

Thesis progression

Thesis title: Bat Virome Analysis and Development of a Platform for Assessing the Spillover Potential of Coronaviruses Identified from Bats in Thailand

Thesis progression title: The assessment of AXL-binding potential coronavirus

Student: Narathit Chanraeng **Student ID:** 667070027-8

Advisor: Assoc. Prof. Dr. Supranee Phanthanawiboon

Co-advisor: Prof. Dr. Chamsai Pientong

Date: January 7th, 2026

1. Introduction

Emerging infectious diseases (EIDs) can be described as infections that are recently recognized or increased in population in a new area. EIDs are a serious threat to global health and an economic burden (Taylor et al., 2001). 75% of EIDs are zoonoses or zoonotic diseases, which originate from pathogens in animals and transmit to humans (Gebreyes et al., 2014). The majority of zoonotic diseases are caused by the zoonotic spillover viruses to humans (Kreuder Johnson et al., 2015). Zoonotic spillover can be defined as the cross-species transmission of pathogens from wild animals to humans. The potential of zoonotic spillover can be determined by interactions among several factors including disease dynamics in the reservoir host, pathogen exposure, and the within-human factors that affect susceptibility to infections. Moreover, key factors determining spillover potential include cellular and tissue tropism, virulence and characteristics of the pathogen, and the ability of the pathogen to adapt and evolve within a novel host environment (Escudero-Pérez et al., 2023).

Bats are important reservoir hosts of many zoonotic viruses that can infect humans and other domestic or wild mammals. Bats belong to the order Chiroptera. Bats are around 20% of all living mammals, with more than 1,300 species, and are found worldwide except in extreme polar regions (Wilson & Reeder, 2005). Bats are the only mammals that can truly fly because of their adaptations in anatomy and physiology including their elongated fingers with stretched skin, strong flight muscles, flexible wings, lightweight skeleton, and a high metabolic rate (Teeling, 2009). Bats are natural reservoirs host for several spillover viruses to humans including the Nipah virus, Hendra virus, Ebola virus, Pandemic avian influenzas, West Nile virus, and some coronavirus such as severe acute respiratory syndrome coronavirus (SARS-CoV), middle east respiratory syndrome

coronavirus (MERS-CoV), and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Han et al., 2015). A broad-scale conclusion from metagenomic studies showed that bats may be particularly prone to carry viral families commonly associated with zoonotic disease. Of the more than 16,600 bat-associated viral sequences on NCBI/GenBank, 30% are the families Coronaviridae (Van Brussel & Holmes, 2022).

In the past two decades, coronaviruses have been associated with human emergence. The Coronaviruses are subdivided into four genera. The genus Alphacoronavirus (α) contains the human coronavirus 229E (HCoV-229E), one other human coronavirus NL63 (HCoV-NL63), and many animal viruses. The genus Betacoronavirus (β) includes the prototype mouse hepatitis virus (MHV), the three human coronaviruses including human coronavirus OC43 (HCoV-OC43), SARS-CoV, and human coronavirus HKU1 (HCoV-HKU1), and MERS-CoV, together with several animal coronaviruses. The genus Gammacoronavirus (γ) contains viruses of cetaceans (whales) and birds, and the genus Deltacoronavirus (δ) contains viruses isolated from pigs and birds (Burrell et al., 2016). Three major outbreaks of coronaviruses: SARS-CoV and MERS-CoV caused significant human morbidity and mortality in 2002 and 2012 respectively and most recently, the SARS-CoV-2 caused coronavirus disease 2019 (COVID-19) pandemic in 2019 that has caused millions of cases and deaths (Pustake et al., 2022). Moreover, most of the virus spillover risk ranking of the top 50 wildlife viruses, including viruses known to be zoonotic and those with unknown zoonotic potential are coronaviruses (Grange et al., 2021). The associated costs of these preventive efforts would be substantially less than the economic and mortality costs of responding to these pathogens once they have emerged. Thus, efforts to increase preparedness and improve surveillance for emerging coronaviruses represent a priority for global health programs (Bernstein et al., 2022). Common receptor usage and the ability of viruses to enter and replicate in human cells are major factors linked to spillover potential. To assess the spillover potential of coronaviruses, human proteins or receptors that have the potential to support zoonotic spillover (restriction factors, receptors, other cellular proteins) are necessary to identify and determine whether those are few or many (Escudero-Pérez et al., 2023).

This study aims to develop a platform for assessing the spillover potential of coronaviruses by investigating common receptor usage of coronaviruses among animal hosts and humans. And analyze the virome profile in bats in Thailand and assess the spillover potential of identified

coronaviruses in bats. This is important for pandemic preparedness and improve surveillance to prevent future coronavirus spillover.

2. Hypothesis

1. Bats in Thailand are reservoirs of many zoonotic viruses including coronavirus
2. Coronaviruses identified from bats in Thailand are capable of spillover to humans

3. Objectives

1. To analyze virome profile in bats in Thailand
2. To develop a platform for assessing the spillover potential of coronaviruses
 - 2.1. To investigate potential receptors of coronaviruses that are shared among animal hosts and human
 - 2.2. To construct potential receptor-expressing cell lines and pseudovirus
3. To investigate the spillover potential of identified coronaviruses from bats in Thailand using the developed platform

4. Scope and limitation

According to the plan and budget, the bat samples in the study are collected in three different shedding routes from two provinces, including Chanthaburi and Chiang-rai. The receptors for assessing spillover potential are selected and constructed based on the available database.

