https://acadpub.com/abs

Acta
Biological Sciences

Identification and Characterization of a Novel Xi'an Ifla-like Virus Strain in Itch Mite (*Sarcoptes scabiei*)

Elele Kingsley, Onovo Izuchukwu Humphrey

Department of Bioinformatics, School of Health and Life Sciences, Teesside University, Middlesbrough, TS1 3BX, United Kingdom

About the Article

Open Access

Research Article

How to Cite: Kingsley E, Humphrey OI. Identification and characterization of a novel Xi'an Ifla-like virus strain in itch mite (Sarcoptes scabiei). Acta Biol Sci. 2025;1:28-41.

Keywords:
Functional domains, iflaviridae, phylogenetic analysis, polyprotein, RNA virus, *Sarcoptes scabiei*

Corresponding author:
Elele Kingsley
Department of Bioinformatics, School of Health and Life Sciences, Teesside University, Middlesbrough, TS1 3BX, United Kingdom
Tel: +447435987505

ABSTRACT

Objective: The goal of this study was to identify and characterize a novel Xi'an Ifla-like virus strain in Sarcoptes scabiei using information from the SRA dataset SRR19592729.

Materials and Methods: The viral genome, consisting of 7,388 nucleotides, encodes a polyprotein of 2,203 amino acids, shorter than the 2,612 amino acids in the closest known relative, the Xi'an Ifla-like virus (QPN36934.1). Codon usage analysis revealed the polyprotein is encoded in the 3'→5' frame, an atypical feature in positive-sense RNA viruses, suggesting distinct regulatory mechanisms. De novo sequencing produced 37,872 mapped reads.

Results: Despite 93.73% similarity at the protein level, the absence of the RNA-dependent RNA polymerase (RdRp) domain and multiple amino acid mismatches in conserved regions such as SF3 helicase and 3C protease support the classification of SRR19592729 as a new strain. De novo sequencing produced 37,872 mapped reads, revealing a 13.74% nucleotide divergence and frequent C→G and A→G substitutions. Phylogenetic analysis using the Neighbour-Joining method placed SRR19592729 within the Iflaviridae family but on a distinct branch from its closest relatives, with a node distance of 0.113.

Conclusion: These findings, including truncated genome features and structural variations, indicate that SRR19592729 represents a genetically divergent, possibly host-adapted viral strain. This sequence exhibits a hallmark characteristic of the Iflaviridae family having polyproteins such as capsid protein, SF3 helicase, as well as truncated protein sequence and missing RNA-dependent RNA polymerase. Therefore, continued surveillance and future functional studies are essential to understand its evolutionary origins, host specificity and potential implications for parasite-host-virus interactions.

INTRODUCTION

Viruses are obligate intracellular parasites composed of nucleic acids either DNA or RNA encased within a protein shell and in some cases, surrounded by a lipid envelope. Lacking the cellular machinery necessary for autonomous replication, viruses rely entirely on host cells for reproduction[1,2]. Despite their structural simplicity, viruses are among the most abundant biological entities on Earth and infect organisms across all domains of life, including bacteria, archaea, protists, fungi, plants and animals[3,4]. Their diverse genome architectures, high mutation rates and ability to adapt rapidly to new hosts or environmental pressures have rendered them essential agents of genetic variation and drivers of evolutionary innovation in cellular life forms[5-7].

The classification and study of viruses have evolved dramatically with the advent of Next-generation Sequencing (NGS) and metagenomics[8,9]. These tools have enabled the discovery of vast numbers of previously unknown viruses in a wide range of hosts and environments, expanding our understanding of viral taxonomy and ecology. Among the most diverse and extensively studied RNA

ACADEMIA Publications

viruses are those belonging to the order Picornavirales, which includes several families of medical, agricultural and ecological significance[10-12].

One family within Picornavirales, Iflaviridae, has garnered attention for its association with arthropods. Iflaviruses are small, non-enveloped viruses that possess a positive-sense single-stranded RNA (+ssRNA) genome, typically ranging from 8 to 11 kilobases in length[13,14]. Their genomes are monopartite and contain a single Open Reading Frame (ORF) that encodes a large polyprotein. This polyprotein is enzymatically cleaved into functional subunits such as capsid proteins, proteases, helicases and RNA-dependent RNA polymerase (RdRp), all of which are essential for the viral replication cycle[15-18]. Unique structural features include a 5 untranslated region (UTR) containing Internal Ribosome Entry Sites (IRES), which allow for cap-independent initiation of translation and a 3 UTR ending in a poly(A) tail that stabilizes the RNA genome and enhances translation efficiency[19-21].

Historically, iflaviruses were first isolated from economically important insects such as the honeybee (*Apis mellifera*) and the silkworm (*Bombyx mori*)[22]. However, with the rise of viral metagenomics, these viruses have been discovered across a broader range of arthropod taxa. New iflavirus-like sequences have been identified in representatives of Diptera (flies), Coleoptera (beetles), Hemiptera (true bugs) and Acari (mites and ticks)[23,24]. While many iflaviruses exist asymptomatically within their hosts, some have been associated with disease phenotypes affecting host development, behaviour and survival. Moreover, studies have revealed the capacity for recombination and host-switching among iflaviruses, suggesting their role in dynamic virus-host interactions and evolutionary plasticity[25,26].

Arthropods, the most speciose and ecologically diverse phylum in the animal kingdom, serve as reservoirs for a vast and largely unexplored diversity of RNA viruses[23,27]. Their frequent interactions with plants, animals and humans make them not only important vectors of disease but also critical players in ecosystem viral exchange[28,29]. The use of high-throughput sequencing platforms has allowed researchers to uncover novel viral genomes in arthropods, providing new insights into host-virus coevolution and viral biodiversity.

Sarcoptes scabiei, a parasitic mite responsible for scabies in humans and mange in animals, is one such arthropod of medical significance. This mite infests the skin of mammals, burrowing into the epidermis and causing intense itching, inflammation and lesions[30]. The global burden of scabies exceeds 200 million individuals at any given time, with the highest prevalence in low-resource settings[31]. Beyond the direct pathology, *S. scabiei* infestations predispose individuals to secondary bacterial infections, particularly with pathogens like *Staphylococcus aureus* and *Streptococcus pyogenes*, compounding disease severity and increasing the risk of invasive complications[32]. Despite its global importance, the biology and genomics of this mite remain poorly understood due to difficulties in maintaining viable mite colonies in laboratory conditions and the limited availability of high-quality molecular data[33]. Recent advances in genomic and transcriptomic analysis have begun to illuminate the molecular landscape of *S. scabiei*, including its symbiotic associations and microbial constituents. One understudied but potentially significant aspect of its biology is the virome, the complete community of viruses associated with the organism. Viruses within the host microbiome can have profound effects on physiology, immunity and disease susceptibility. In arthropods, this includes modulating vector competence, shaping immunity and influencing interspecific interactions[23]. However, little is currently known about the viral communities associated with mites such as *S. scabiei.*

A recent study reported the presence of a previously uncharacterised iflavirus-like virus in *S. scabiei* samples collected from Xi'an, China, hinting at the mite's potential to harbor novel RNA viruses[26]. This discovery underscores the importance of continued surveillance and characterisation of mite-associated viruses, both for understanding their ecology and for anticipating potential public health implications.

Building on these findings, the current study explores the virome of *S. scabiei* collected from a distinct geographic location, Melbourne, Australia. Using RNA-seq data and a suite of bioinformatic tools, this research aims to characterise potential ifla-like viral sequences associated with *S. scabiei*. The findings contribute to the expanding field of arthropod virology, enhancing our understanding of viral evolution, taxonomy and the complex relationships between hosts and their viral symbionts.

MATERIALS AND METHODS

Acquisition of datasets: The research commenced with the acquisition of raw sequencing data from the National Centre of Biotechnology Information (NCBI) database (https://www.ncbi.nlm.nih.gov/sra), specifically from the Sequence Read Archive (SRA), under the accession number SRR19592729. This RNA-seq dataset was generated from *Sarcoptes scabiei* (itch mite) adult females, a parasitic mite species responsible for causing scabies in humans and mange in animals. The dataset was submitted by the University of Melbourne, who had collected the mite samples as part of their ongoing research on the biology and health impacts of the organism.

The RNA-seq dataset was selected due to its potential to reveal novel or under-characterised viruses associated with arthropod hosts, especially those within the Iflaviridae family or other related viral groups that have not been thoroughly studied in this context. The viral sequences of

arthropods, particularly mites, remain largely unexplored, despite their potential to harbour a wide array of novel viruses.

