Title

The evolutionary and molecular history of a chikungunya virus outbreak lineage

Authors

Janina Krambrich[1,A]
Filip Mihalič[1]
Michael W. Gaunt[2]
Jon Bohlin[3]
Jenny Hesson[1,4,A]
Åke Lundkvist[1,A]
Xavier de Lamballerie[5]
Cixiu Li[6,7]
Weifeng Shi[7,8]
John H.-O. Pettersson[9,10,11,A*]

Corresponding author

John H.-O. Pettersson (john.pettersson@uu.se | @folkhalsomyndigheten.se)

Author affiliations

1) Department of Medical Biochemistry and Microbiology, Uppsala University, SE-751 23 Uppsala, Sweden;
2) Solena Ag, 300 Lincoln Center Dr. Foster City, CA 94404, USA;
3) Infectious Disease Control and Environmental Health, Norwegian Institute of Public Health, Oslo, Norway;
4) Biologisk Myggkontroll, Nedre Dalälvens Utvecklings AB, Kölnavägen 25, 81197 Gysinge, Sweden;
5) Unité des Virus Émergents (UVE), Aix-Marseille University - IRD 190 - Inserm 1207, Marseille, France;
6) Key Laboratory of Emerging Infectious Diseases in Universities of Shandong, Shandong First Medical University & Shandong Academy of Medical Sciences, Taian 271000, China;
7) Department of Infectious Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China;
8) Shanghai Institute of Virology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China;
9) Department of Medical Science, Uppsala University, SE-751 85 Uppsala, Sweden
10) Department of Clinical Microbiology and Hospital Hygiene, Uppsala University Hospital, SE-751 85 Uppsala, Sweden.
11) Department of Microbiology, Public Health Agency of Sweden, Nobels väg 18, SE-171 82 Solna, Sweden.
A) Members of the Zoonosis Science Center, Uppsala University, Sweden.

Abstract

In 2018–2019, Thailand experienced a nationwide spread of chikungunya virus (CHIKV), with approximately 15,000 confirmed cases of disease reported. Here, we investigated the evolutionary

and molecular history of the East/Central/South African (ECSA) genotype to determine the origins of the 2018–2019 CHIKV outbreak in Thailand. This was done using newly sequenced clinical samples from travellers returning to Sweden from Thailand in late 2018 and early 2019 and previously published genome sequences. Our phylogeographic analysis showed that before the outbreak in Thailand, the Indian Ocean lineage (IOL) found within the ESCA, had evolved and circulated in East Africa, South Asia, and Southeast Asia for about 15 years. In the first half of 2017, an introduction occurred into Thailand from another South Asian country, most likely Bangladesh, which subsequently developed into a large outbreak in Thailand with export to neighbouring countries. Based on comparative phylogenetic analyses of the complete CHIKV genome and protein modelling, we also identified amino acid substitutions that may be associated with immune evasion, increased spread, and virulence. We identified several mutations in the E1/E2 spike complex, such as E1 K211E and E2 V264A, which are highly relevant as they may lead to changes in vector competence, transmission efficiency and pathogenicity of the virus. A number of mutations (E2 G205S, Nsp3 D372E, Nsp2 V793A), that emerged shortly before the outbreak of the virus in Thailand in 2018 may have altered antibody binding and recognition due to their position. This study not only improves our understanding of the factors contributing to the epidemic in Southeast Asia, but also has implications for the development of effective response strategies and the potential development of new vaccines.

Author Summary

We investigated the evolutionary and molecular history of the East/Central/South African (ECSA) genotype to determine the origins of the 2018–2019 chikungunya virus (CHIKV) outbreak in Thailand. We used newly sequenced clinical samples from travellers returning to Sweden from Thailand in late 2018 and early 2019 together with previously published genome sequences. Our phylogeographic analysis shows that the Indian Ocean lineage (IOL), found within ECSA, evolved in Eastern Africa, Southern Asia, and Southeast Asia for about 15 years before the outbreak in Thailand in 2018. We have also identified amino acid substitutions that may be associated with immune evasion, increased spread, and higher virulence that occurred prior to the outbreak and may have played a critical role in the rapid spread of the virus. Our study concludes that monitoring and understanding CHIKV dynamics remains critical for an effective response to the previously unpredictable outbreaks of the virus.

