</sup>

Each strain of hCoVs was developed by multiple complicated mutations events during evolution. This is due to a higher error rate of copying RNA genome during viral multiplications into the host cell. Hence, RNA viruses have a better ability to generate variants than DNA viruses. The most important mutations are those occurring in genes coding for the outer structural proteins, such as S protein, which is the responsible protein for binding with the specific host cell receptor. This increases the chance of different host transmission for the newly generated variants.<sup>3</sup> The evolution reports of SARS-CoV-2, which caused coronavirus disease-2019 (COVID-19) pandemic, supposed its relation to bat CoVs lineages as a genetic reservoir that transmitted to humans after chances of series mutations.<sup>4</sup>

Furthermore, many single nucleotide variants have diverged from the first identified isolate of SARS-CoV-2 from Wuhan, China (NC\_045512.2). Hence, it was expected that further CoVs variants will diverge, and they may get the chance to acquire different pathogenicity and immunogenicity characteristics as compared to the originated isolates. So, it is important to follow up the transmission of different geographic isolates of diverged CoVs, especially for species closely related to humans and bats. Full genome sequencing is the traditional technique to search for CoVs variants. However, it is an expensive technique and difficult to apply to low-income labs for surveying a large number of samples. Therefore, it is important to find out a simple and less expensive technique to diagnose and identify different types of CoVs as a preliminary step to forward the new isolates for full genome sequencing.

Nucleotide sequences of a variety of conserved genes among related organisms are targeted for phylogenetic and identification studies. In detail, this nucleotide variation is used to distinguish different holotypes of an organism by a simple method termed DNA barcoding, which is based on PCR- amplicon sequencing using universal primers. The validity of DNA barcoding protocol depends on obtaining distinguishable reference sequences for each holotype and confirming the taxonomy of the related organisms.<sup>5</sup>

Generally, according to information announced by the ICTV, all *Coronaviridae* members shared the genetic structure and organization of the replicase gene, which is encoding for essential proteins involved in the viral replication processes into the host cell. The structure of this gene is a complex of 2 overlapped open reading frames (ORFs); 1a and 1b. The translation products of them comprised 2 polyproteins (pp1a and pp1ab) that are subsequently cleaved by posttranslational processes to produce 16 mature products, called nonstructural proteins (NSPs). One of these proteins is RNA-dependent RNA polymerase (RdRp) which is comprised of NSP12. Moreover, ICTV cited that, the nucleotide sequence encoding for NSP12 is considered one of the most important taxonomic elements to classify CoVs according to the phylogeny. This means that the NSP12 gene contains variable nucleotide sequences that can distinguish the related and the unrelated species to an unknown isolate, depending on the estimated genetic distance of this gene.

Thus, the present study aims at evaluating alpha and beta CoVs genomes, especially in the region of NSP12, to extract a conserved gene containing species-specific barcoding sequence for typing these CoVs with respect to their classification phylogeny, to facilitate the rapid identification of new CoV isolates.

## 2 | MATERIALS AND METHODS

### 2.1 | In silico investigations and experimental design

The design of the presented DNA barcoding technique is based on targeting amplicons containing species-specific sequences by locating universal PCR primers on the nearby homologous flanking regions. Firstly, the genomic sequence data of six different species of hCoVs were retrieved from The National Center for Biotechnology Information (NCBI) database. These six species include Human coronavirus 229E (KY996417.1), Human coronavirus NL63 (AY567487.2), Human coronavirus OC43 (AY391777.1), MERS (KF192507.1), severe acute respiratory syndrome coronavirus (SARS-CoV) (AY559086.1), and SARS-CoV-2 (NC\_045512.2). Then, multiple sequence alignment was carried out by the BLAST tool to determine regions that exhibit similarities among different types of hCoVs, to eliminate high mutation rate sequences from the analysis.