The sequence data were downloaded in FASTQ format directly from NCBI's SRA portal. FASTQ format is a widely used file format that stores both sequence information (nucleotide sequences) and quality scores, which are crucial for assessing the reliability of the sequence data. The downloaded dataset served as the foundation for subsequent analyses, including sequence translation, alignment, annotation and phylogenetic assessments. High-throughput sequencing technologies used to generate this dataset enable the identification of genetic material from a broad spectrum of viruses that may infect the host, providing a rich source of information for the study.

Protein translation: The raw nucleotide sequence retrieved from the SRA dataset SRR19592729 was translated using the ExPASy Translate Tool (https://web.expasy.org/translate/) All six possible reading frames were scanned to identify potential Open Reading Frames (ORFs). The translation process was used to generate protein sequences, which were subsequently exported for downstream domain annotation, protein alignment and phylogenetic analysis. The identified ORFs were manually reviewed to select the most likely candidate for a complete polyprotein, considering frame continuity and sequence length.

Sequence alignment: The next critical step in the bioinformatics analysis pipeline was sequence alignment, which is crucial for identifying regions of similarity and homology between the sequencing data and known reference sequences. This was accomplished using two primary tools: BLAST N and BLAST P.

BLAST N (Basic Local Alignment Search Tool for Nucleotides) was employed to align the nucleotide sequence against the NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi). focusing on identifying regions of similarity between the query (the RNA-seq data from the *S. scabiei* sample) and the nucleotide sequences of known viruses. The tool facilitates the identification of homologous sequences and provides an E-value score that indicates the statistical significance of the match. An E-value of zero indicates a perfect match, while higher values suggest weaker alignments. By comparing the sequence to publicly available genomic data, we were able to narrow down possible viral candidates present in the S. scabiei sample.

Additionally, BLAST P (Basic Local Alignment Search Tool for Proteins) was used for aligning the translated protein sequences from the RNA data. This tool is designed to identify homologous protein sequences and determine the extent of similarity with known proteins from viral families. This alignment provided further insight into the functional domains of the viral proteins and helped classify the virus at a protein level, facilitating functional annotation.

After performing BLAST N and BLAST P, the CLC Genomics Workbench was utilised for further sequence alignment and visualization. The CLC tool allows for comprehensive data analysis, including multiple sequence alignments, visual inspection of sequence data and the construction of phylogenetic trees. The tool also enabled us to evaluate discrepancies in the sequence lengths between the translated protein and the polyprotein from the Xi'an Ifla-like virus reference. By aligning the sequences, we were able to gain a clearer understanding of how closely related the viral sequences from the S. scabiei sample were to those of previously identified viruses.

Protein mutations: To identify protein-level mutations and missing regions, the translated polyprotein from SRR19592729 was aligned against the reference polyprotein (QPN36934.1) using CLC Genomics Workbench. This alignment enabled direct comparison of amino acid sequences, highlighting mismatches and truncations. InterProScan was further employed to annotate and confirm functional domains within both protein sequences. The combined analysis helped detect conserved motifs, assess structural differences and pinpoint absent or altered domains in the novel polyprotein, particularly regions corresponding to essential viral enzymes such as RdRp.

Functional annotation

InterPro analysis: Functional annotation is a key step in understanding the biological roles of the proteins encoded by the viral genome. In this study, InterProScan (https://www.ebi.ac.uk/interpro/) was used to predict the functional domains and families present in the translated protein sequences. InterProScan is a powerful tool that integrates data from multiple databases such as Pfam, SMART and ProDom, to identify conserved protein domains, families and motifs. These domains are essential for determining the biological function of the protein and understanding its role in the virus's replication cycle.

The translated protein sequence from the RNA data was submitted to InterProScan, which compared the sequence against a wide array of known viral proteins and functional motifs. This comparison helped identify the presence of key functional domains typically associated with viral proteins, such as capsid proteins, RNA-dependent RNA polymerase (RdRp) and helicase. These functional elements are crucial for viral replication, host cell interaction and maintaining the structural integrity of the virus.

Functional annotation with InterProScan allows for a deeper understanding of the protein's potential role in the viral lifecycle. The results from InterProScan were cross-referenced with published research to validate the predictions and strengthen the understanding of the protein's functional significance. By identifying and annotating functional regions, this step provided valuable information that could be used for further experimental validation or comparative analyses.

Missing region analysis

HHpred analysis: After performing functional annotation, a further analysis was conducted to detect potential missing regions in the translated protein sequence. The tool HHpred (https://toolkit.tuebingen.mpg.de/tools/hhpred) was used to perform this analysis. HHpred uses hidden Markov models (HMMs) and profile-based techniques to detect evolutionary relationships between the query protein and a vast array of known structural and functional domains.

HHpred was employed to identify gaps in the protein sequence, particularly regions that may correspond to key viral functional domains that were missing or unannotated. The translated protein sequence was compared against a large database of known structural and functional protein families, focusing on evolutionary relationships that might suggest missing regions or functional motifs that are essential for viral function.

This missing region analysis was important for refining the overall understanding of the protein structure and functionality. By identifying potential missing elements, we were able to expand on the functional annotation and gather more information about critical components of the virus, such as the RNA polymerase, which is essential for RNA replication. The results from HHpred also provided further insight into the viral genome's completeness and suggested directions for further experimental work.

Evolution/phylogenetic analysis: Phylogenetic analysis was employed to explore the evolutionary relationships of the identified viral sequence and determine its position within the broader context of related viruses, particularly those from the Iflaviridae family. Phylogenetic trees are invaluable tools for understanding the evolutionary history and relatedness of different species or strains. By constructing a phylogenetic tree, we were able to visualise the relationships between the viral sequences from *S. scabiei* and other known viruses.

To generate the phylogenetic tree, the translated protein sequence from the SRR19592729 dataset was aligned with a selection of similar viral sequences from public databases. These sequences were selected based on their similarity and potential relevance to the study. The CLC Genomics Workbench was used to perform the alignments and construct the tree. The software facilitated multiple sequence alignments, ensuring that the sequences were accurately aligned based on evolutionary distance.

A rooted phylogenetic tree was constructed in this analysis. Rooting a tree is an important step as it helps provide a clear understanding of evolutionary relationships by placing a reference group (outgroup) in the tree, allowing for accurate interpretation of evolutionary distance. For this study, two families from the same order, Picornavirales, were selected as outgroups: Dicistroviridae and Marnaviridae. These families were chosen as outgroups because they are distantly related to the viruses being studied but share a common evolutionary ancestor. The inclusion of outgroups provides a stable baseline for understanding the evolutionary divergence of the virus and ensures that the tree is rooted accurately.

De novo sequencing (Map to Contigs): Finally, the RNA-seq dataset, which included paired-end reads (forward and reverse), was processed using the CLC Genomics Workbench for reference-based assembly. The nucleotide sequence served as the reference genome for the mapping process, which allowed for more accurate reconstruction of the viral genome by aligning the RNA-seq reads directly to the reference genome rather than assembling the genome from scratch.

The first step involved performing quality control on the raw sequence reads. This process involved filtering out low-quality reads and adapter sequences to ensure that only high-quality data were used for downstream analysis. The cleaned forward and reverse reads were then mapped to the reference genome using the Map to Contigs function in CLC Genomics Workbench. This function identified regions of overlap between the RNA-seq reads and the reference genome, enabling the construction of contigs that represent the genomic structure of the virus.

Mapping the sequence reads to a reference genome also allowed for the identification of potential sequence variations, insertions, or deletions relative to the known genome. The mapping process generated contigs that formed the basis for all subsequent analyses, including protein translation, alignment with known viral sequences and functional domain prediction. Coverage statistics were generated to assess the sequencing depth and evaluate any regions of low coverage or discrepancies. These statistics helped ensure that the entire viral genome was adequately represented and identified potential areas that required further analysis

RESULTS

Genome composition: The nucleotide sequence with SRR19592729 number possesses a genome with a total length of 7,388 nucleotides (nt). It encodes a single, large amino acid polyprotein of 2, 203 amino acids flanked by 3 and 5 untranslated regions (UTRs) as translated by Expasy (https://web.expasy.org/translate).The base pair composition analysis revealed that the sequence had an A /T ratio of 0.5548, a GC content of 0.4038, base pairs of A:T was 2,048 pairs, base pairs of G:C was 1,261 pairs, Paired bases count of 6,618 bases and an unpaired read count of 770 bases (Table 1)

The genome encodes a single polyprotein, which undergoes proteolytic cleavage to generate both structural and non-structural proteins. Using InterPro domain analysis,

several key proteins were identified. The analysis showed presence of a RHV-like domain, CRPV capsid protein-like domain, Superfamily 3 helicase of positive ssRNA viruses domain profile, Picornavirales 3C/3C-like protease domain profile and a 3C cysteine protease (picornain 3C) domain.