5. Anticipated outcome

1. Knowledge of common receptor usage of coronaviruses shared among animal hosts and human
2. List of viruses in bats in Thailand
3. Diversity of the identified coronaviruses in bats in Thailand
4. List of potential spillover coronaviruses from bat to human

6. Conceptual framework

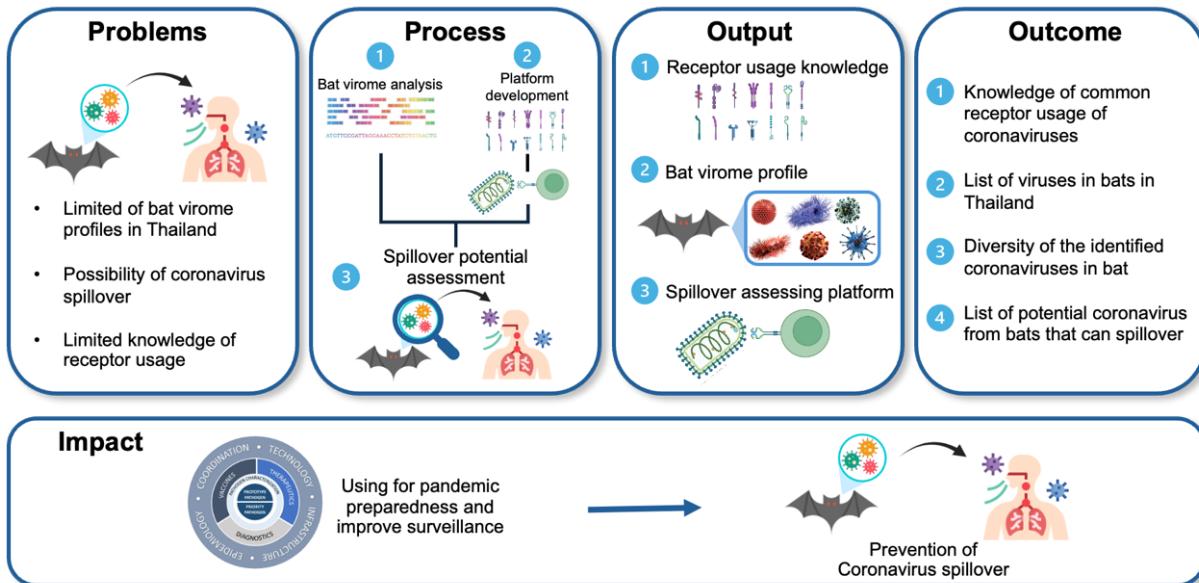


Figure 1. Conceptual framework

Table 1. Detail of conceptual framework

Step	Detail
Problems	The problems of this study are the bats virome profiles in Thailand remain limited and there are possibility of coronavirus to transmit to humans. Moreover, the knowledge of receptor usage of coronavirus which is the key factor of virus spillover remains limited.
Process	To fix the problems metagenomic analysis will be used to analyze the virome profiles of bats in Thailand. Then, the coronavirus receptor usage will be retrieved and investigate potential receptor and developed platform that will be used to analyze the potential coronavirus spillover.
Output	<ol style="list-style-type: none"> 1. Receptor usage knowledge of coronavirus 2. The virome profiles of bats in Thailand 3. The platform for assesment of coronavirus spillover potential
Outcome	<ol style="list-style-type: none"> 1. Knowledge of common receptor usage of coronaviruses 2. List of viruses in bats in Thailand

3. Diversity of the identified coronaviruses in bat

4. List of potential coronaviruses spillover

Impact

Virome profile in bat and list of coronavirus that have potential spillover will be used for pandemic preparedness and for improvement of the surveillance leading to the prevention of coronavirus spillover in the future.

7. Experimental design

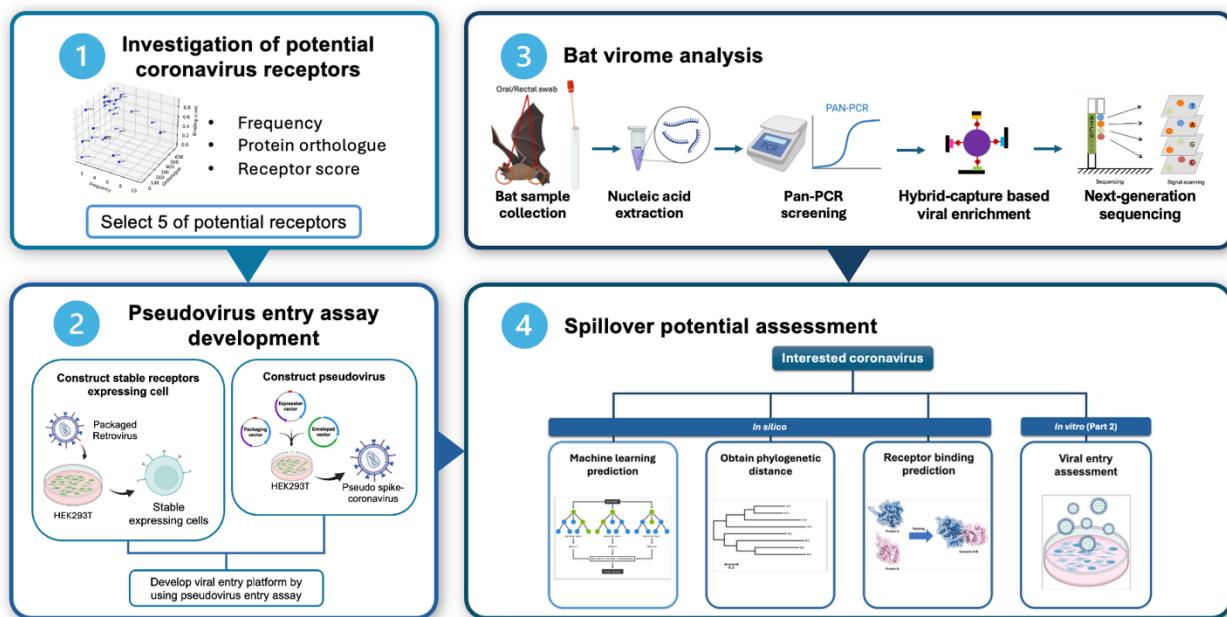


Figure 2. Experimental design

As shown in the flowchart, the experimental design was subdivided into four parts, including Part 1 Investigation of potential coronavirus receptors, Part 2 Pseudovirus entry assay development, Part 3 Bat virome analysis by metagenomic analysis, and Part 4 Spillover Potential Assessment.