Introduction

Chikungunya virus (CHIKV, *Togaviridae*), is a single-strand positive-sense mosquito-borne RNA virus with a genome of approximately 12 kb that comprises two open reading frames (ORFs) encoding non-structural and structural proteins respectively [1]. The virus is transmitted to humans mainly through the bites of infected mosquitoes, such as *Aedes aegypti* and *Ae. albopictus*, which are widely distributed in tropical and subtropical regions around the world. [2-5]. These mosquito species are also responsible for the transmission of other well-known viruses, for example dengue virus, Zika virus, and yellow fever virus [2]. CHIKV was first discovered in Tanzania in 1952 and has since its discovery both been identified and/or suggested to be the causative agent of multiple outbreaks in Africa and Asia for several decades, if not centuries [6, 7]. However, since 2004 the virus has spread rapidly to new geographic regions and cases are now reported from over 100 countries in Asia, Africa, Europe, and the Americas [8, 9]. The geographic distribution of CHIKV is primarily determined by the presence and spread of its mosquito vectors.

In African forests, a sylvatic cycle of CHIKV occurs between mosquitoes and non-human primates [10, 11]. This sylvatic cycle may lead to sporadic spill-over events, where the virus is transmitted to humans, initiating a separate urban cycle [12]. In the urban cycle, non-human primates are not necessary to sustain the epidemic, since the virus is transmitted between humans and *Aedes* mosquitoes. The sylvatic and the urban cycles can exist separately, contributing to the complex transmission dynamics of CHIKV in African regions.

Following the expansion of CHIKV since 2004, outbreaks have occurred throughout the tropical- and subtropical regions of the world, becoming a significant public health concern. Between 2004 and 2020, 3.4 million suspected and confirmed CHIKV cases were reported from various countries [13]. The actual number of CHIKV infections is likely considerably higher due to underreporting and asymptomatic cases. CHIKV infections, although rarely fatal, can lead to prolonged and incapacitating joint pain, lasting months or even years in some cases [9, 14, 15]. There are several vaccines for CHIKV that are currently under development. The Coalition for Epidemic Preparedness Innovations and the European Commission are currently supporting the development of a live-attenuated, single-dose vaccine that is designed by deleting a part of the CHIKV genome (Ixchiq, VLA1553 by Valneva). In November 2023, the US Food and Drug Administration approved and authorized this vaccine in the US [16, 17]. The Jenner Institute research group has developed another CHIKV vaccine using a combination of recombinant chimpanzee adenoviruses and Modified vaccinia Ankara (MVA), which however is not approved for use yet [18]. Other prevention efforts focus primarily on reducing mosquito populations and avoiding mosquito bites. As of now, there is no specific treatment for CHIKV [19].

Based on phylogenetic analyses, CHIKV is commonly divided into three major lineages: the East/Central/South African (ECSA), the West African, and the Asian lineages [20, 21]. The ECSA lineage gave rise to the Indian Ocean lineage (IOL), which has been responsible for epidemics in the Indian Ocean islands, South and Southeast Asia, and Europe since 2005 [21, 22]. The first CHIKV outbreak in Thailand was reported in Bangkok in 1958, and the Asian genotype was identified as the cause of that outbreak [23]. The next notable outbreak occurred in southern Thailand between 2008 and 2009, followed by a smaller local spread in 2013 in north-eastern Thailand, both caused by the ECSA genotype [24, 25]. The overall number of reported cases remained low until just before the start of the 2018–2019 outbreak, according to the Bureau of Epidemiology in Thailand. In June 2018, the number of monthly reported chikungunya cases in Thailand began to increase and a nationwide spread of CHIKV was observed with approximately 15,000 confirmed cases reported between 2018 and 2019 [26]. The virus primarily affected urban and semi-urban areas, with high transmission rates observed in densely populated regions. Due to international travel and the popularity of Thailand and other tropical regions as tourist destinations, an increase in imported CHIKV cases to other countries, including Europe and the United States, was observed both after the outbreak in Thailand in 2018 and after the CHIKV outbreak in the Caribbean and South America in 2014 [27-32].

To improve our understanding of the factors that contributed to this epidemic in Southeast Asia, and, in particular, the sudden increase of cases in the 2018 Thailand outbreak, we conducted a comprehensive phylogeographic and mutational outbreak lineage analysis focusing on the ECSA and IOL. We performed phylogeny, time estimation and mutational profiling. We also carried out protein folding predictions from genes related to transmission and virulence. The analyses were

performed on previously sequenced genomes as well as newly sequenced clinical samples from travellers who returned to Sweden from Thailand in late 2018 and early 2019.