Table 1 presents the nucleotide composition and structural characteristics of the SRR19592729 sequence. The table includes details such as total base pairs, A:T and G:C content, paired and unpaired bases, A/T and G/C ratio and the total sequence length

Genomic mutations: The genomic mutations of the nucleotide sequence SRR19592729 were analysed using BLASTn alignment and de novo assembly in CLC Workbench. The BLASTn search returned a single significant match, aligning with a known Xi'an Ifla-like virus strain (Accession number MT757479.1) with 86.26% sequence identity and an E-value of 0.0, confirming a close relationship to this reference strain. De novo sequencing produced a total of 37,872 mapped reads, with an average coverage depth of 4728-fold which were assembled into contigs and analysed for genetic variations. The sequence divergence of 13.74% from the reference strain consisted of mismatches (substitutions), insertions and deletions (indels). Further analysis revealed Single Nucleotide Variants (SNVs), with C→G and A→G substitutions being the most frequent (Fig. 1a and b). These mutations, representing both transversions and transitions, may impact viral protein structure, replication efficiency, or host adaptation. Given that RNA viruses exist as quasi species, these variations likely result from the error-prone replication of RNA-dependent RNA polymerase (RdRp). The observed mutations could play a role in viral adaptation, immune evasion and persistence within the itch mite (*Sarcoptes scabiei*) host. Further structural and functional analysis is required to determine their biological significance of these mutations, as well as their impact on viral fitness and pathogenicity. The summary statistics of the de novo sequencing on CLC workbench showed a total reads count of 69, 067.366, total percentage read of 100.0 was noted, total read average length of 99.36 was observed, total read number of bases was 6, 862, 543, 894 and total read percentage of bases stood at 100.00; the broken paired reads also showed a count of 11, 380, percentage of reads 0.02, average length 98.89, number of bases 1, 125, 408 and percentage of bases 0.02%; the reads in pair also revealed a count of 26, 492, percentage of reads was 0.04, average length was 214.63, number of bases observed was 2, 636, 369 and percentage of bases was 0.04%; furthermore, the not mapped reads had a count of 69,029,494 percentage reads of 99.95, average length of 99.36, number of bases,6,858,782,118 and percentage of bases at 99.95; the mapped reads count stood at 37,872 percentage of reads, 0.05, average length of 99.33, number of bases showed 3,761,177 and percentage of bases at 0.05; finally, References had a count of 1, nothing showed for percentage of reads, average length was 7,388.00, number of bases, 7,388 and nothing for percentage of bases (Table 2).

An overview of the de novo sequencing results obtained from Map to Contigs analysis using CLC Workbench.

Protein mutations: The protein sequence of the virus was analysed by translating the nucleotide sequence SRR19592729 and performing a BLASTp alignment against known viral protein sequences. Based on the BLASTp results, the translated protein was identified as a polyprotein from the Xi'an Ifla-like virus (Accession number QPN36934.1) with 93.73% sequence identity and an E-value of 0.0, indicating a strong similarity to related strains.

Further analysis was performed by aligning the translated protein with the polyprotein from the BLAST result using CLC. This alignment revealed a variance in length, with the polyprotein being 2612 amino acids long, while the translated protein was shorter, at 2203 amino acids. To investigate the missing length in the translated protein, an additional HHpred analysis was conducted which identified the presence of an RNA-dependent RNA polymerase (RdRp) domain with a probability score of 100, E-value of 4.7e-51, score of 395.73, secondary structure score (SS) of 42.9 and 389 aligned columns. This domain was not detected in the InterPro analysis of the translated protein sequence.

The InterPro analysis also identified other important domains, including the capsid protein, 3C picornavirus protease and Superfamily 3 helicase, which are critical for the structure and replication of the virus.

Evolution: A phylogenetic analysis was performed to determine the evolutionary relationship of the Xi'an Ifla-like virus using the translated protein sequence of SRR19592729. The analysis was conducted using CLC Workbench with the Neighbor-Joining (NJ) method and the Jukes-Cantor protein distance model. A total of 100 bootstrap replicates were used to assess the tree's reliability. The translated protein from SRR19592729 clustered with the Xi'an Ifla-like virus, confirming its classification within the Iflaviridae family. A distinct branch was observed between the translated protein and the Xi'an Ifla-like virus polyprotein, which was separate from all other branches in the tree.

Table 1: Nucleotide composition and sequence characteristics

Base pairs	SRR19592729
A : T	2048
G : C	1261
Paired bases	6618
Unpaired base	770
A/T Ratio	0.5548
GC content	0.4038
Total sequence length	7388 nt

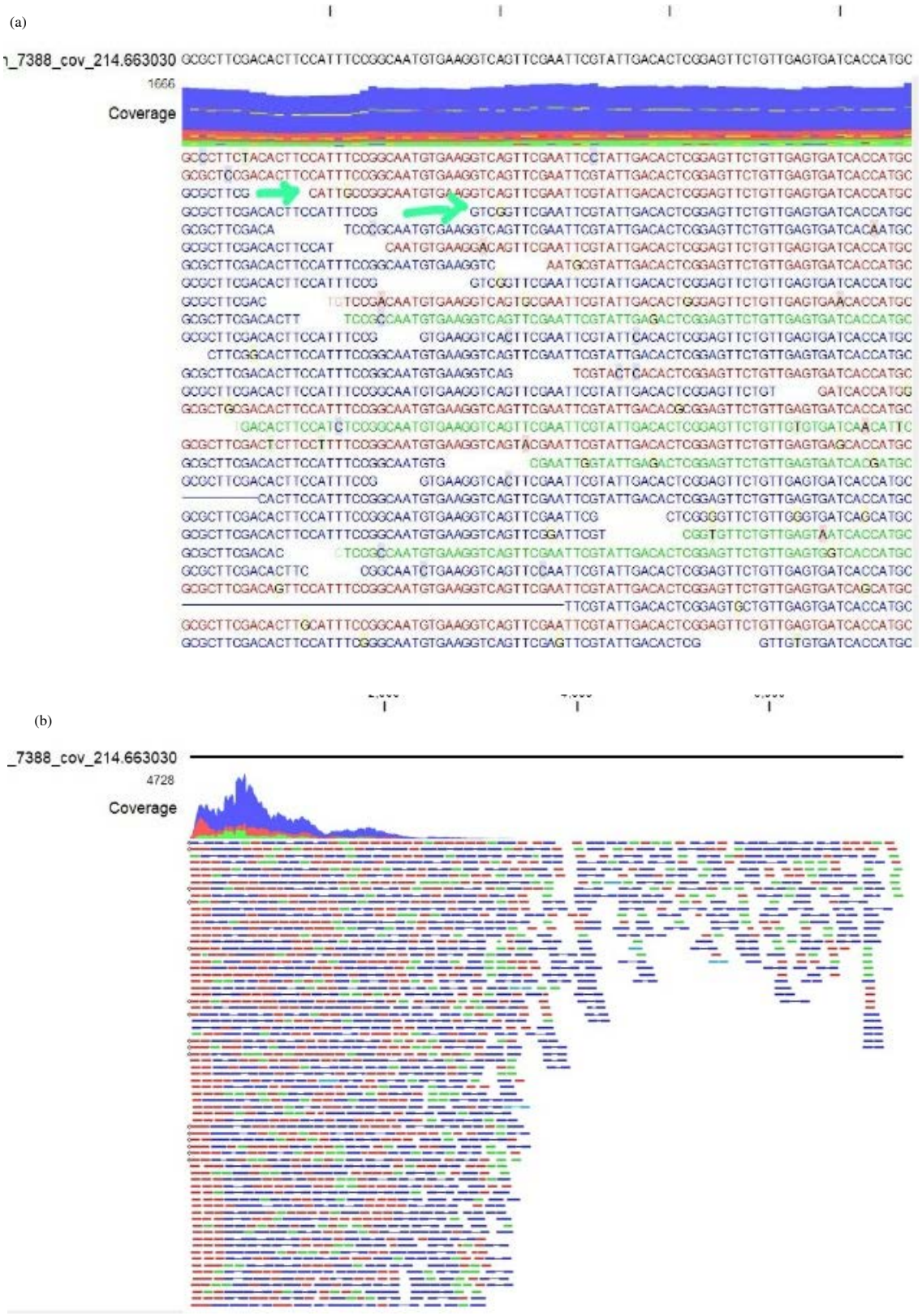


Fig. 1(a, b): Coverage and mismatch analysis from de novo sequencing.