8. Previous progression

The Random Forest Classifier can effectively predict the human infection potential of coronaviruses based on spike protein sequences. Through extensive cross-validation and external validation, the Random Forest Classifier consistently achieved excellent predictive performance across all key metrics, confirming its generalizability and reliability. SHAP-based feature importance analysis further identified specific k-mer trimers, including LEP and KIQ, as strongly associated with human infection and mapped to the N-terminal domain (NTD) and HR1 regions of the S protein. Based on these findings, we focused on the potential role of the NTD in mediating human infection. Notably, the NTD has been reported to interact with AXL, an alternative receptor that enables ACE2-independent SARS-CoV-2 entry, suggesting that the NTD may contribute to coronavirus spillover through AXL-mediated pathways.

9. Materials and Methods

To assess the AXL-binding potential of coronaviruses, key residues involved in the AXL–SARS-CoV-2 spike interaction were identified through a literature review (Lei et al., 2023; Wang et al., 2021). Total 5391 complete spike protein sequences of coronaviruses, excluding SARS-CoV-2, were retrieved from the NCBI Virus database. Each sequence was then pairwise aligned to the SARS-CoV-2 spike protein using PairwiseAligner, and the corresponding key AXL-interaction residues were extracted to identify coronaviruses with potential AXL-binding capability. Selected coronaviruses were subsequently used to construct a phylogenetic tree based on whole-genome nucleotide sequences using MAFFT for multiple sequence alignment and IQ-TREE for phylogenetic construction with 1000 bootstrap. In addition, cophylogenetic analyses were performed using the RdRp and spike protein sequences to compare their evolutionary histories. Similarity plot analysis was conducted to examine sequence similarity patterns across the spike gene using SimPlot. Finally, spike protein structures of the selected coronaviruses were modeled using SWISS-MODEL, and protein–protein docking against the human AXL receptor was performed using the HADDOCK server.

10. Results

1. Key residue of AXL-SARS-CoV-2 interaction

Table2. Key residue of AXL-SARS-CoV-2 interaction

AXL	SARS-CoV-2 NTD	Interaction types
E70	K147	
E70	K150	
I68	K150	
H61	S247	Hydrogen bond
E59	S247	
E59	R246	
E85	S256	
P57	W152	
P57	P251	
P58	W152	
P58	P251	Hydrophobic
I68	W152	
I68	P251	
F113	W152	
F113	P251	

2. Key residue identity relative to SARS-CoV-2 of coronaviruses

From a total of 5,391 complete spike protein sequences, the key residue identities relative to SARS-CoV-2 for the top 50 coronaviruses with AXL-binding potential are shown in Figure 3, together with the percentage of conserved key residues and overall sequence identity. The result shown eleven coronaviruses exhibited complete conservation of key residues across all analyzed positions. However, only nine of this also showed high overall sequence identity. In contrast, two coronaviruses displayed low overall sequence identity while retaining identical key residues. Therefore, these eleven coronaviruses were selected for further analyses.