Consequently, inside the selected regions, well-aligned sequences were evaluated to locate universal primers matching with homologous flanking regions surrounding the species-specific sequences, using BioEdit sequence alignment software.<sup>6</sup> Candidate taxonomic amplicon sequences were evaluated to be distinguishable for each species by estimating the difference between the max score and the bit score of each aligned sequence, compared to the same region of other CoVs lineages (only sequences with 100% coverage), using a bit-scoring system of BLAST web-based tool (<https://blast.ncbi.nlm.nih.gov/>). Furthermore, sequence variability among different CoVs amplicons (involved hCoVs and the other related CoVs resulted by BLAST) was evaluated by estimating the entropy values of each nucleotide site, using BioEdit software.<sup>6</sup>

The selected PCR primers were validated by in silico amplification of targeted CoVs genomes, using NCBI Primer-BLAST web-based tool.<sup>7</sup> The produced amplicon sequences of 96 different holotypes of alpha and beta CoVs were submitted for phylogenetic analysis using the Maximum Composite Likelihood method<sup>8</sup> conducted in MEGA6 software.<sup>9</sup>

The unintended matching of the selected primers was investigated using the Primer-BLAST tool against the nucleotide sequence of the human genome and other human viruses (which associated with the respiratory tract) such as Adenoviridae (taxid:10508), Influenza A virus (taxid:11320), Influenza B virus (taxid:11520), Human respirovirus 3 (human parainfluenza virus, taxid:11216), Enterovirus (Rhinoviruses, taxid:12059), Respiratory syncytial virus (taxid:12814), Human bocavirus (taxid:329641), and Metapneumovirus (taxid:162387).

### 2.2 | Real-time PCR experiment

Previously identified RNA samples of 5 human CoV species (NL63, 229E, OC43, MERS, and SARS-CoV-2) were obtained from the Microbiology Reference Lab (as mentioned in the acknowledgment



**FIGURE 1** BLAST graphic viewer showed the whole-genome alignment (upper part) of SARS-CoV-2 (NC\_045512.2) with Severe acute respiratory syndrome coronavirus (AY559086.1), Middle East respiratory syndrome coronavirus (KF192507.1), Human coronavirus OC43 (AY391777.1), Human coronavirus NL63 (AY567487.2), and Human coronavirus 229E (KY996417.1). Red regions indicate aligned sequences with a score >200. Delimited region with forward (Fd) and reverse (Rv) arrows, which indicate primers directions of the PCR amplicon sequence (332 bp), was focused for nucleotide resolution of sequence alignment at the lower part of the Figure

**TABLE 1** Data analysis of in silico PCR results of the designed primers targeted alpha and beta-human coronaviruses

Target organism (abbreviation)	NCBI accession number	CoV classification		PCR amplicon size/ GC%	PCR amplicon range
		Genus	Sub-genus		
Human coronavirus 229E (229E)	KY996417.1	Alpha	Duvinacovirus	332/38	14 024–14 355
Human coronavirus NL63 (NL63)	AY567487.2	CoVs	Setracovirus	332/36	13 927–14 258
Severe acute respiratory syndrome coronavirus (SARS)	AY559086.1	Beta	Sarbecovirus	332/39	14 882–15 213
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)	NC_045512.2	CoVs	Sarbecovirus	332/37	14 968–15 299
Middle East respiratory syndrome coronavirus (MERS)	KF192507.1		Merbecovirus	332/42	14 934–15 265
Human coronavirus OC43 (OC43)	AY391777.1		Embecovirus	332/38	14 832–15 163

Abbreviations: CoV, coronaviruses; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

section). The reverse transcription step was carried out using FIRE-Script® RT cDNA synthesis kit (product# 06-15; Solis BioDyne), contained 10 ng of RNA sample and 25 Pmol of each primer (sense: 5'-GGHAARGCHVGNBNTAYTAYGA, antisense: 5'-GGRTARTCCC

AACCCAT), followed by PCR amplification using iQ™ SYBR® Green Supermix (Bio-Rad Laboratories, Inc.).