Mapping of Illumina paired-end reads (SRR19592729) to contigs using CLC Genomics Workbench. The figure shows reference-guided De novo assembly performed using high-throughput Illumina sequencing data. Forward and reverse reads were imported and mapped to contigs derived from SRR19592729. Mapping parameters included a minimum distance of 1 and maximum distance of 100 between paired reads, match score of 1, mismatch cost of 2 and insertion/deletion cost of 3. Length and similarity fractions were set at 0.95, with reads mapped randomly. Arrows highlight frequent C→G and A→G single nucleotide mismatches identified during assembly.

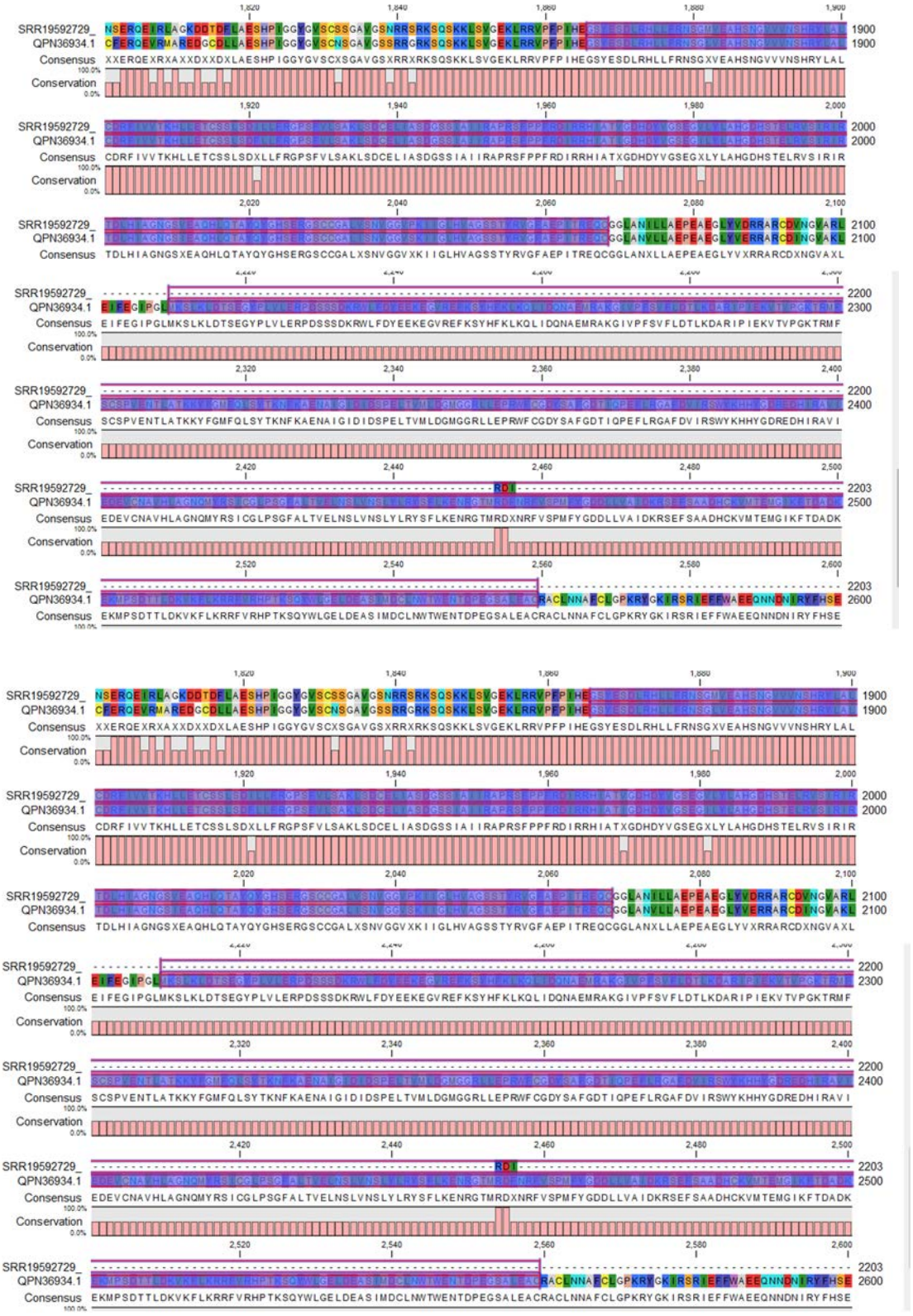


Fig. 1c: Annotated protein alignment between the translated polyprotein (SRR19592729) and Xi'an Ifla-like virus polyprotein (QPN36934.1).

Highlighted regions indicate conserved functional domains in both sequences, with matching positions aligned by amino acid number. The Rhv-like structure is encoded between 363-500, while the capsid protein spans residues 682-892. The SF3 helicase domain occurs at 1203-1371 and the 3C-like protease domain, containing a trypsin-like serine protease, is located between 1866-2068 in both sequences. A DNA/RNA polymerase domain is detected at 2101–2185 in the translated protein, while in QPN36934.1, the corresponding RdRp region spans 2210-2559. The translated protein terminates at residue 2203, lacking the full RdRp domain observed in the reference polyprotein.

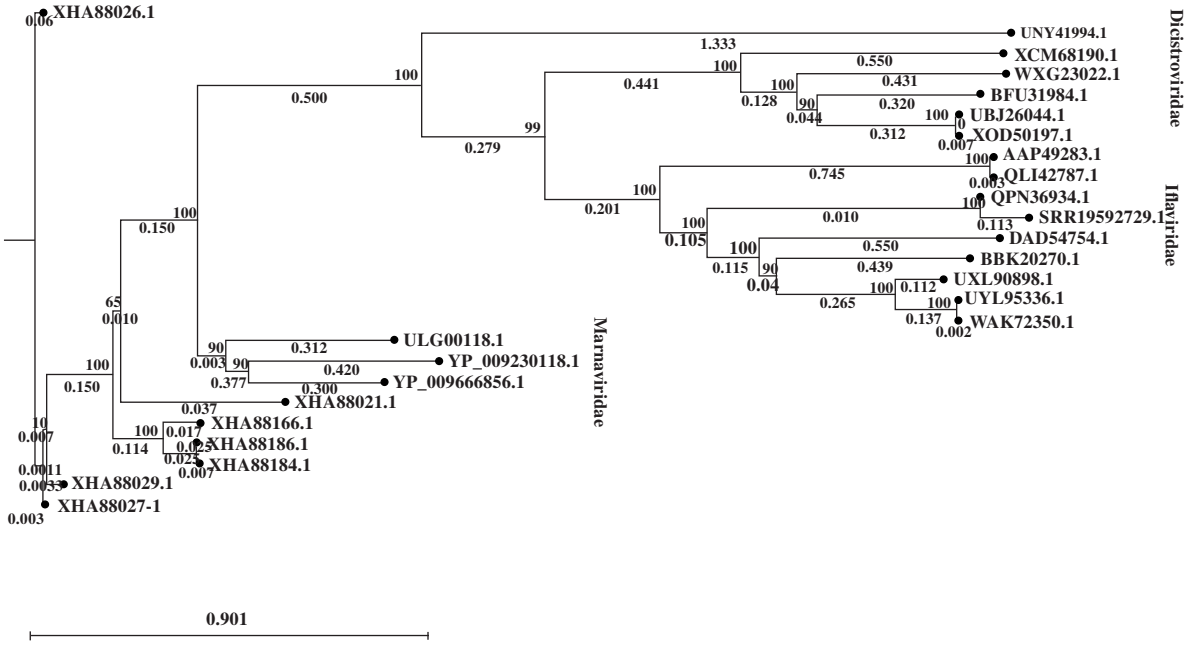


Fig. 1d: Phylogenetic tree showing the evolutionary relationships between the translated polyprotein of (SRR19592729) and closely related viruses from the Iflaviridae family.