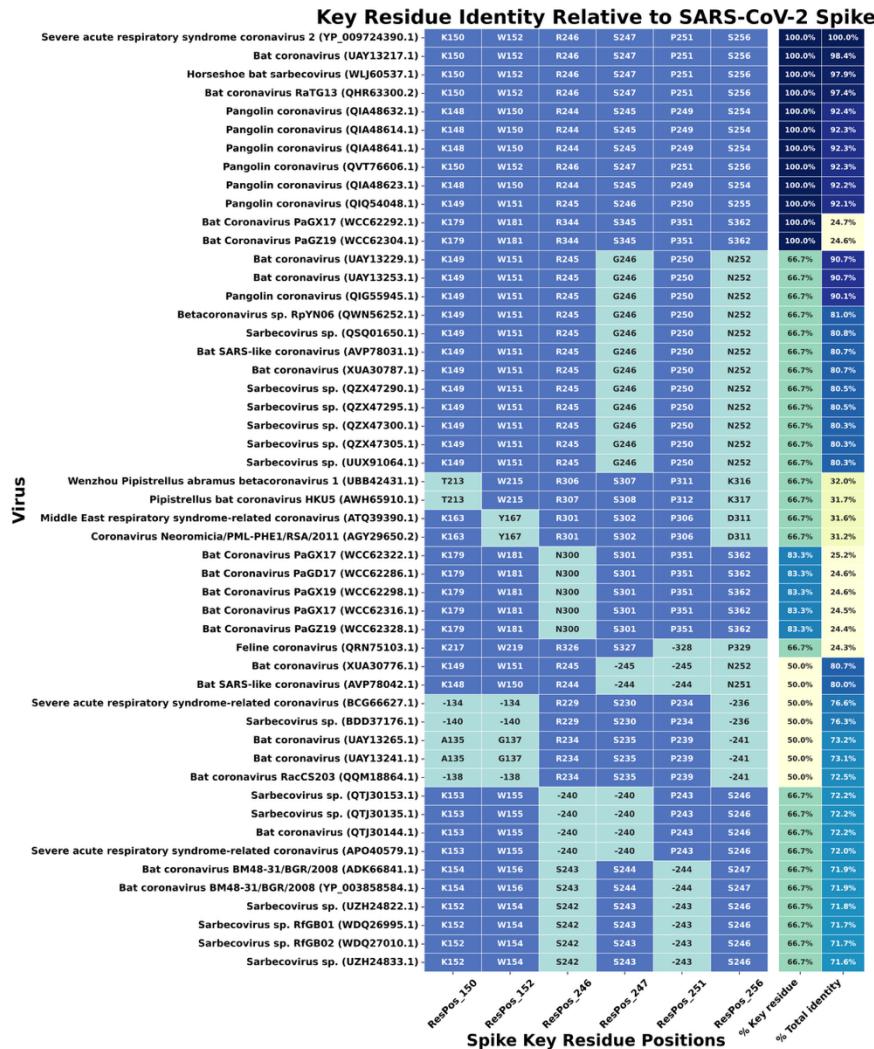


Figure 3. Key residue identity relative to SARS-CoV-2 of coronaviruses alongside with the percentage of conserved key residues and overall sequence identity.

3. Phylogenetic tree of selected coronavirus based on whole genome

To examine the evolutionary relationships of coronaviruses with AXL-binding potential, a phylogenetic tree was constructed based on whole-genome nucleotide sequences. The analysis shows that the coronaviruses are clearly separated into the four established genera including Alphacoronavirus, Betacoronavirus, Gammacoronavirus, and Deltacoronavirus, indicating that the tree topology is consistent with current coronavirus taxonomy. The nine coronaviruses with high overall spike sequence identity cluster within the Betacoronavirus genus, including pangolin coronaviruses, bat coronaviruses, horseshoe bat sarbecovirus, and bat coronavirus RaTG13, group

closely with SARS-CoV-2, forming a well-supported clade. In contrast, the two coronaviruses with low overall sequence identity, bat coronavirus PaGX17 and PaGZ19, cluster within the Alphacoronavirus clade (Figure 4).

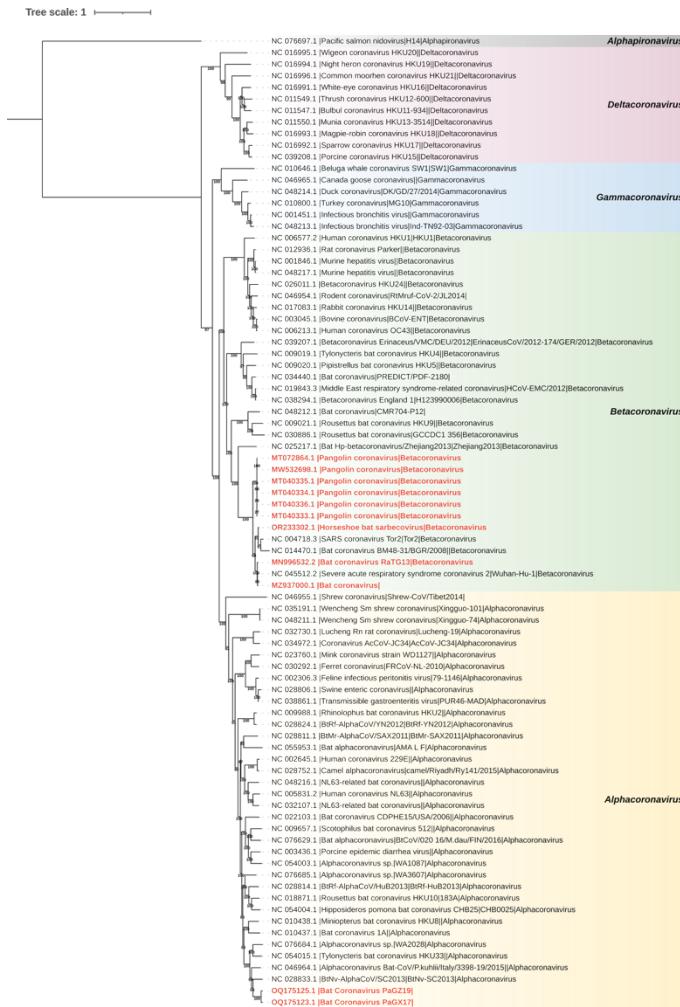


Figure 4. Phylogenetic tree of selected coronavirus based on whole genome. Selected coronaviruses were label as red color.

4. Cophylogenetic analysis using the RdRp and Spike protein

To evaluate the evolutionary concordance between conserved genomic regions and the receptor binding protein of virus, a cophylogenetic analysis was performed using phylogenetic trees constructed from the RdRp gene and the spike protein. The analysis shows that coronaviruses are broadly separated into the four established genera including alphacoronavirus, betacoronavirus, gammacoronavirus, and deltacoronavirus in RdRp tree, indicating overall consistency in evolutionary relationships. However, the spike-based tree shows greater topological

rearrangements, particularly within alphacoronavirus that separate into two clade, reflecting increased evolutionary plasticity of the spike protein. Several coronaviruses maintain similar positions in both trees, suggesting congruent evolutionary histories for RdRp and spike. In contrast, other viruses display discordant placements between the two trees, indicating differential evolutionary pressures acting on the spike protein relative to the conserved RdRp gene. Among the coronaviruses of interest, pangolin coronaviruses, bat coronaviruses, the horseshoe bat sarbecovirus, and bat coronavirus RaTG13 consistently cluster within the Betacoronavirus genus in both trees and remain closely associated with SARS-CoV-2, indicating broadly concordant evolutionary histories at both the genomic and spike levels. This consistency supports their close evolutionary relationships and high overall genomic similarity. In contrast, bat coronavirus PaGX17 and PaGZ19 cluster within the Alphacoronavirus genus in both trees and are clearly separated from the Betacoronavirus clade. Therefore, only nine coronaviruses consistently cluster with SARS-CoV-2 were subsequently selected for future analysis.