The run of PCR was carried out into Rotor-Gene Q instrument (QIAGEN), the thermal profile was programmed as 95°C for

**TABLE 2** BLAST scoring of different lineages of alpha and beta coronaviral amplicons (332 bp), reflects the resolution of sequence differentiation between inter- and intra-species

Genus	Sub-genus	Lineages	hits <sup>a</sup>	BLAST bit-score(Max score = 600) <sup>b</sup>	
				Intra-species maximum difference score	Inter-species minimum difference score
Alphacoronavirus	Duvinacovirus	Human coronavirus 229E	143	37	113
		229E-related bat coronavirus	43	167	127
		Rousettus aegyptiacus bat coronavirus 229E-related	1	-	118
		Camel alphacoronavirus	26	23	68
	unclassified Alphacoronavirus	Alphacoronavirus sp.	55	379	194
		Bat alphacoronavirus	38	348	271
	Myotacovirus	BtMr-AlphaCoV/SAX2011	1	-	235
	Pedacovirus	Pedacovirus sp.	725	303	23
	Decacovirus	Bat coronavirus HKU10	8	28	100
		unclassified Decacovirus sp.	4	276	109
	Setracovirus	Human coronavirus NL63	118	37	190
	Luchacovirus	Lucheng Rn rat coronavirus	7	253	220
	Minunacovirus	Miniopterus bat coronavirus	22	82	203
		Miniopterus bat coronavirus 1	9	19	28
	Rhinacovirus	BtRf-AlphaCoV/YN2012	1	-	28
		Rhinolophus bat coronavirus HKU2	7	5	14
Betacoronavirus	Embecovirus	Human coronavirus OC43	311	14	32
		Betacoronavirus HKU24	8	86	68
		Human coronavirus HKU1	124	86	149
		Longquan RI rat coronavirus	7	73	95
		Murine coronavirus	48	96	100
		Rabbit coronavirus HKU14	5	10	140
	Sarbecovirus	Bat betacoronavirus	15	84	77
		Severe acute respiratory syndrome-related coronavirus	1005	66	59
		Pangolin coronavirus	7	190	190
	Merbecovirus	Bat coronavirus HKU4	26	32	50
		Middle East respiratory syndrome-related coronavirus	725	239	55
		Pipistrellus bat coronavirus HKU5	28	55	230
	Hibecovirus	Hypsugo bat coronavirus HKU25	2	7	212
		Bat Hp-betacoronavirus/Zhejiang2013	1	-	208

<sup>a</sup>Hits: The number of organisms categorized under the same lineage.

<sup>b</sup>The estimation limited to sequences have 100% coverage of the amplicon range.

15 min, 40 cycles of 95°C (30 s) –55°C (20 s) –72°C (40 s), and the fluorescent reading was acquired by the green channel at the end of each cycle. High resolution melting (HRM) analysis was carried out at the end of PCR cycles, to visualize the melt

profile of each PCR product. The validity of the PCR design and experiment was verified using capillary electrophoresis sequencing (MacroGen) and gel electrophoresis of each PCR product.

### 3 | RESULTS

#### 3.1 | Bioinformatics evaluation

The whole-genome sequence alignment of different human CoVs showed that RNA replicase gene (encoded by ORF1ab) has conserved sequences as compared to other regions, particularly the part of ORF1b, which indicates the lower mutation rate of this region during the evolution of different CoVs. However, inside the ORF1b at the NSP12 coding sequence, there are species-specific sequences between short lengths of somewhat homologous sequences which are selected for locating primers to be universal for targeting all types of alpha and beta CoVs. Therefore, the selected primers have some ambiguous nucleotide bases (Figure 1).

In silico PCR amplification of the selected primers produced a single PCR product of 332 bp for hCoVs genomes, without any unintended matching with human genomes and other human respiratory viruses (Table 1). The matching score of each amplicon sequence against other related CoVs genomes was estimated to evaluate the suitability of the selected amplicon to be a taxonomic sequence distinguishable for each type of targeted CoVs. The resolution of discrimination between different species of CoVs was estimated according to the difference between the max score of the complete aligned sequence and the bit score of the related species. The results indicate that each amplicon sequence is distinguishable for the species, with sub-species resolution in some lineages (Table 2). The entropy plot of this taxonomic sequence reflected high variability among 96 analyzed sequences of different CoVs species (Figure 2). To illustrate, the number of the analyzed nucleotide after subtracting primers bases equals 293 and the count of variable nucleotides equals 180 (61.4%), which have average entropy value equals 0.723979278.