The tree was constructed based on protein sequence alignments using the Neighbour-Joining method with the Jukes-Cantor distance model and 100 bootstrap replicates. The translated polyprotein from SRR19592729 clusters within the Iflaviridae family, alongside the polyprotein from the BLASTp search results. The tree was rooted using outgroups from the Dicistroviridae and Marnaviridae families, ensuring accurate evolutionary comparison. Branch lengths represent evolutionary distances and bootstrap values at the branch points provide support for the clades. The root is clearly indicated to avoid any ambiguity in the evolutionary analysis, highlighting the position of the novel virus relative to other Iflaviruses, including Xi'an Ifla-like virus (QPN36934.1).

Table 2: Summary statistics of de novo sequencing (Map to Contigs) on CLC workbench

	Count	Reads (%)	Average length	No. of bases	Bases (%)
References	1	7	7,388.00	7,388	-
Mapped Reads	37,872	0.05	99.33	3,761,777	0.05
Not mapped reads	69,029,494	99.95	99.36	6,858,782,118	99.95
Reads in pais	26,492	0.04	214.63	2,636,369	0.04
Broken paired reads	11,380	0.02	98.89	1,125,404	0.02
Total reads	69,067,366	10000	99.36	6,862,543,895	100.00

The node distance between QPN36934.1 (Xi'an Ifla-like virus polyprotein) and SRR19592729 was 0.113, indicating the evolutionary distance and the amount of genetic change that has occurred between these two sequences. Additionally, Dicistroviridae and Marnaviridae from the Picornavirales order were included as outgroups. The phylogenetic tree clearly separated the viruses into their respective families, reinforcing the classification of the Xi'an Ifla-like virus within Iflaviridae.

DISCUSSION

Taxonomy of the translated protein: The analysis of SRR19592729 provides compelling evidence that the virus represents a new strain within the Iflaviridae family. Iflaviruses are known for their single-stranded, positive-sense RNA genomes and this virus shares several hallmark features with other members of the family, such as polyprotein structure encoding essential proteins like the SF3 helicase, 3C protease and a capsid protein[34]. The complete genome length of 7,388 nt and the presence of these signature proteins strongly suggest that this virus is an Iflavirus variant. The sequence GC content and A/T ratio values align with ranges observed in other positive-sense RNA viruses but are sufficiently specific to support its differentiation[35]. Phylogenetic analysis further supports this affiliation, as the virus clusters closely with other Iflaviruses, particularly the Xi'an Ifla-like virus.

Furthermore, the hallmark of the genome and proteome features support the classification of SRR19592729 as a new strain within the Iflaviridae family, rather than a simple variant of the Xi'an Ifla-like virus. Most notably, the translated polyprotein of SRR19592729 is shorter, consisting of 2,203 amino acids, compared to 2,612 in the reference Xi'an Ifla-like virus (QPN36934.1). This

difference is particularly evident in the absence of the RNA-dependent RNA polymerase (RdRp) region in the terminal segment, suggesting either an incomplete genome or a naturally truncated variant. The four mismatches in the SF3 helicase region and eight mismatches in the 3C protease region of the protein further emphasize these structural differences, indicating minor yet consistent amino acid substitutions in conserved motifs. These substitutions, found in critical domains of the virus, could have significant implications for its functional properties, such as replication efficiency and host interactions.

Additionally, differences in codon usage or ORF arrangement were observed in SRR19592729. The protein is encoded in the 3'→5' reading frame (Frame 1), which is atypical for positive-sense RNA viruses and may reflect altered regulatory or translational mechanisms, further supporting the divergence of SRR19592729 from the reference strain.

The overall protein-level similarity between SRR19592729 and the reference polyprotein was 93.73%, with 100% coverage and a significant E-value of 0.0, suggesting a close but distinct relationship. Furthermore, nucleotide-level comparison revealed a 13.74% divergence from the reference strain, including numerous non-synonymous mutations and insertions/deletions (indels), which were identified through CLC Workbench analysis. Frequent C→G and A→G transitions were observed, potentially contributing to structural or functional variation[36,37].

These findings are in line with the known high mutation rates of RNA viruses, which are a consequence of the error-prone activity of RNA-dependent RNA polymerase[38]. The translated protein sequence also displayed slight deviations in domain positions and alignment breakpoints relative to the reference virus, as confirmed by InterProScan outputs. These cumulative differences, though subtle individually, collectively reflect evolutionary divergence and support the conclusion that SRR19592729 represents a new strain within the Iflaviridae family.

The use of BLASTn and BLASTp for nucleotide and protein sequence comparison revealed that the virus exhibits high levels of sequence homology with other Iflaviruses, suggesting that it belongs to the broader viral clade. Nonetheless, the strain-specific mutations and its divergence in key genomic regions establish SRR19592729 as distinct from its closest relatives, particularly in its interaction with the host. This underscores the dynamic and ever-changing nature of viral genomes, especially within a family as diverse as Iflaviridae[39].

Phylogenetic analysis, based on both nucleotide and protein sequence alignments, reveals that SRR19592729 sits within the evolutionary tree of the Iflaviridae family but is sufficiently distinct from existing strains. The virus's position within the tree suggests that it shares a common ancestor with other Iflaviruses but its divergence points to an evolutionary pathway that warrants its classification as a new strain.

Function of the translated protein: The functional elements of SRR19592729 were extensively analysed to understand its replication, maturation and potential pathogenicity. Like other Iflaviruses, this virus encodes a polyprotein that is cleaved into functional viral proteins. The capsid protein, SF3 helicase and 3C protease are typical features found in members of the Picornavirales order and their presence in SRR19592729 suggests that it shares essential replication and structural functions with other viruses in the family[15]. The capsid protein plays a critical role in viral entry into host cells, encapsulating the viral RNA and facilitating its protection and delivery to the host cell machinery[40]. The SF3 helicase, involved in unwinding RNA, is crucial for viral replication, as it facilitates the synthesis of new RNA genomes from the viral RNA template. Similarly, the 3C protease is vital for the maturation of the viral polyprotein, which is cleaved into smaller functional units during the viral replication cycle[41].

However, an intriguing finding in the study was the identification of a truncated protein in SRR19592729, particularly in the region encoding the RNA-dependent RNA polymerase (RdRp). The absence of the RdRp domain in the protein sequence could imply an incomplete genome assembly or a naturally truncated viral variant. The RdRp region is highly conserved among RNA viruses and its absence in the translated protein sequence is unusual for a positive-sense RNA virus[40]. This missing segment was further confirmed through alignment with the Xi'an Ifla-like virus, which contains a complete RdRp region.

The potential loss of this critical region raises several questions. It may suggest that the virus has evolved a different mechanism for replication, potentially using a host-derived polymerase or adopting alternative replication strategies[42]. Alternatively, the missing RdRp segment could be a result of sequencing or assembly artifacts. While this would require further validation through long-read sequencing technologies, such as Oxford Nanopore or PacBio, the possibility of an incomplete genome assembly also brings attention to the challenges faced when studying novel viruses using short-read sequencing methods[43].

Notably, the absence of the RdRp region does not invalidate the virus's ability to replicate. Many RNA viruses exhibit diverse strategies for replication and transcription. For instance, some viruses rely on

alternative translation mechanisms, such as ribosomal frameshifting or read-through events, to produce functional viral proteins[44].

Functional domain analysis of the polyprotein revealed additional insights into the virus's ability to replicate and process its proteins. The protease and helicase domains play essential roles in virus maturation, enabling the cleaving of the polyprotein into smaller functional units. The presence of these domains suggests that the virus utilizes a well-conserved mechanism to process its polyprotein into active proteins, enabling successful replication and assembly of the viral particle. These features are commonly seen across multiple viral families, highlighting the importance of these domains in viral life cycles[45].

The VP1 capsid protein also raised questions regarding viral host interaction. The unique amino acid substitutions in this protein could affect its ability to interact with the host cell machinery, particularly in terms of immune evasion. Such substitutions may enhance the virus's ability to persist in the host or increase its virulence. The structural diversity in the capsid protein could potentially affect the virus's stability, infectivity and its interaction with host immune systems, emphasizing the need for further investigation into the protein's function[34].

Itch mite (*Sarcoptes scabei*): The identification of *Sarcoptes scabiei* as the host for SRR19592729 presents significant insights into the virus's ecology and host interaction. *Sarcoptes scabiei*, the itch mite, is an arthropod that acts as a natural reservoir for various pathogens, including viruses[46]. The presence of SRR19592729 in this host suggests that the virus has adapted to the arthropod environment and the observed genetic variations may be linked to its interaction with the host's immune system.