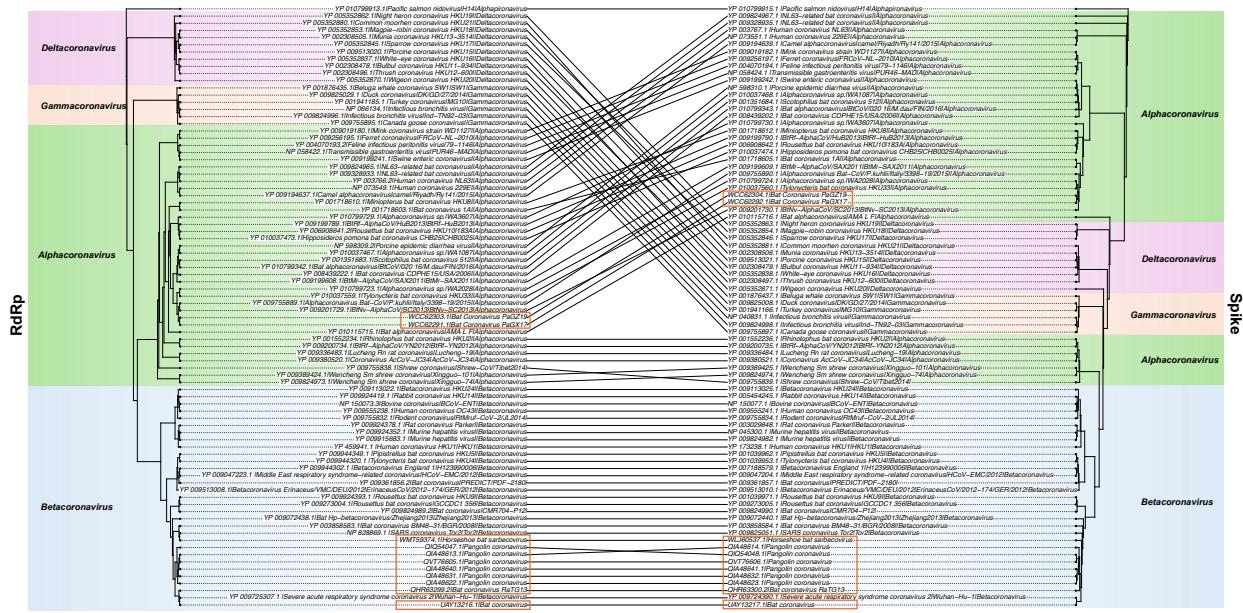


Figure 5. Cophylogenetic analysis using the RdRp (Left) and Spike protein (Right). The selected coronaviruses were indicated in the red box.

5. Similarity plot analysis

To examine sequence similarity patterns across the spike protein, a similarity plot analysis was performed comparing selected coronaviruses against SARS-CoV-2. The results reveal region-specific divergence patterns, with the most pronounced variability observed in the early N-terminal domain (NTD) and the receptor-binding domain (RBD). SARS-CoV-2 shows consistently higher sequence similarity to bat coronavirus and horseshoe bat sarbecovirus across most of the spike gene, whereas bat coronavirus RaTG13 exhibits reduced similarity within the RBD. Pangolin coronaviruses display moderate to high overall similarity to SARS-CoV-2 across much of the spike protein but show reduced similarity in the early NTD, followed by relatively stable similarity of approximately 90% across the reported AXL-binding region. This is followed by a marked decrease in similarity within the RBD and increased similarity in parts of the S2 subunit. Overall, the high sequence similarity observed at reported AXL-binding positions suggests that these viruses may retain structural features interacting with AXL receptor.