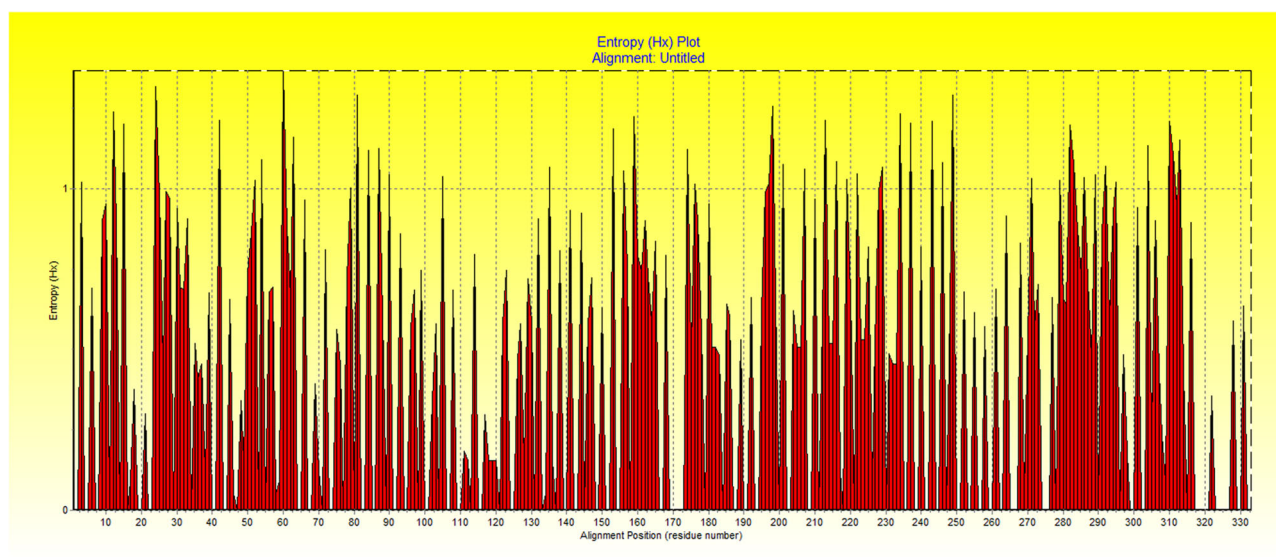
Phylogenetic analysis of alpha and beta CoVs amplicons showed excellent topology for the classification of each subgenus individuals into closely related clades. The taxa involved 96 nucleotide sequences of different CoVs strains, including 18 species of alpha CoVs and 78 several species of all beta CoVs subgenera; *Embecovirus*, *Sarbecovirus*, *Merbecovirus*, *Nobecovirus* and *Hibecovirus*. The resulting clades of the UPGMA tree put the closely related species together with homogenous distances according to their classification (Figure 3).