The genetic mutations identified in the virus, particularly the single nucleotide variants (SNVs), may enhance the virus's ability to adapt to the host. These mutations could play a role in immune evasion, allowing the virus to persist in the host without triggering an overt immune response. This adaptation could facilitate the establishment of a chronic infection, allowing the virus to continue replicating and evolving within the host population. The error-prone replication process of RNA viruses, driven by the RNA-dependent RNA polymerase (RdRp), results in a high mutation rate, leading to viral diversity[47]. This diversity is advantageous for the virus, as it allows rapid adaptation to changing environmental conditions and host immune pressures.

Interestingly, the geographic separation between the virus found in Melbourne, Australia and its closest relative in China raises intriguing questions about the virus's global distribution. This suggests that SRR19592729 might be part of a broader viral lineage that spans different geographic regions, possibly reflecting a widespread presence of Iflaviruses in *S.scabiei* populations[25]. The viral mutations identified in the study, particularly the presence of C→G and A→G substitutions, could influence the virus's ability to replicate and transmit across different populations. This further highlights the potential for SRR19592729 to evolve independently in different geographic regions, suggesting the possibility of distinct viral strains emerging in separate populations[48].

While the virus's ability to infect *Sarcoptes scabiei* is established, its infectivity in other potential hosts remains speculative. The base composition of the sequence can influence viral replication kinetics, codon preference and RNA structure within the host, thereby contributing to the virus's persistence and host-specific fitness[35]. The host range of the virus, which includes the mite as a primary host, may extend to other arthropods or even mammals, though further studies are needed to confirm this. The high mutation rate in RNA viruses like SRR19592729 often facilitates cross-species transmission, which could enable the virus to adapt to new hosts over time.

CONCLUSION

The SRR19592729 sequence represents a new strain within the Iflaviridae family, displaying distinct genetic divergence from its closest relatives. The virus exhibits hallmark characteristics typical of Iflaviruses, including a polyprotein encoding essential viral proteins such as the capsid protein, SF3 helicase and 3C protease but also shows unique genetic features, including a truncated protein sequence and missing RNA-dependent RNA polymerase (RdRp) region. This new strain appears to be well-adapted to its *Sarcoptes scabiei* host, with evidence suggesting that its high mutation rate could play a role in host adaptation and immune evasion. Additionally, the genome's base composition characteristics-specifically its GC content of 40.38% and A/T ratio of 0.5548 highlight a genomic profile that may confer replication adaptability and translational optimization within the *Sarcoptes scabiei* host.

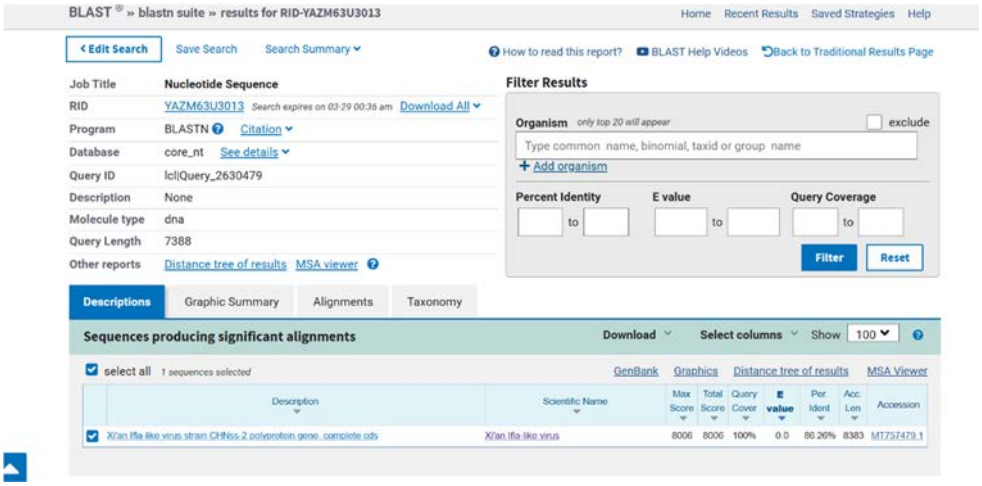
Given the geographical separation of this virus from its closest relatives, it is likely that this strain has evolved independently, although future studies will be necessary to determine its full evolutionary history and potential for cross-species transmission. The findings contribute to the growing body of knowledge about Iflaviridae and underscore the importance of continuing to explore viral evolution, especially in arthropod vectors.

Future research should focus on sequencing the remaining regions of the genome using long-read technologies and performing infectivity assays to confirm the virus's host range and pathogenicity. Structural modeling of the viral proteins, particularly the VP1 capsid protein, could also provide insights into how the virus interacts with host cells and immune systems. This comprehensive understanding will be crucial for assessing the broader implications of this new strain and its potential impact on animal and human health.