Methods

Preparation of patient material

Serum samples from a total of 12 patients who had travelled from Thailand to Sweden between December 2018 and April 2019 and who were PCR-positive for CHIKV in a real-time PCR screening were included in the study (supplementary Table S1). Total RNA was extracted from all patient samples by automated magnetic bead total nucleic acid extraction using a MagLEAD system (Precision System Science Co.) from the aqueous phase after Trizol–chloroform separation. RNA-seq libraries were then prepared using the Trio RNA-Seq Library Preparation Kit (NuGen) according to the manufacturer's instructions and subsequently sequenced on one Illumina X10 lane. RNAseq library preparation and high-throughput sequencing were performed by BGI, Hong Kong.

Sequence processing

First, low-quality reads were removed with Trimmomatic v.0.36 using the default settings [33]. All quality-checked sequence data libraries were then mapped with Bowtie v.2.3.4. using the default local settings against a NCBI CHIKV reference sequence (NCBI GenBank accession number: MF773566), whereupon 50% majority consensus sequences were generated. All CHIKV sequences were deposited to the NCBI GenBank under accession numbers PP193832–PP193843 and the raw data (excluding human reads) was deposited under NCBI SRA accession nr: PRJNA1066385.

Evolutionary analyses

First, 2,564 CHIKV sequences of ≥8000bp, with known collection date and geographic location together with the 12 CHIKV genomes generated above were aligned using MAFFT v.7.520 [34, 35] utilizing the L-INS-i algorithm, where the 5' and 3' ends were trimmed. To reduce the number of sequences prior to temporal analyses, we constructed a maximum likelihood phylogenetic tree with IQ-TREE v.2.2.0 [36] using the Generalized time-reversible model of Tavaré 1986 (GTR) with empirical base frequencies, invariant sites, and invariant sites plus FreeRate model with tree categories (GTR+F+I+I+R3) following the ModelFinder implemented in IQ-TREE [37]. We then subsampled the phylogenetic tree to include a total of 218 CHIKV ECSA genotype sequences (see coloured terminal nodes in Supplementary Fig. S1 for sequences included), which were used for the temporal and evolutionary analyses. The temporal structure of the subsampled dataset, sampled between the years 1953 and 2023, was then assessed using TempEst v.1.5.3 (Supplementary Fig. S2) [38]. Finally, the evolutionary history of the subsampled dataset was assessed using BEAST v.1.10.4 [39, 40] by performing a single run of 250 million MCMC generations, sampling every 5k generations, using terminal node calendar dates (i.e. tip dates) as temporal calibration, GTR with invariant sites and four gamma variables with default flat Dirichlet priors as a model of nucleotide evolution, an uncorrelated lognormal relaxed molecular clock with default prior distribution, and a non-parametric Gaussian Markov random field Bayesian Skyride tree prior [41]. Following 10% burn-in, the run was checked using Tracer v.1.7.2 [42] to confirm that the effective sample size for all parameters was >200. Finally, we used TreeAnnotator v.1.10.4 [40] to compute a maximum clade credibility tree and calculate median node heights. The resulting tree was viewed and annotated in FigTree v.1.4.4 [43]. Posterior probabilities ≥ 0.95 are presented in Supplementary Fig. S3).