#### 3.2 | Molecular identification of CoVs by PCR product sequencing

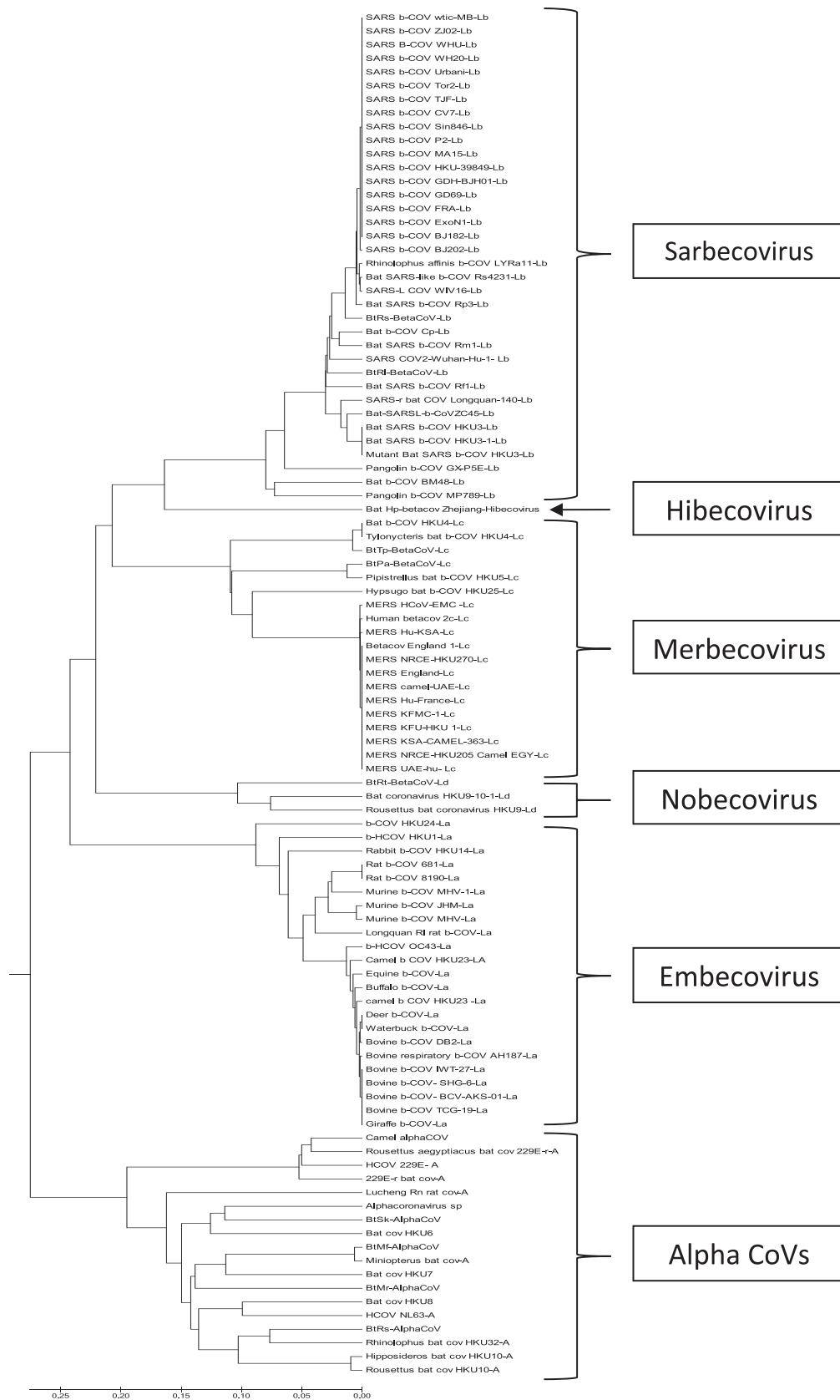
The validity of the given PCR technique for identifying CoVs genomes was practically confirmed by direct sequencing of the obtained PCR products with the same primers in both directions, followed by BLAST analysis of the obtained sequences. The results revealed that the product of this PCR contains a species-specific sequence with a resolution of more than 250 nucleotide bases of recognizable bar code sequence for each CoV species. Moreover, the validity of the PCR design for producing a single PCR product, for each examined CoV genome, was confirmed by gel electrophoresis and HRM analysis. In particular, each PCR product exhibits a single band at molecular weight equaling 332 bp (Figure 4), and a distinctive melting profile (Figure 5).