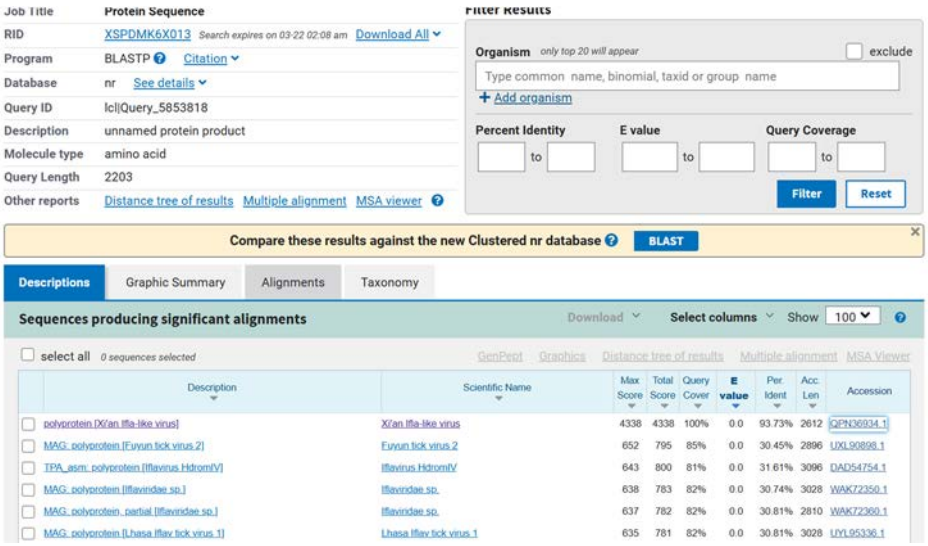
APPENDIX

```
ID   VIRT-2674220          Unreviewed;        2203 AA.
AC   VIRT-2674220;
DE   Translation of nucleotide sequence generated on Expasy
DE   3'5' Frame 1, start_pos=102
DE   on Thu Mar 27 19:09:54 CET 2025 by 192.33.215.76
CC   -!- This virtual protein sequence will automatically be deleted
CC       from the server after a few days.
SQ   SEQUENCE   2203 AA; C74340ED3ECA78DE CRC64.
     MSQESNVAPI SGLNTEEQVS SVLLESGVKS VDSIMAPPTP AQVVEGTLTS LGSDLTNRWN
     LLERFDWSVT SSFGENLKTW KIPEDLFESS LLKINLLPYK VYCYGNCGVF LRLQVNANRF
     HVGQLQFSWY YHGLDDQDFA LRDNVYCASQ RPNGIFQVGT SQTCELYIPF KSPLTHLTFR
     KESMEPFEAK NFALVQLKVL SALSTSEASL KSASVTVMFR LADTNFMGLR SIIAEPQMLP
     MRMVVKAAEA ALNEAGKEFN RDFPTNPFNK LSVVPTGSQS WASGIGDVIS LNSLRLDPRA
     TTPHPNYRND EMLVSNITNH WGLIGKVQWR AFQDQQLTSL PVAPVHSKLL RLRSRASFTQ
     GLLSTRILAP TSVVASMFLY WRGPIEFRFD IVVGPFHTGR LVASFIPGPK KSVSVQEMMQ
     SYYSVFDVTD QTSFIFEVPY IANKSWQYVT QGQRNYGDEG IPIGYLYLSS LNQLVVCDGV
     SRNVDILIYM RAPKIEFSVP CNCLVGLGRD REFFHPSYMR IKTGYTSIYA GNWHDFALSA
     KIICRYGPGS DHIVQVEGGK LDEVYISEPL KNVTFKDAQG KHVADLNFYA IANHKEFGDY
     RYLIPFDTLE KARAYVTSGA NGSANWDLCL DYVSDSAYND GSKWMFVPVE VQEKFEFIEP
     QGDEVKEEVK AQPIKTIAHP SALSFNEDFS DLKSLCRRFN VYFESQYEIT EGYELGRYFC
     KLPLRVDGLK MNFARDRRAS ALQNRLRNGM IAVLSSGFRF FRGSLRFKIV VDTETDSTLM
     VTHIPDETNF TECRVLSESN ADYVGAGYAC YGQSMQVNRV LEIEVPYYLP YEFGLIGNVS
     SSSGLGRSAS LGCLGFSLLT KPRKKIAIDV MVYYAIGDDM RFSSFMGFPE MVLRTEIPDI
     EHQGLFGLFP DAEKISKDVA GEVVGSVEEI KTKVNDSVNL ITDRTNSVLS NIHHRAEDVM
     TKVLEVLSSL AKGFGDKSSD FLQDCAFELA HLVANPNIRT LGVSLAKILV KLGIITSEFF
     SALSKWAGEA ITWALPQKQQ EGEVVHQGLS NEGITLMSTL IGMISSSLNI AVNAPRRGFL
     SYLACSLSSG LRDMAMTSNQ MLIFLKNNLS VINSIWYWLL GQADPSASYI RDLTESMPQL
     ENWAKRGEEL TRLVNRDEVK RDPIKQKDAF ICAALGSAIV MRMSLSGVTD PRMNYIRSLA
     SKLDKLQDEL KSECLSPMVK FEPFVFQLSG PPGIGKSFMS ESLGRRLLKK VGHRYYGEST
     FVRTPGVAYW NGVRNQPILI YDDLGCTATA LKEQVGEIFS IKSPQPFNPP IAELENKKLR
     YNPFIMALCS NVSYFRDPEI TDPVAFNRRR DTLIECRLTE SVKRKYGAGV TARMIGSQDL
     ENYGHLEFRF STNPSNEAGP YSDWMRYEEM LSILEKQFEI YYNAERKNYI RKLTELGDGY
     PTDDHAISLA DLARMVNNDF GDIDWLMETD TREIMRTFAV QEQPALLAEP QGLNVVMDLA
     KDEATGLLKI DDTVEFDMSY EEIPWNSSLM CAHRGIKKAY AFFRKPGQKY GVFYDPMRHA
     DICLEEDPII GISEAPCSNC LLHDQENVVK LISQYYVNFI YVRTNGPIKI QNKIVPLALR
     NFHVNVQVID PFLLPEKKFV LTDKMKAFWE KVTSTTLKVV MKVWGLLKPC LGWLRKSFKY
     LVGTFALAGA AVMITRPFRK SEVDDSPTGF ADERVATRTN RSSDCGLLSN KNAGARSVAD
     HPLLSKILAD REKTKVFASE DNSFCKNVFG SRALSNLPVR LWSPSNQEYS ISTDRGGDSE
     NSERQEIRLA GKDDTDFLAE SHPIGGYGVS CSSGAVGSNR RSRKSQSKKL SVGEKLRRVP
     FPIHEGSYES DLRHLLFRNS GMVEAHSNGV VVNSHRYLAL CDRFIVVTKH LLETCSSLSD
     ILLFRGPSFV LSAKLSDCEL IASDGSSIAI IRAPRSFPPF RDIRRHIATV GDHDYVGSEG
     VLYLAHGDHS TELRVSIRIR TDLHIAGNGS VEAQHLQTAY QYGHSERGSC CGALVSNVGG
     VPKIIGLHVA GSSTYRVGFA EPITREQCGG LANILLAEPE AEGLYVDRRA RCDVNGVARL
     GVVDDSFTVR LPYRSRIVPT LLHDRIYKHV TEPAPLSRDD PRIAGGSGRS PFQCALDKIG
     RTPGSVDIGL LNRAIADLTA KVRYECVPIL PTHGELLTKK RDI
//
```

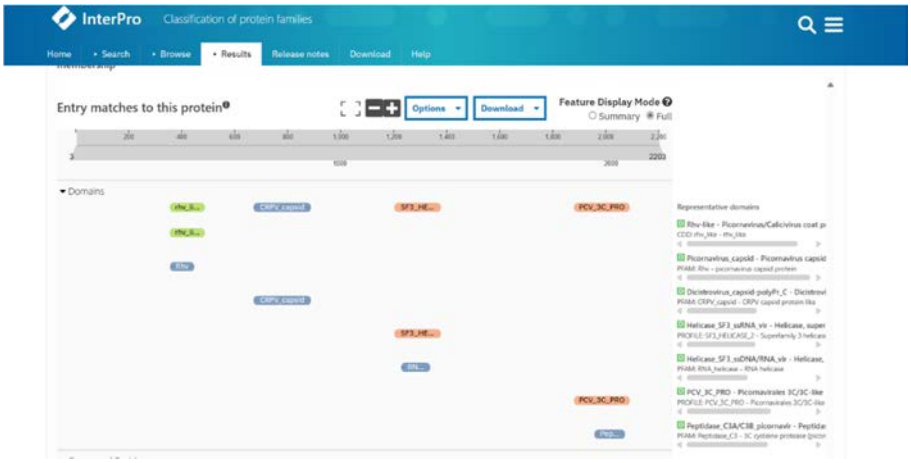
Appendix 1: Translation of SRR19592729 using Expasy