Mutational analysis

In order to analyse emergent mutations arising in the IOL, we mapped them on the existing structural models of chikungunya protein complexes. This approach was used to visualise mutations in the proteins of the replication and spike complexes respectively. To gain insight into mutations that were located outside of the experimentally resolved regions we additionally predicted the structural models using Colabfold [44]. Colabfold predictions were in excellent agreement with the solved crystal structures in regard to globular domains and also allowed us to visualise the disordered regions found in viral proteins. The replication complex mutations were mapped on the structure accessible under PDBid entry 7y38 (complex of Nsp1, Nsp2 helicase domain and Nsp4) [45] and 4ztb (Nsp2 protease domain) [46]. It should be noted that not all of the amino-acid residues are resolved in these crystal structures. For the E1-E2-E3 trimer spike complex we visualised mutations based on the PDBid entry 6jo8 [47] and 6nk6 [47]. Finally, for the Nsp3, 6K and CP we mapped the mutations directly onto the Colabfold structure prediction due to the high degree of disorder (Nsp3, CP) or lack of other structural information (6K). For both Nsp3 and CP, we superimposed Colabfold predictions onto the relevant solved crystal structures of their respective globular domains to confirm the quality of the predictions and found them to be in excellent agreement (Nsp3 Macrodomain, PDBid: 6vuq, RMSD: 0.344; Nsp3 Zinc-binding domain, PDBid: 4gua, RMSD: 0.58; CP protease domain, PDBid: 5h23, RMSD: 0.505). PyMOL (The PyMOL Molecular Graphics System, Version 2.5.4 Schrödinger, LLC) was used to visualise the protein structures. Different protein sequences were used for the experimentally determined structures meaning that the stick model of the mutation used for figures does not always corresponds to the amino acids involved in the mutation we discuss in text (example: for the Nsp4 structure in Fig. 2D the T75A mutation, a methionine (M) is the actual residue in the Nsp4 sequence that was used for Cryo-electron microscopy experiment, and is thus visualised in the figures). With that in mind the highlighted residues are intended to indicate the position of the mutation rather than imply any amino-acid change or impact of the mutation itself.

Ethical statement

This study was in part conducted at the Public Health Agency of Sweden supported by the ordinance (2021:248:§37) from the Swedish Parliament to study and monitor the situation and development of infectious diseases. It should be noted that, apart from the country of infection and the date of sample collection, no information or data from this project can be linked or traced to a specific individual included in the study. Therefore, the CHIKV-positive samples were used in accordance with the regulations governing the use of such material and in accordance with the mandate of the Swedish Parliament.

Results

Evolutionary history of CHIKV

Based on the phylogenetic analysis, CHIKV can be divided into three different genotypes (Supplementary Fig. S1), the West African, the Asian, and the ECSA genotype, which are supported by high posterior probabilities (Fig. S3). High posterior probabilities (≥ 0.95) were also observed for all nodes, which are described in the results section. The ECSA genotype is characterised by three clusters, as shown in Fig. 1: the paraphyletic group ECSA 1 (light green) and the monophyletic groups ECSA 2 (red) and IOL (blue) (see also Supplementary Fig. S1). When analysing the complete ECSA genotype (consisting of light green, red, and blue, see Supplementary Fig. S2), the root-to-tip analysis showed that our investigated data set exhibited a significant

temporal structure (correlation coefficient = 0.90; R-squared = 0.82, p<0.001, Supplementary Fig. S2). To investigate the dispersal history and temporal structure of the CHIKV ECSA genotype and the IOL in particular, we then performed a phylogeographic analysis of 218 viral genomes of this genotype, covering approximately 70 years of CHIKV evolution and movement between regions (Fig. 1). Fig. 1 shows the resulting time-calibrated phylogenetic tree with maximum clade credibility of CHIKV together with a map indicating the major routes of CHIKV movement for ECSA lineages 1 and 2 (green and red) and the IOL (blue).

For the ECSA genotype as a whole, i.e., including the IOL, the lognormal, uncorrelated relaxed clock estimated the mean evolutionary rate to be 4.4 x 10^{-4} substitutions/site/year (95% highest posterior density [95% HPD] = 4.1–4.8 x 10^{-4}) with a most recent common ancestor emerging around 1952 [95% HPD = 1948–1953] (Fig. 1). For the IOL, the mean evolutionary rate was estimated to be 2.8 x 10^{-4} substitutions/site/year (95% HPD = 2.2–3.5 x 10^{-4}) and a most recent common ancestor emerging around February 2003 (95% HPD 2001-07 to 2004-02). However, we recognise, that there are rate variation between lineages even within the ECSA genotype [48], affecting the African clades, which have a lower sampling frequency. It is noteworthy that the earliest recorded CHIKV isolates are from West and East Africa, emphasising the historical importance of this region in the evolutionary trajectory of the virus.

