Identifying Mutations in ACE2 That Influence Susceptibility to SARS-CoV-2

(L) Zak White

Introduction / SARS-Cov-2

COVID-19 is caused by the virus **SARS-CoV-2**

- Likely of zoonotic origin
- Some non-human animals are known to be susceptible; some are not

Introduction / Motivation

Motivation:

- Why are some animals susceptible to SARS-CoV-2 and some are not?

Big-eared horseshoe bat (L) is susceptible; the greater horseshoe bat (R) is not susceptible. Both belong to genus Rhinolophus

Introduction / ACE2

The SARS-CoV-2 spike protein's target in potential hosts is the ACE2 protein (angiotensin-converting enzyme 2)

- Differences in ACE2 sequences can result in different interaction with the spike protein, or possibly inhibit interaction altogether

Introduction / Objective

Objective:

- Identify variations in ACE2 sequences (ie, mutations) that influence a host's susceptibility to SARS-CoV-2

Introduction / Related Work

Li et al. (2005)

- Researched SARS-CoV spike protein interaction with ACE2 chimera sequences
- Identified mutations that restricted interaction: D31, A41, 82-84NFS, A353, H353, and A357

Liu et al. (2021)

- Structural analysis of human ACE2 to identify critical residues
- Concluded acids at positions 31, 35, 38, 82, and 353 were crucial for supporting interaction

Methods / Data

25 ACE2 sequences were collected from NCBI and UniProt:

- 12 susceptible mammal hosts (incl. human)
- 6 insusceptible mammal hosts
- 7 insusceptible partial sequences

Each was aligned with the human sequence to standardize indexing

- Complete: Needleman-Wunsch
- Partial: Smith-Waterman

Methods / Algorithm

Iterative sequence comparison algorithm

- Compared acids at each positions
- Assigned weights to mutations based on number of mutations in the sequence
- Summed weights to identify most influential mutations

```
for each negative sequence S⁻ do
    for each i < length of S⁻ do
        isInfluentialIndex ← True;
        for each positive sequence S⁺ do
            if i^th acid in S⁻ == i^th acid in S⁺ then
                isInfluentialIndex ← False;
            end
        end
        if isInfluentialIndex then
            // The mutation may be influential
            Record i, i^th acid in S⁻
        end
    end
end
```

Results

The analysis identified 8 influential mutations

- 5 of the results overlapped with the findings of other studies (D31, A41, F83, H353, S426)

Pos	Mutation		
31	Lysine (K)	→	Aspartate (D)
41	Tyrosine (Y)	→	Alanine (A)
66	Glycine (G) Arginine (R)	⇒	Alanine (A)
83	Tyrosine (Y)	→	Phenylalanine (F)
113	Serine (S) Arginine (R)	⇒	Asparagine (N)
353	Lysine (K)	→	Histidine (H)
426	Proline (P)	→	Serine (S)
679	Isoleucine (I)	→	Valine (V)

Discussion

Each mutation appeared in at least two insusceptible sequences and no susceptible sequences

- Each insusceptible sequence contained at least one of these mutations

The presence of these mutations may be indicators of a host's insusceptibility to the virus

Three mutations identified have not been mentioned in past studies (A66, N113, V679)

Pos	Mutation		
31	Lysine (K)	→	Aspartate (D)
41	Tyrosine (Y)	→	Alanine (A)
66	Glycine (G) Arginine (R)	⇒	Alanine (A)
83	Tyrosine (Y)	→	Phenylalanine (F)
113	Serine (S) Arginine (R)	⇒	Asparagine (N)
353	Lysine (K)	→	Histidine (H)
426	Proline (P)	→	Serine (S)
679	Isoleucine (I)	→	Valine (V)

Discussion / Considerations

Limited amount of data was a main restriction

- Not enough research into which animals are or are not susceptible
- More sequences could produce more accurate results

Effect of mutations in consecutive positions

- Some studies (incl. Li et al.) show that mutation in positions 82-84 together can influence susceptibility
- The algorithm could be expanded to account for this

Conclusion

This analysis identified eight ACE2 mutations that may influence susceptibility to SARS-CoV-2

- The presence of any of these mutations can be an indicator for a host's immunity

Pos	Mutation		
31	Lysine (K)	→	Aspartate (D)
41	Tyrosine (Y)	→	Alanine (A)
66	Glycine (G) Arginine (R)	⇒	Alanine (A)
83	Tyrosine (Y)	→	Phenylalanine (F)
113	Serine (S) Arginine (R)	⇒	Asparagine (N)
353	Lysine (K)	→	Histidine (H)
426	Proline (P)	→	Serine (S)
679	Isoleucine (I)	→	Valine (V)

Questions

Thanks for listening!