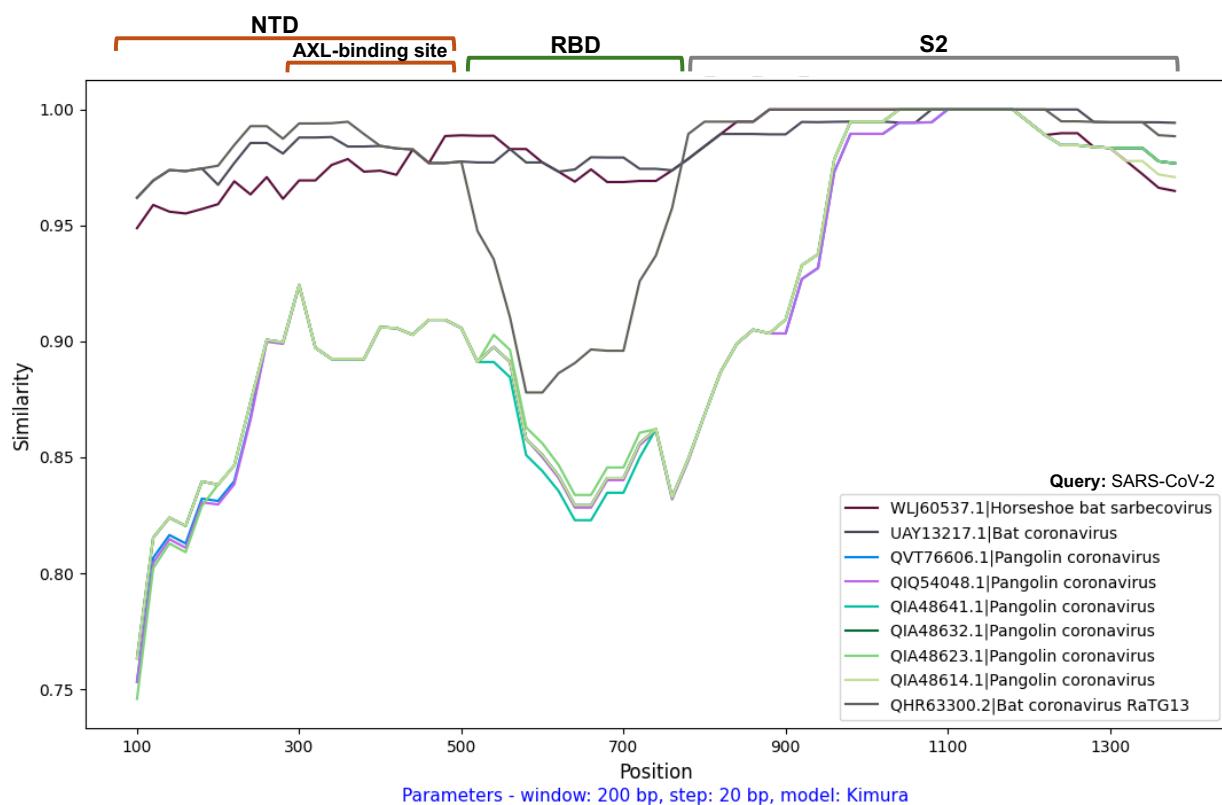


Figure 6. Similarity plot analysis

6. Molecular docking

To evaluate whether spike proteins from selected coronaviruses exhibit AXL-binding potential comparable to SARS-CoV-2, molecular docking analysis was performed using HADDOCK server. The results show that among the animal coronaviruses, the bat coronavirus (UAY13217.1) exhibits a HADDOCK score (-113.6) close to that of SARS-CoV-2 (-118.0), together with a low RMSD (1.6 Å), comparable van der Waals energy, and a similar buried surface area. Followed by horseshoe bat sarbecovirus (WLJ60537.1) with HADDOCK scores -107.8. Bat coronavirus RaTG13 also shows a similar interaction (-106.5), but with higher RMSD values, suggesting greater conformational variability in the docking solutions. Pangolin coronaviruses display a broader range of HADDOCK scores (-92.1 to -107.3). Several pangolin coronavirus spike proteins show interaction energies and buried surface areas comparable to SARS-CoV-2, but generally exhibit higher RMSD values. Overall, bat and pangolin coronaviruses demonstrate comparable docking result to SARS-CoV-2, supporting the potential of these coronavirus to bind with AXL receptor (Table 3).

Based on these findings, bat coronavirus UAY13217.1 was selected for detailed structural analysis. The bat coronavirus-AXL interaction was visualized and analyzed using ChimeraX. The results show a well-defined binding interface between human AXL (chain A, blue) and the spike protein of the bat coronavirus (chain B, pink), with interactions predominantly involving residues within the N-terminal domain (NTD). Interface analysis identified multiple hydrogen bonds that contribute to stabilization of the complex. Several polar and charged AXL residues, including GLU59, ARG64, GLN67, GLU70, and LEU71, form hydrogen bonds with NTD residues of the spike protein. Notably, GLU59 forms hydrogen bonds with HIS134 and ARG135 at short distances (1.7-2.9 Å), consistent with strong electrostatic interactions. In addition, ARG64 and GLN67 interact with ASP241 and SER243, while GLU70 forms hydrogen bonds with ARG234, THR238, and GLY245, and LEU71 interacts with GLY240. Importantly, key AXL residues involved in this interaction, including GLU59 and GLU70, overlap with residues previously reported in SARS-CoV AXL binding (Fang et al., 2023; Lei et al., 2023; Wang et al., 2021). Collectively, these interactions support the presence of a stable AXL-NTD binding interface in the bat coronavirus model, suggesting potential compatibility of this spike protein with the human AXL receptor (Table 4, Figure 7)

Table 3. The result of molecular docking analysis

Accession	Virus name	HADDOCK score	Cluster size	RMSD from the overall lowest-energy structure	Van der waals energy	Electrostatic energy	Desolvation energy	Restraints energy	Buried surface area	Z-score
YP_009724390.1	SARS-CoV-2	-118.0	28	1.1	-61.6	-349.3	2.3	111.4	1860.3	-2.3
UAY13217.1	Bat coronavirus	-113.6	11	1.6	-61.9	-257.5	-10.7	104.9	1766.5	-1.8
WLJ60537.1	Horseshoe bat sarbecovirus	-107.8	14	0.8	-55.6	-239.6	-11.8	76.4	1646.0	-1.2
QHR63300.2	Bat coronavirus RaTG13	-106.5	47	9.2	-60.7	-213.4	-13.3	101.8	1691.6	-1.2
QIA48632.1	Pangolin coronavirus	-99.5	16	19.8	-52.4	-257.6	-4.4	87.3	1927.1	-1.5
QIA48614.1	Pangolin coronavirus	-92.1	17	19.7	-49.0	-282.1	5.3	80.0	1610.4	-1.5
QIA48641.1	Pangolin coronavirus	-107.3	17	0.7	-58.0	-250.1	-6.2	69.6	1910.7	-1.9
QVT76606.1	Pangolin coronavirus	-104.3	10	20.4	-52.2	-291.2	-2.2	83.2	1907.1	-1.7
QIA48623.1	Pangolin coronavirus	-105.3	33	0.9	-56.1	-204.6	-17.2	89.1	1545.1	-1.8
QIQ54048.1	Pangolin coronavirus	-102.0	30	21.4	-43.3	-307.7	-9.4	122.4	1542.6	-1.6