Early evolution and movement of the ECSA genotype

Following the divergence from a West African ancestor (Supplementary Fig. S1) [48], our phylogeographic analysis of the ECSA genotype reveals a highly structured dispersal network with a significant interregional spread of CHIKV. Since the emergence of the most recent common ancestor of the ESCA genotype, CHIKV has circulated to date, moving between many sub-Saharan African countries, particularly in eastern, central, and southern African countries. The first CHIKV discoveries and isolations were limited to East Africa in 1953 (Tanzania) and West Africa in 1964 (Nigeria) (Fig. 1, Supplementary Fig. S1). Between the 1950s and the 1980s, African ECSA-CHIKV viruses continued to diversify, followed by a long period of sporadic detection but possibly continuous sylvatic circulation [10]. The first of the evolved ECSA lineages, African ECSA 1 (light green), showed intermediate-range dissemination within sub-Saharan Africa, spreading from Central Africa to the Americas (red) in the late 1980s and early 2010s (1995 in the USA and 2014 in Brazil) (Fig. 1). Subsequently, the African ECSA 1 lineage spread eastwards to the Indian Ocean islands (Madagascar, Mauritius, Mayotte, Comoros) in the 1960s and to India in the 1980s (blue) (Fig. 1). A more recent African cluster with representatives from Cameroon, Gabon and the Democratic Republic of Congo, probably emerged more than 20 years ago and circulated in this region until at least 2018. A similar but independent cluster in Angola and the Democratic Republic of Congo, was separated by more than 30 years of evolution before it emerged in 2011. A similar scenario occurred when the pathogen was introduced to Brazil via Western Africa or the USA after circa 30 years of unaccounted evolution, which later led to a significant and sustained outbreak in South America [49]. The Southeast Asian region in particular, including India, emerged as a major inter-regional transmission hub, facilitating the spread of CHIKV to other regions. Sub-Saharan Africa, where the virus originally emerged in the 1920s, also played a central role in the global spread and maintenance of the virus, albeit to a lesser extent than Southeast Asia and India. Bidirectional transmission events between Africa and Asia have been documented, including early transmission from Africa to India in the late 1920s, with subsequent introductions in 1986 and 2000.

Spread and recent emergence of the IOL in Southeast Asia

The emergence and spread of CHIKV in the Indian Ocean islands, the Indian subcontinent, and Southeast Asia are associated with a significant increase in cases. Here we examine the evolutionary history of CHIKV to trace the origins of the outbreak lineage that caused a substantial number of cases in Thailand and other Southeast Asian countries in 2018 and 2019 [24, 26]. The 2004 outbreak of CHIKV in the Indian Ocean islands was the first documented outbreak in the IOL [50]. The outbreak was mainly observed in urban and semi-urban areas, for example on the Comoros Islands where more than 5,000 cases were reported [51]. Seroprevalence studies from 2011 indicate that 20% of the population on Ngazidja (Grande Comore), the largest island in the Comoros with a population of approximately 316,600, were infected with CHIKV on [52]. The outbreak then spread to other islands in the Indian Ocean, including Madagascar, Mauritius, and the Seychelles, and eventually to other parts of the world, including Europe [53, 54]. The IOL has been circulating in South and Southeast Asia for two decades now, with several sub-lineages and variants having emerged and spread throughout the region (Fig. 1). The last common ancestor of the IOL of the ECSA genotype is estimated to have originated in coastal Kenya and the Mascarene islands, around early 2003 (95% HPD: 2001-07–2004-02), which is consistent with previous estimates [50, 55]. A new IOL sub-lineage, distinct from the previous IOL that originated from the Kenyan coast, was found to have originated in India and circulated during 2008–2016, with subsequent spread to Pakistan, Bangladesh, Thailand, and Italy [56]. Several introductions and re-introductions of IOL strains to Africa (Kenya, Djibouti, and Sudan) and the Arabian Peninsula have been observed over the years. Most outbreaks in South Asian countries since 2005 have reportedly been caused by IOL strains, and new clades have evolved in multiple Southeast Asian countries over time, indicating a significant presence of the IOL in the region.