### 4 | DISCUSSION

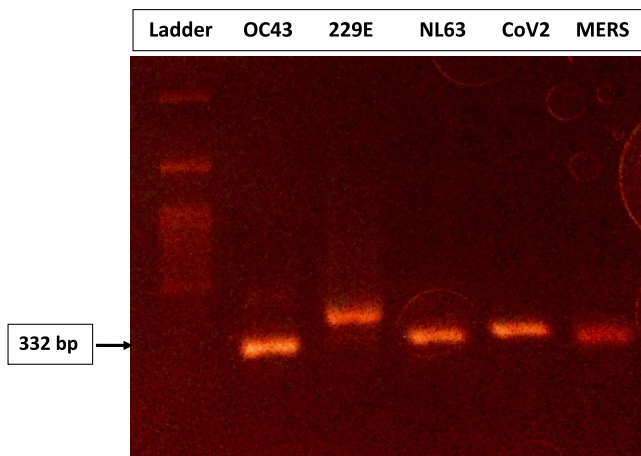
According to the literature, evolution reservoirs of SARS-CoV-2 belong to bat-related viruses. Researchers need to survey a wide range of variants of each geographic isolates of these viruses around the



**FIGURE 2** The entropy plot of 332 bp nucleotide sequence resulted by multiple sequence alignment of 96 different alpha and beta CoVs genomes within the amplicon range, using BioEdit software



**FIGURE 3** Phylogenetic UPGMA tree of 96 alpha and beta CoVs amplicon sequences. The length of branches was drawn upon the scale and was estimated according to the evolutionary distance, using the Maximum Composite Likelihood method, which represents the number of base substitutions per site (total = 332 bp)

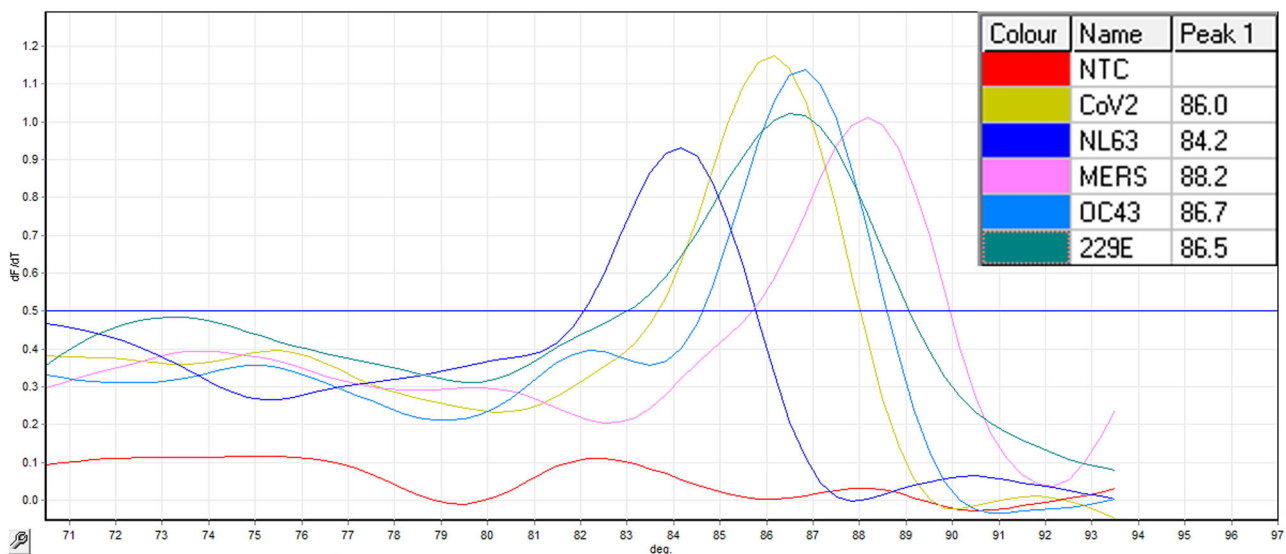


**FIGURE 4** Molecular screening agarose gel electrophoresis loaded by PCR products of five different hCoV species; human coronavirus OC43, human coronavirus 229E, human coronavirus NL63, SARS-CoV-2, and MERS. MERS, Middle East respiratory syndrome coronavirus; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2

bats' CoVs, which are represented by 96 genomic sequences of different types of alpha and beta CoVs. To identify species DNA bar code, amplicon sequence was analyzed using BLAST.