Table 4. The structural analysis of bat coronavirus-AXL interaction

AXL residue (A: Blue)	Atom	NTD residue (B: Pink)	Atom	VDW Overlap	Distance (Å)
GLU59	HN	HIS134	ND1	0.687	1.953
GLU59	N	HIS134	ND1	0.338	2.927
GLU59	OE2	ARG135	HN	0.375	1.705
GLU59	OE2	ARG135	N	0.054	2.651
ARG64	NH1	ASP241	OD1	0.04	2.665
GLN67	HE21	ASP241	O	0.28	1.8
GLN67	NE2	ASP241	O	0.031	2.674
GLN67	NE2	SER243	HN	0.034	2.591
GLU70	OE2	ARG234	HE	0.315	1.765
GLU70	OE2	ARG234	HH22	0.078	2.002
GLU70	OE1	THR238	HG1	0.255	1.825
GLU70	OE1	GLY245	O	0.069	2.891
LEU71	HN	GLY240	O	0.049	2.031
GLN78	CD	ASN136	HD21	0.016	2.684

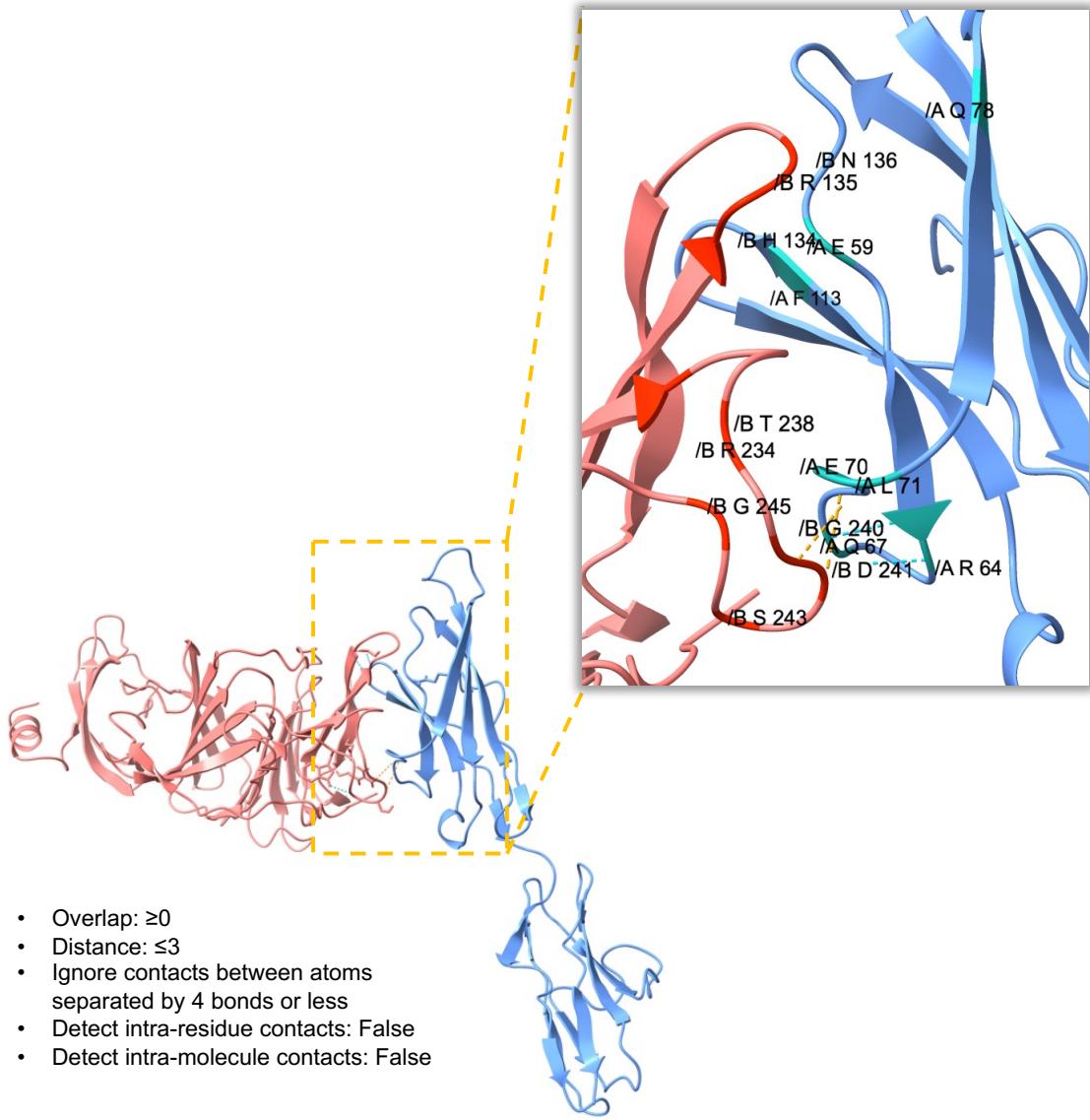


Figure 7. The structural analysis of bat coronavirus-AXL interaction

- Overlap: ≥ 0
- Distance: ≤ 3
- Ignore contacts between atoms separated by 4 bonds or less
- Detect intra-residue contacts: False
- Detect intra-molecule contacts: False

11. Thesis plan

12. Reference

Bernstein, A. S., Ando, A. W., Loch-Temzelides, T., Vale, M. M., Li, B. V., Li, H., Busch, J., Chapman, C. A., Kinnaird, M., Nowak, K., Castro, M. C., Zambrana-Torrelio, C., Ahumada, J. A., Xiao, L., Roehrdanz, P., Kaufman, L., Hannah, L., Daszak, P., Pimm, S. L., & Dobson, A. P. (2022). The costs and benefits of primary prevention of zoonotic pandemics. *Science Advances*, 8(5), eabl4183. <https://doi.org/10.1126/sciadv.abl4183>

Burrell, C. J., Howard, C. R., & Murphy, F. A. (2016). Coronaviruses. *Fenner and White's Medical Virology*, 437 - 446.