Emergence of mutations in the IOL preceding the Thai outbreak in 2018

The CHIKV sequences isolated in this study from Swedish travellers, as well as other sequences isolated during the 2018–2019 outbreak in Thailand, belonged to the IOL but show marked differences from the strains responsible for the massive CHIKV outbreak in Thailand in 2008–2009, indicating a clear, non-local origin. The outbreak was likely due to the introduction of a viral strain from South Asia, possibly Bangladesh (Fig. 1) [56], since phylogenetic analysis of the isolates revealed that the Thai sequences diverged from a Bangladeshi ancestor around April 2017 (95% HPD: January–May 2017). Following its introduction and epidemic spread in Thailand, CHIKV also spread to Cambodia, Malaysia, Myanmar and China starting in mid-2018 (Fig. 1). Our analysis of the IOL outbreak strains led to the detection of mutations in the ancestral strain, distinguishing the outbreak strains from ESCA 1 and 2 at the last common ancester node (Fig. 1, Node A). Our results revealed a high frequency of amino acid substitutions in both structural and non-structural genes of CHIKV. Nine substitutions were detected in non-structural proteins and nine in structural proteins. Over the subsequent history of the IOL, 18 additional amino acid substitutions were introduced on 14 occasions in the main lineage (green boxes in Fig. 1), while 16 different substitutions were introduced on 12 occasions in different subclades (orange boxes in Fig. 1).

Genetic diversity in the IOL

Throughout the CHIKV IOL divergence, we identified 47 mutations that occurred at different time points in viral evolution. To rationalise the impact of the mutations on viral proteins function of, we mapped the emergent mutations that occurred in the IOL to solved crystal structures or Colabfold-

predicted models. We analysed the replication complex and the trimeric E1-E2-E3 spike separately, as they consist of highly interactive complexes and high-quality structural models are available for these complexes. For the Nsp3, CP, and 6K proteins, which had a high number of mutations in the experimentally unsolved regions, we mapped the mutations to the Colabfold-predicted models and confirmed minimal deviation from the experimentally determined structures of their globular domains to ensure a high quality of prediction. Not surprisingly, we found that most mutations occurred in the surface-exposed regions of the viral proteins and that the mutations were predominantly conservative in nature, with a few notable exceptions (Fig. 2–4). We found most mutations in the E1-E2-E3 spike complex, where mutations were evenly distributed across all three proteins, and the fewest mutations in protein 6K, where only one emergent mutation was observed. An overview of all identified mutations is shown in Table 1.

Table 1: Amino acid substitutions in the IOL.

Protein	Aminoacid			Conservation	Location	Effect	Node	Published functional studies
	Pos.	Ori.	Sub.					
Nsp1	128	T	K	polar to charged	surface exposed	minimal	G	
	290	I	V	conservative	hydrophobic core	may effect protein stability	AA	
	376	T	M	polar to hydrophobic	ambigous	may effect docking of Nsp4, into Nsp1 ring	C	
	488	Q	R	polar to charged	surface exposed	minimal	A	
Nsp2	54	S	N	conservative	Nsp4 interface	may impact Nsp2, Nsp4 interaction	A	
	130	H	Y	semi-conservative	surface exposed	minimal	N	
	145	E	D	conservative	surface exposed	minimal	U	
	495	N	S	conservative	surface exposed	may effect substrate recognition	Z	
	539	L	S	hydrophobic to polar	surface exposed	may alter protein stability or protein protein interactions, may effect substrate recognition	H	
	566	S	F	polar to aromatic	surface exposed	may alter protein stability or protein protein interactions, may effect substrate recognition	Y	
	793	A	V	semi-conservative	surface exposed	likely unfavorable	X	
Nsp3	59	M	T	hydrophobic to polar	surface exposed	may have stabilizing effects	D	
	217	Y	H	semi-conservative	surface exposed	may effect RNA replication and essembly	AA	
	337	T	I	polar to hydrophobic	surface exposed	may be new interaction sites, or minimal effect	A	
	338	T	M	polar to hydrophobic	surface exposed	may be new interaction sites, or minimal effect	F	
	372	D	E	conservative	surface exposed	may be new interaction sites, or minimal effect	W	
	461	L	P	hydrophobic to cyclic	surface exposed	may be new interaction sites, or minimal effect	A	
	471	P	S	cyclic to polar	surface exposed	may be new interaction sites, or minimal effect	A	
Nsp4	55	S	N	conservative	surface exposed	minimal	T	
	75	T	A	semi-conservative	surface exposed	minimal	A	
	82	R	S	charged to polar	surface exposed	minimal	I	
	85	R	G	charged to small flexible	surface exposed	may affect Nsp1-Nsp4 interface, may impact dimer formation	O	
	254	T	A	semi-conservative	surface exposed	minimal	A	
E1	55	I	V	conservative	hydrophobic core	may interfere with folding	M	
	136	L	F	semi-conservative	surface exposed	minimal	S	
	211	K	E	polar to charged	surface exposed	effects adaptability to *Ae. aegypti*	L	[68-73]
	226	A	V	semi-conservative	Mxra8 interface	increases fitness in *Ae. albopictus*	B, E	[71, 74-76]
	269	M	V	conservative	surface exposed	minimal	A	
	284	D	E	conservative	surface exposed	minimal	A	
	317	I	V	conservative	surface exposed	minimal	R	
E2	74	M	I	conservative	Mxra8 interface	minimal	S	
	76	A	T	semi-conservative	Mxra8 interface	minimal	P	
	205	G	S	small flexible to polar	surface exposed	may effect immune evasion	V	
	210	L	Q	hydrophobic to polar	surface exposed	may effect immune evasion	D	
	211	I	T	hydrophobic to polar	surface exposed	may effect immune evasion	A	
	252	K	Q	charged to polar	E3 interface	may affect E2, E3 interface	H	
	264	V	A	semi-conservative	Mxra8 interface	may affect MXRA8, spike interaction	K	[71, 95-99]
	312	T	M	polar to hydrophobic	surface exposed	minimal	A	
	375	S	T	conservative	surface exposed	minimal	A	
	386	V	A	semi-conservative	surface exposed	minimal	A	
E3	39	V	I	conservative	E2 interface	may affect E2, E3 interface	J	
	56	P	S	cyclic to polar	surface exposed	minimal	J	
6K	8	V	I	conservative	surface exposed	minimal	A	
CP	23	P	S	cyclic to polar	surface exposed	minimal	G	
	27	V	I	conservative	surface exposed	minimal	G	
	73	K	R	conservative	surface exposed	minimal	Z	
	79	N	S	conservative	surface exposed	minimal	Q	