Interestingly, by exploring the results of the present study, it is well expected that the provided protocol is suitable to diagnose the upcoming and the unknown variants of mammals' CoVs. This is due to the design of PCR primers targeting the conserved regions inside the NSP12 coding sequence. As NSP12 is encoding for RdRp which is important for virus replication into the host cells, its nucleotide sequence contains clusters of short conserved sequences among different lineages of CoVs. Consequently, the designed primers are suitable to be universal for all mammals CoVs, whether ancestors or variants.

However, the designed primers contained ambiguous bases to avoid the mismatching of the variable bases of diverged sequences. It is easy to replace these ambiguous bases with specific bases when targeting a distinct lineage of CoVs. For example, the derived specific primers targeting SARS and SARS-CoV-2 have no ambiguous bases and match completely with all sub-species variants.



**FIGURE 5** Screenshot of the HRM analysis output of hCoVs PCR products using Rotor-Gene Q instrument

globe and follow their evolutions in relation to the capability to infect human cells. The danger of the upcoming unknown variants of such viruses will be extreme if the outbreak occurred globally before the readiness to overcome the disease.

The best strategy to overcome such a pandemic is the early detection of the pathogen and the tendency to prepare billions of doses of suitable vaccines. Thus, the premeditation to have a step before pathogen outbreak needs to find easy and fast diagnosis tools. The coupling of both diagnosis and identification processes in a tool is the best for large survey studies.

The given protocol in the present study focused on diagnosing all types of mammals' CoVs with a particular emphasis on humans and

Moreover, the genomes of all SARS-CoV-2 variants do not exhibit frequent nucleotide polymorphisms in the amplicon range (14968:15299), till the date. This reflects that the present and the upcoming variants will suit this diagnostic tool. Furthermore, it is well expected that the evolution of a new CoV strain diverged from another mammalian-related CoV to humans will also be detectable by this protocol.

Generally, the fingerprinting assessment, for typing organisms by DNA barcoding, is based on selecting distinguishable amplicon sequences inside a conserved gene of an organism that contains enough variability as compared to other related species, such as 16S ribosomal DNA gene in bacterial species. In this regard, the design of

CoVs-typing protocol, in the present study, achieved the goal of discriminating alpha and beta CoVs in a species or subspecies-based resolution, by DNA barcoding of a conserved gene (NSP12) using PCR universal primers. This will facilitate the rapid diagnosis and identification of these CoVs as a preliminary step before full genome sequencing for the discovered isolates, to reduce the cost of survey studies.

The validity of the provided DNA barcoding protocol was verified by estimating amplicon sequence variability among related species, and by applying amplicon sequences of taxonomically known CoV species for phylogenetic analysis, the obtained tree confirmed the classification of each beta CoVs lineages and separate them from alpha CoVs species.

Information and the outcome data concluded from this study are logically in harmony with the previously announced essential research theories. Firstly, the RdRp region contains taxonomic sequences used for detection, identification, and classification of CoVs,<sup>10</sup> and reveals the highest percentage of conserved sequences among SARS-CoV-2 variants<sup>11</sup> and other beta CoVs species.<sup>12</sup> A similar previous study published by Wacharapluesadee et al.<sup>13</sup> investigated the diversity of bats' CoVs using the RdRp gene for classifying them. However, their PCR design is not matched well with some CoVs, especially hCoVs such as NL63, MERS, and OC43. Therefore, the present study tends to modify this protocol to match all alpha and beta CoVs.

It is inferred from experience that the use of a specific diagnosis tool restricted for targeting a species, such as using species-specific PCR primers, may lead to losing the ability to detect its diverged variants before their subsequent outbreak. This probability is referred to the chance of mutations that may result in nucleotide substitutions inside the PCR priming sites. Thus, by the approached tool in the current research, it is easy to target a specific species of CoVs, by designing a specific probe inside the variable region of the amplicon sequence, to be included in the PCR. This will facilitate the ability to identify other CoVs genomes (if present) in probe-unmatched samples which showed positive PCR amplification.