Escudero-Pérez, B., Lalande, A., Mathieu, C., & Lawrence, P. (2023). Host-Pathogen Interactions Influencing Zoonotic Spillover Potential and Transmission in Humans. *Viruses*, 15(3). <https://doi.org/10.3390/v15030599>

Fang, Q., He, X., Zheng, X., Fu, Y., Fu, T., Luo, J., Du, Y., Lan, J., Yang, J., Luo, Y., Chen, X., Zhou, N., Wang, Z., Lyu, J., & Chen, L. (2023). Verifying AXL and putative proteins as SARS-CoV-2 receptors by DnaE intein-based rapid cell-cell fusion assay. *Journal of Medical Virology*, 95(7), e28953. <https://doi.org/https://doi.org/10.1002/jmv.28953>

Gebreyes, W. A., Dupouy-Camet, J., Newport, M. J., Oliveira, C. J. B., Schlesinger, L. S., Saif, Y. M., Kariuki, S., Saif, L. J., Saville, W., Wittum, T., Hoet, A., Quessy, S., Kazwala, R., Tekola, B., Shryock, T., Bisesi, M., Patchanee, P., Boonmar, S., & King, L. J. (2014). The Global One Health Paradigm: Challenges and Opportunities for Tackling Infectious Diseases at the Human, Animal, and Environment Interface in Low-Resource Settings. *PLOS Neglected Tropical Diseases*, 8(11), e3257. <https://doi.org/10.1371/journal.pntd.0003257>

Grange, Z. L., Goldstein, T., Johnson, C. K., Anthony, S., Gilardi, K., Daszak, P., Olival, K. J., O'Rourke, T., Murray, S., Olson, S. H., Togami, E., Vidal, G., Expert, P., Consortium, P., Mazet, J. A. K., Anderson, K., Auewarakul, P., Coffey, L., Corley, R.,...Fine, A. (2021). Ranking the risk of animal-to-human spillover for newly discovered viruses. *Proceedings of the National Academy of Sciences*, 118(15), e2002324118. <https://doi.org/10.1073/pnas.2002324118>

Han, H.-J., Wen, H.-I., Zhou, C.-M., Chen, F.-F., Luo, L.-M., Liu, J.-w., & Yu, X.-J. (2015). Bats as reservoirs of severe emerging infectious diseases. *Virus Research*, 205, 1-6. <https://doi.org/https://doi.org/10.1016/j.virusres.2015.05.006>

Kreuder Johnson, C., Hitchens, P. L., Smiley Evans, T., Goldstein, T., Thomas, K., Clements, A., Joly, D. O., Wolfe, N. D., Daszak, P., Karesh, W. B., & Mazet, J. K. (2015). Spillover and pandemic properties of zoonotic viruses with high host plasticity. *Sci Rep*, 5, 14830. <https://doi.org/10.1038/srep14830>

Lei, M., Ma, Y., Chen, H., Huang, P., Sun, J., Wang, X., Sun, Q., Hu, Y., & Shi, J. (2023). Emerging SARS-CoV-2 variants of concern potentially expand host range to chickens: insights from AXL, NRP1 and ACE2 receptors. *Virol J*, 20(1), 196. <https://doi.org/10.1186/s12985-023-02123-x>

Pustake, M., Tambolkar, I., Giri, P., & Gandhi, C. (2022). SARS, MERS and CoVID-19: An overview and comparison of clinical, laboratory and radiological features. *J Family Med Prim Care*, 11(1), 10-17. https://doi.org/10.4103/jfmpc.jfmpc_839_21

Taylor, L. H., Latham, S. M., & Woolhouse, M. E. (2001). Risk factors for human disease emergence. *Philos Trans R Soc Lond B Biol Sci*, 356(1411), 983-989. <https://doi.org/10.1098/rstb.2001.0888>

Teeling, E. C. (2009). Bats (Chiroptera). In *The Timetree of Life* (pp. 0). Oxford University Press. <https://doi.org/10.1093/oso/9780199535033.003.0078>

Van Brussel, K., & Holmes, E. C. (2022). Zoonotic disease and virome diversity in bats. *Current Opinion in Virology*, 52, 192-202. <https://doi.org/https://doi.org/10.1016/j.coviro.2021.12.008>

Wang, S., Qiu, Z., Hou, Y., Deng, X., Xu, W., Zheng, T., Wu, P., Xie, S., Bian, W., Zhang, C., Sun, Z., Liu, K., Shan, C., Lin, A., Jiang, S., Xie, Y., Zhou, Q., Lu, L., Huang, J., & Li, X. (2021). AXL is a candidate receptor for SARS-CoV-2 that promotes infection of pulmonary and bronchial epithelial cells. *Cell Res*, 31(2), 126-140. <https://doi.org/10.1038/s41422-020-00460-y>

Wilson, D. E., & Reeder, D. M. (2005). *Mammal Species of the World*. Johns Hopkins University Press. <http://www.press.jhu.edu>