Replication complex

The replication complex consists of the proteins Nsp1, Nsp2, and Nsp4, which form a disc-like structure that docks into the neck of the ultrastructures packed with viral RNA, the so-called spherules (Fig. 2A) [45, 57]. In this complex, eleven monomers of the RNA capping enzyme Nsp1

form an outer ring to which the RNA-dependent RNA polymerase Nsp4 is docked. The viral protease Nsp2 also associates with the complex from the cytoplasmic side (Fig. 2A) [45]. In Nsp1, two mutations, T128K and Q488R, are located on the protein surface, which probably have only minimal effects on the stability and function of the protein (Fig. 2B). The conservative mutation I290V was found to be buried in the hydrophobic core of the protein and may have slight effects on protein stability, while the T376M mutation is located near the Nsp1-Nsp4 interface and thus may affect the docking of Nsp4 to the oligomeric ring of Nsp1.

Of the seven mutations found in Nsp2, three were located in the N-terminal helicase domain (S54N, H130Y and E145D) and four in the C-terminal protease domain (N495S, L539S, S566F and A793V; Fig. 2C). While H130Y and E145D are conservative surface mutations that likely have limited effects on Nsp2 function, S54N is located at the interface between Nsp2 and Nsp4 so the mutation could potentially have an impact on the interaction between Nsp2 and Nsp4 and on the overall stability of the complex. Of the four mutations found in the protease domain, N495S retains a hydrophilic character and is unlikely to affect the function of the protein. The non-conservative surface mutations L539S and S566F could alter the stability of the protein or interactions with potential binding partners, and A793V introduces a larger hydrophobic moiety on the surface of the short, disordered C-terminal peptide of the Nsp2 protease domain, which is unlikely to be favourable. Being in moderate proximity to the Nsp2 active site, it is possible that the N495S, L539S and S566F mutations affect substrate recognition, as the exact substrate binding interface for CHIKV Nsp2 is not clear. Interestingly, the A793V mutation reverted to A in the last common ancestor of the 2018 Thai outbreak lineage.

The Nsp4 is largely devoid of emergent mutations, with the exception of the N-terminal domain, which extends into the replication spheroid space and for which an interaction with the RNA template has been proposed [45] (Fig. 2D). All of the mutations we identified in Nsp4 (S55N, T75A, R82S, R85G and T254A) are surface mutations, with R82S and R85G being the most likely to affect Nsp4 protein function due to their non-conservative nature. The R85G mutation in particular is located close to the Nsp1-Nsp4 interface and could impair the respective dimer formation.

