# Wuhan-coronavirus homologue map

### Kou Amano

## Background and Objective

Background: At the beginning of 2020, global risk of infection of coronavirus is spreded. The pandemic started in 2019 and governments announced a state of emergency. In Japan, the government adopted PCR as a diagnosis method of the infection. But the selection of prmers infruence the accuracy greatly. Objective: Therefor, I provide a "map" of homological region of coronavirus to other viruses and annimals to help the primer design.

## Data

#### Vertebrates:

- Bat (NW\_017738920.1 .. NW\_017802358.1; 63439frgs)
   Dog (NC\_006583.3 .. NC\_006621.3; 39chrs)
- Beluga (NW\_022097992.1 .. NW\_022103895.1; 5904frgs)
- Camel (NC\_044511.1 .. NC\_044547.1; 37chrs)
- Cat (NC\_018723.3 .. NC\_018741.3; 19chrs) Viruses:
- 7554 genomes of viruses or phages
- Ferret (NW\_004569142.1 .. NW\_004576923.1; 7782frgs)
- Human (NC\_000001.4 .. NC\_000024.3; 24chrs)
- Mouse (NC\_000067.6 .. NC\_000087.7; 21chrs)
- Wuhan-corona virus genome (MN908947.3)
- Rabbit (NC\_013669.1 .. NC\_013690.1; 22chrs)
- Pig (NC\_010443.5 .. NC\_010462.3; 20chrs)
- Rock Pigeon (NW\_004973171.1 .. NW\_004988092.1; 14922frgs)
- Turkey (NC\_015011.2 .. NC\_015042.2; 32chrs)

#### Method

#### BLAST

DB: makeblastdb -in <<input file>> -out <<DB name>> -dbtype nucl -parse\_seqids Query: megablast -d <<DB name>> -i <<query sequence>> -W 10 > <<output file>>

#### Self-BLAST

Fragmentation: fragment bf=<<input file>> S=25 G=25 cs=1 > <<output file>> DB: same as above.

Query: same as above.

#### Window-fourrier

Fragmentation: 30 fragments – 1000 bases Conversion: "A" -> 1, "T" -> -1, "G" -> I, "C" -> -I Fourier transform: Ft (<< each fragment>>)

## Selection of frequent homologues Vertebrate: base-hit count >= 5; seq-lenght >= 5

Viruses: base-hit count >= 18; seq-lenght >= 5

							The homologues
No.	homologue	start	end	count	in CDS		
1	AATTTTA	1787	1793	6	Yes		
2	GATGAGGATGAAGAAGAAGGT	3047	3067	6	Yes		
3	CTACAAAGAAACAGTTAC	5866	5884	5	Yes		
4	ATAAATATTATAATTTG	6944	6960	7	Yes		
5	TTGCAT	7425	7430	5	Yes		
6	CCATCCATCTTTACTTTGATAAA	7764	7786	7	Yes	No.	homologue start end count in CDS
7	TTTTTGTTGCTGCTATTTTCTATTTA	8607	8632	9	Yes	15	CAAGATCTCAATGGTAACTGGTATGATTTCGGTGATT 14 068 14 104 21 Yes
8	ATTTCTCTGTTTGTTTTTGT	11 167	11 186	7	Yes	16	TGGTAATGCTGC 14775 14786 18 Yes
9	AAAAGT	12 198	12 203	5	Yes	17	ACAAAACGTAATGTCATCCCTACTATAACTCAAATGAATCTTAAGTATGCCATTAGTGCAAAGAATAGAGCTCGCACCGTAGCTGGTGTCTCTAT 15 031 15 125 32 Yes
10	ATAAAATAGAAGAA	19 121	19 134	6	Yes	18	CTTATGGGTTGGGATTATCCTAATACCTTTTTTAACAT
11	TGTTTGTTTTTCTTGTTTTATTGC	21564	21 587	14	Yes	19 20	GATGCCACAACTGCTTATGCTAATAGTGTTTTTAACAT  CAAAACAATGTTTTTATGTCTGAAGCAAAATGTTGGACTGAGACTGACCTTACTAAAGGACCTCATGAATTTTGCTCTCAACATACA  15 490 15 527 19 Yes  CAAAACAATGTTTTTATGTCTGAAGCAAAATGTTGGACTGAGACTGACCTTACTAAAGGACCTCATGAATTTTGCTCTCAACATACA 15 805 15 891 26 Yes
12	TATTAAAATATAATGAA	22 389	22 405	5	Yes	21	GGTTGTGATGGTGGCAGTTTGTATGTAAATAAACATGCATTCCACACACCC 19 276 19 325 27 Yes
13	TTTTCT	26 495	26 500	5	No	22	TATGAGAGTTATACATTTTGGTGCTGGTTCTGATAAAGGAGTTGCACCAGGTAC 20 850 20 903 21 Yes
14	CCTAA	29 375	29 379	5	Yes	23	CCGAGGCCACGCGGAGTACGATCGAGTGTACAG 29 732 29 764 32 No
Но	omologue count in 12 vertebrates (Max: 14)	1	2 		4 5 m.n.kh.\l\lh.k.	7	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$
	Turkey  Rock Pigeon ■ ■						
	Pig	•			_	•   _ •	
	Mouse					-   -	
apped region	n to 12 vertebrates Human  Ferret			•	• • •	<u>-</u>	
	Dog	-			•		
	Cat Camel ■		•			_	
	Beluga	•	•	•	•	•	
han-corona	genome (PCR primers: black dot)	-			•		
	Coding region (star end: )			5000			10000 20000 25000
	Mapped region to 7556 viruses (total)						
	Mapped region (length <= 35) to viruses =	-					
	Homologue count in viruses (Max: 37)						<sup>┸</sup> ╅┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸╫╫┎┧╟╏┧╴┉╸╴┎╟┪╟╶╏┸╶┅╌┧┅╅╻╟╏╍╏┉┟┸┲┰┉┉┰┼┪╻┉┰┰┪┎┸┪╟┸┩ <sup>┸</sup> ┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸┸
							16
							15
							25
		I	and the second of the second o	and the second s	1 1 1	1. I	
	المحرابة المارية المار	المراجع		المنافل لينهين والمنطيان		المعتقد والطالس الد	
	Window Fourier (size: 1000 base)						
	Window Fourier (size: 1000 base)					complete district	
	Window Fourier (size: 1000 base)				والوراد المحمد الأوراد الاحتجاز والمحمد المحمد المح	(Իրոս <sup>գ</sup> ակակցի) (հուսաբոսի	
	Window Fourier (size: 1000 base)	gagliaer, ffi. galid filte filter, etia etia galida filte	ار المراجعة المحاولة	ec <sub>e</sub> mmi midd fall ei endag et egel e <sub>e</sub> gewydd fal fall fall fal fall fall fal fal fal	والوراد المحمد الأوراد الاحتجاز والمحمد المحمد المح	(Իրոս <sup>գ</sup> ակակցի) (հուսաբոսի	
	Window Fourier (size: 1000 base)	gagliaer, ffi. galid filte filter, etia etia galida filte		ec <sub>e</sub> mmi Pridigia (ecentra), esperante de la compositiva della com	والوراد المحمد الأوراد الاحتجاز والمحمد المحمد المح	(հուսագրական) (դուսանացի	