Appendix 2: BLAST N alignment of SRR19592729



Appendix 3 : BLAST P alignment of SRR19592729



Appendix 4: Inter Pro scan analysis of the translated Protein

REFERENCES

1. Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR, Carstens EB et al. Virus taxonomy in the age of metagenomics. Nat Revs MicroBiol. 2017;15:161-168.
2. Defilippis VR, Villarreal LP. An Introduction to the Evolutionary Ecology of Viruses. In: Hurst CJ, editor. *Viral Ecology*. Cincinnati, Ohio: Elsevier; 2000. p. 125-208.
3. Koonin EV, Dolja VV. A virocentric perspective on the evolution of life. Curr Opin Virol. 2013;3(5):546-557.
4. Suttle CA. Marine viruses-Major players in the global ecosystem. Nat Revs MicroBiol. 2007;5:801-812.
5. Lauring AS and ino R. Quasispecies theory and the behavior of RNA viruses. PLoS Pathog. [Internet]. 2010;6. Available from: https://doi.org/10.1371/journal. ppat.1001005
6. Forterre P, Prangishvili D. The major role of viruses in cellular evolution: Facts and hypotheses. Curr Opin Virol. 2013;3:558-565.
7. Broecker F, Moelling K. What viruses tell us about evolution and immunity: Beyond darwin? Ann NY Acad Sci. 2019;1447:53-68.
8. Tan JK, Servellita V, Stryke D, Kelly E, Streithorst J, Sumimoto N et al. Laboratory validation of a clinical metagenomic next-generation sequencing assay for respiratory virus detection and discovery. Nat Commun. [Internet]. 2024;15. Available from: https://doi.org/10.1038/s41467-024-51470-y
9. Kosmopoulos JC, Anantharaman K. Microbial and Viral Ecology Analysis for Metagenomic Data. Quant Biol. [Internet]. 2024.
10. Zell R, Delwart E, Gorbalenya AE, Hovi T, King AMQ, Knowles NJ et al. ICTV virus taxonomy profile: Picornaviridae. J Gen Virol. 2017;98:2421-2422.
11. Sanfaçon H, Wellink J, Gall OL, Karasev A, Vlugt RVD, Wetzel T. Secoviridae: A proposed family of plant viruses within the order Picornavirales that combines the families Sequiviridae and Comoviridae, the unassigned genera Cheravirus and Sadwavirus and the proposed genus Torradovirus. Arch Virol. 2009;154:899-907.
12. Berman JJ. Group IV Viruses: Single-Stranded (+)Sense RNA. In: Berman JJ, editor. *Taxonomic Guide to Infectious Diseases*. San Diego, CA: Elsevier Science Publishing Co Inc; 2012. p. 237-246.
13. Jia W, Wang F, Li J, Chang X, Yang Y, Yao H et al. A novel iflavirus was discovered in green rice leafhopper Nephotettix cincticeps and its proliferation was inhibited by infection of rice dwarf virus. Front MicroBiol. [Internet]. 2021;11. Available from: https://doi.org/10.3389/fmicb.2020.621141
14. ICTV. Iflaviridae. In: ICTV Report. [Internet]. Available from: https://ictv.global/report/chapter/iflaviridae/iflaviridae
15. van Oers MM. Genomics and Biology of Iflaviruses. In: Asgari S, Johnson KN, editors. *Insect Virology*. UK: Caister Academic Press; 2010. p. 231-250.
16. Geng P, Li W, Lin L, Miranda JRD, Emrich S, An L et al. Genetic characterization of a novel Iflavirus associated with vomiting disease in the Chinese oak silkmoth Antheraea pernyi. PLoS ONE. [Internet]. 2014;9. Available from: https://doi.org/10.1371/journal.pone.0092107
17. Xiao J, Zhu K, Ma R, Zhang C, Lv K, Ge D et al. Full genome sequence of a novel iflavirus from the aster leafhopper Macrosteles fascifrons. Virus Genes. 2022;59:338-342.
18. Ye ZX, Li YH, Chen JP, Zhang CX, Li JM. Complete genome analysis of a novel iflavirus from a leaf beetle, Aulacophora lewisii. Arch Virol. 2020;166:309-312.
19. Jiang H, Coleman J, Miskimins R, Srinivasan R, Miskimins WK. Cap-independent translation through the p27 5'-UTR. Nucleic Acids Res. 2007;35:4767-4778.
20. Davis WG, Bowzard JB, Sharma SD, Wiens ME, Ranjan P, Gangappa S et al. The 3' untranslated regions of influenza genomic sequences are 5'PPP-independent ligands for RIG-I. PLoS ONE. [Internet]. 2012;7. Available from: https://doi.org/10.1371/journal.pone.0032661
21. Chen JH, Zhang RH, Lin SL, Li PF, Lan JJ, Song SS et al. The functional role of the 3 untranslated region and poly(a) tail of duck hepatitis a virus type 1 in viral replication and regulation of ires-mediated translation. Front MicroBiol. [Internet]. 2018;9. Available from: https://doi.org/10.3389/fmicb.2018.02250
22. Kalynych S, Plidal A, Pálková L, Levdansky Y, Miranda JRD, Plevka P. Virion structure of iflavirus slow bee paralysis virus at 2.6-angstrom resolution. J Virol. 2016;90:7444-7455.
23. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX et al. Redefining the invertebrate RNA virosphere. Nature. 2016;540:539-543.
24. Santos ERD, Trentin LB, Ecker A, Silva LA, Borges M, Mowery JD et al. An iflavirus found in stink bugs (Hemiptera: Pentatomidae) of four different species. Virology. 2019;534:72-79.
25. Mordecai GJ, Brettell LE, Pachori P, Villalobos EM, Martin SJ, Jones IM et al. Moku virus; A new Iflavirus found in wasps, honey bees and Varroa. Sci Rep. [Internet]. 2016;6. Available from: https://doi.org/10.1038/srep34983
26. Zhang Y, Ye ZX, Feng XX, Xu ZT, Chen JP, Zhang CX et al. Prevalence of reversed genome organizations for viruses in the family Iflaviridae, order Picornavirales. Microbiol Spectr. [Internet]. 2023;11(3). Available from: https://doi.org/10.1128/spectrum.04738-22
27. Li CX, Shi M, Tian JH, Lin XD, Kang YJ, Chen LJ et al. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. Elife. [Internet]. 2015;4. Available from: https://doi.org/10.7554/elife.05378

28. Lefèvre T, Sauvion N, Almeida RP, Fournet F, Alout H. The ecological significance of arthropod vectors of plant, animal and human pathogens. Trends Parasitol. 2022;38:404-418.
29. Ferraguti M, Magallanes S, Suarez-Rubio M, Bates PJJ, Marzal A, Renner SC. Does land-use and land cover affect vector-borne diseases? A systematic review and meta-analysis. Landscape Ecol. 2023;38:2433-2451.
30. Arlian LG, Morgan MS. A review of sarcoptes scabiei: Past, present and future. Parasites Vectors. [Internet]. 2017;10. Available from: https://doi.org/10.1186/s13071-017-2234-1
31. Engelman D, Kiang K, Chosidow O, McCarthy J, Fuller C, Lammie P et al. Toward the global control of human scabies: Introducing the international alliance for the control of scabies. PLoS Negl Trop Dis. [Internet]. 2013;7. Available from: https://doi.org/10.1371/journal. pntd.0002167
32. Swe PM, Zakrzewski M, Waddell R, Sriprakash KS, Fischer K. High-throughput metagenome analysis of the sarcoptes scabiei internal microbiota and in-situ identification of intestinal Streptomyces sp. Sci Rep. [Internet]. 2019;9. Available from: https://doi.org/10.1038/s41598-019-47892-0
33. Mofiz E, Deborah C, Seemann T, Currie BJ, Fischer K, Papenfuss AT. Genomic resources and draft assemblies of the human and porcine varieties of scabies mites, Sarcoptes scabiei var. Hominis and var. Suis. GigaSci. [Internet]. 2016;5. Available from: https://doi.org/10.1186/s13742-016-0129-2
34. Valles SM, Chen Y, Firth AE, Guérin DMA, Hashimoto Y, Herrero S et al. ICTV virus taxonomy profile: Iflaviridae. J Gen Virol. 2017;98:527-528.
35. Kustin T, Stern A. Biased mutation and selection in RNA viruses. Mol Biol Evolution. 2020;38(2):575-588.
36. Samuel CE. Adenosine deaminases acting on RNA (ADARS) are both antiviral and proviral. Virology. 2011;411(2):180-193.
37. Ward SV, George CX, Welch MJ, Liou LY, Hahm B, Lewicki H et al. RNA editing enzyme adenosine deaminase is a restriction factor for controlling measles virus replication that also is required for embryogenesis. Proc Natl Acad Sci. 2010;108:331-336.
38. Peck KM, Lauring AS. Complexities of viral mutation rates. J Virol. [Internet]. 2018;92(14). Available from: https://doi.org/10.1128/jvi.01031-17
39. Dolja VV, Krupovic M, Koonin EV. Deep roots and splendid boughs of the global plant virome. Annu Rev Phytopathol. 2020;58:23-53.
40. Machitani M, Yasukawa M, Nakashima J, Furuichi Y, Masutomi K. RNA dependent RNA polymerase, RDRP, a promising therapeutic target for cancer and potentially COVID-19. Cancer Sci. 2020;111(11):3976-3984.
41. Sun D, Wang M, Wen X, Mao S, Cheng A, Jia R et al. Biochemical characterization of recombinant avihepatovirus 3C protease and its localization. Virol J. [Internet]. 2019;16. Available from: https://doi.org/10.1186/s12985-019-1155-3
42. Fujimuro M. The interactions between cells and viruses. Int J Mol Sci. [Internet]. 2024;25(13). Available from: https://doi.org/10.3390/ijms25136886
43. Oxford Nanopore Technologies. Oxford Nanopore sequencing: An introduction. Oxford Nanopore Technologies; 2020. https://nanoporetech.com/resource-centre/introduction-nanopore-sequencing
44. Dinan AM, Lukhovitskaya NI, Olendraite I, Firth AE. A case for a negative-strand coding sequence in a group of positive-sense RNA viruses. Virus Evol. [Internet]. 2020;6(1). Available from: https://doi.org/ 10.1093/ve/veaa007
45. Babé LM, Craik CS. Viral proteases: Evolution of diverse structural motifs to optimize function. Cell. 1997;91(4):427-430.
46. Bhat SA, Mounsey KE, Liu X, Walton SF. Host immune responses to the itch mite, Sarcoptes scabiei, in humans. Parasites Vectors. [Internet]. 2017;10. Available from: https://doi.org/10.1186/s13071-017-2320-4
47. Domingo E, García-Crespo C, Lobo-Vega R, Perales C. Mutation rates, mutation frequencies and proofreading-repair activities in RNA virus genetics. Viruses. [Internet]. 2021;13(9). Available from: https://doi.org/10.3390/ v13091882
48. Kyriakopoulou Z, Amoutzias GD, Dimitriou TG, Tsakogiannis D, Mossialos D, Markoulatos P. Intra- and inter-serotypic recombinations in the 5' UTR-VP4 region of echovirus 30 strains. Arch Virol. 2017;163:365-375.