## 5 | CONCLUSION

The present research approached a simple protocol for molecular typing of mammals' CoVs (including alpha and beta CoVs), using DNA barcoding of a taxonomic sequence inside a conserved gene (NSP12). By this protocol, it is easy to diagnose (positive PCR indicates infected sample) and identifies a wide range of these viruses by short-length sequencing of the obtained amplicon (332 bp). This will be useful for surveying different geographic isolates of the targeted CoVs, and suitable to detect diverged variants.

## ACKNOWLEDGEMENT

Great of thanks to Professor Ehab Serag, Microbiology Reference Lab of the National Blood Transfusion Service (NBTS), Ministry of Health,

Egypt, for providing the authors with purified RNA samples of identified hCoVs.

## AUTHOR CONTRIBUTIONS

Waleed A. Nemr contributes by the practical part of PCR, and writing the paper. Nashwa K. Radwan contributes by the interpretation of results and helps in the in silico design of PCR.

## CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

Waleed A. Nemr  <http://orcid.org/0000-0001-8465-0497>

## REFERENCES

- Hulswit RJ, de Haan CA, Bosch BJ. Coronavirus spike protein and tropism changes. In: Ziebuhr J, (ed), *Advances in virus research*. Academic Press; 2016;1(96):29-57. doi:10.1016/bs.aivir.2016.08.004
- Llanes A, Restrepo CM, Caballero Z, Rajeev S, Kennedy MA, Lleonart R. Betacoronavirus genomes: how genomic information has been used to deal with past outbreaks and the COVID-19 pandemic. *Int J Mol Sci*. 2020;21(12):4546. doi:10.3390/ijms21124546
- Lau SK, Li KS, Tsang AK, et al. Genetic characterization of Betacoronavirus lineage C viruses in bats reveals marked sequence divergence in the spike protein of pipistrellus bat coronavirus HKU5 in Japanese pipistrelle: implications for the origin of the novel Middle East respiratory syndrome coronavirus. *J Virol*. 2013;87(15):8638-8650. doi:10.1128/JVI.01055-13
- Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020; 579(7798):270-273. doi:10.1038/s41586-020-2012-7
- Stoeckle M. Taxonomy, DNA, and the bar code of life. *BioScience*. 2003; 53(9):796-797. doi:10.1641/0006-3568(2003)053%5B0796:TDATBC%5D2.0.CO;2
- Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser*. 1999;41:95-98.
- Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;13:134. doi:10.1186/1471-2105-13-134
- Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences (USA)*. 2004;101:11030-11035. doi:10.1073/pnas.0404206101
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 2013;30:2725-2729. doi:10.1093/molbev/mst197
- Wilkinson DA, Joffrin L, Lebarbenchon C, Mavingui P. Analysis of partial sequences of the RNA-dependent RNA polymerase gene as a tool for genus and subgenus classification of coronaviruses. *J Gen Virol*. 2020;101(12):1261-1269. doi:10.1099/jgv.0.001494
- Miao M, Clercq ED, Li G. Genetic diversity of SARS-CoV-2 over a one-year period of the COVID-19 pandemic: a global perspective. *Biomedicine*. 2021;9(4):412. doi:10.3390/biomedicine9040412



12. Wang R, Chen J, Gao K, Hozumi Y, Yin C, Wei GW. Analysis of SARS-CoV-2 mutations in the United States suggests presence of four substrains and novel variants. *Communications biology* 15. 2021; 4(1):1-4. doi:10.1038/s42003-021-01754-6
13. Wacharapluesadee S, Duengkae P, Rodpan A, et al. Diversity of coronavirus in bats from Eastern Thailand. *Virology* 2015;12(1):1-7. doi:10.1186/s12985-015-0289-1

**How to cite this article:** Nemr WA, Radwan NK. Typing of alpha and beta coronaviruses by DNA barcoding of NSP12 gene. *J Med Virol*. 2022;94:1926-1934.  
doi:10.1002/jmv.27550