

# Wuhan-coronavirus homologue map

Kou Amano

## Background and Objective

**Background:** At the beginning of 2020, global risk of infection of a new coronavirus is spreaded. 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 primers influences the accuracy greatly.

**Objective:** Therefor, I provide a "map" of homological regions of coronavirus genome to other virus and animal genomes to help the primer design.

## Data

### Vertebrates

- Bat (NW\_017738920.1 .. NW\_017802358.1; 63439frgs)
  - Beluga (NW\_022097992.1 .. NW\_022103895.1; 5904frgs)
  - Camel (NC\_044511.1 .. NC\_044547.1; 37chrs)
  - Cat (NC\_018723.3 .. NC\_018741.3; 19chrs)
- Dog (NC\_006583.3 .. NC\_006621.3; 39chrs)
  - Ferret (NW\_004569142.1 .. NW\_004576923.1; 7782frgs)
  - Human (NC\_000001.4 .. NC\_000024.3; 24chrs)
  - Mouse (NC\_000067.6 .. NC\_000087.7; 21chrs)
- 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)

### Viruses

- 7554 genomes of viruses or phages (FTP bulk download)
  - Wuhan coronavirus genome (MN908947.3)
- PCR primer (<https://www.niid.go.jp/niid/images/lab-manual/2019-nCoV20200217.pdf>)

## Method

### BLAST

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

### Self-BLAST

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

### Window-fourier

Fragmentation: 30 fragments; 1000 bases / fragment  
Conversion: "A" -> 1, "T" -> -1, "G" -> I, "C" -> -I  
Fourier transform:  $Abs(Ft(\text{<<each fragment>>}))$

### Selection of frequent homologues

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

## Result