Spike complex

The spike complex is a trimer of E1-E2-E3 heterotrimers, that forms the icosahedral outer envelope (Fig. 3A) of virus particles and is responsible for receptor binding, membrane fusion and viral entry [58-62]. While E1 performs membrane fusion in acidic environments [63], E2 facilitates Matrix remodelling-associated protein 8 (MXRA8) receptor binding [64], and E3 protects premature exposure of the E1 fusion loop and is important for correct E1-E2 maturation [64, 65]. We identified seven, ten and two emergent mutations in the E1, E2, and E3 proteins, respectively. All mutations found in E1 are surface mutations with the exception of I55V, which is located in the hydrophobic core of domain II and could interfere with the correct folding of the protein (Fig. 3B). Of the remaining six mutations, three are located in the domain I and domain III regions of the protein and are located both proximal to the membrane and at the base of the spike. Mutation L136F is located on the surface of domain I, I317V is located on the surface of domain III, and mutation D284E is situated in the linker region between domains I and III. This junction region undergoes substantial rearrangement upon conversion of E1 to the fusion form [66, 67], but given the conservative nature of the mutation, the rearrangement is unlikely to be affected by the

identified mutation. As the surface mutations found in domains I and III are positioned at the outer surface of the spike, they could also affect the packing of neighbouring spike complexes into the icosahedral viral lattice. Of the remaining three mutations, K211E has already been described to have an impact on viral replication and the adaptability to the *Ae. aegypti* vector [68-73]. The conservative mutation M269V is located on the inner spike surface of E1 and probably has no effect on protein function. The A226V mutation of E1, which is associated with increased fitness of CHIKV in *Ae. albopictus* [71, 74-76], and probably contributes to the epidemic potential of CHIKV, was introduced in two different subclades at different time points (Fig. 1, Node B and E). However, in the subclade following node E, two sequences, FJ000067.1 from India and FJ445428.2 from Sri Lanka, do not have valine at position 226 but alanine. In the main IOL, the ancestral alanine remained at position 226.

MXRA8 is a cell surface receptor for several arthritogenic alphaviruses, such as CHIKV [62, 77], and interacts mainly with the outer crown of the spike complex consisting of three E2 proteins. Although the E2 mutations M74I and A76T are located proximal to the MXRA8 interface, they are unlikely to have a significant impact on receptor interaction given their conservative nature. In contrast, V264A is in direct contact with the MXRA8, so this mutation probably has a greater impact on the interaction between the spike complex and the cell surface receptor. Because it is exposed on the surface of the virion, E2 is also the primary target of natural and recombinant antibodies [78-80]. Several of the antibodies target the B domain of E2 [78, 79] in the region where we also found three emergent mutations: G205S, L210Q, and I211T (Fig. 3C). Since the same residues are targeted by antibodies, their mutation could serve as an immune evasion mechanism leading to enhanced viral fitness [78, 79, 81]. Interestingly, the K252Q mutation is located at the interface of E2 and E3 and is in direct contact with the V39I mutation on the E3 protein. Since these mutations do not occur in the same viral clade, they are likely not the result of coevolution, but suggest that some degree of amino acid variation is operative in this region of E1–E3 (Fig. 3C). Finally, T312M is a surface mutation with likely limited effects on protein function and S375T, V386A are located in the transmembrane region of the E2 protein with likely minimal effects on the protein function.

Apart from the V39I mutation, only one other mutation was found in the E3 protein (Fig. 3D). The P56S mutation, which is located in the immediate vicinity of the C-terminal furin cleavage site [64], probably has no major influence on protein function. Both E3 mutations occur only in one clade of CHIKV with node J as the closest common ancestor (Fig. 1).

Nsp3, 6K, and capsid proteins

For the proteins that are neither part of the replication nor the spike complex, we have mapped the mutations to the individual structural models predicted by Colabfold. The Nsp3 protein, which is closely associated with the replication complex [45, 82], consists of two folded globular domains, the N-terminal macro-domain and the zinc-binding domain, followed by an elongated hypervariable C-terminal disordered region (Fig. 4A) [82]. The macro-domain exhibits ADP-ribosylhydrolase activity [83], while the zinc-binding domain, although poorly understood, is associated with various functions in viral genome replication and transcription that are often species- and cell type-specific [84]. The hypervariable C-terminal region has been shown to be intrinsically disordered [85] and serves as a platform for the binding of various host factors [85, 86]. We identified one emergent mutation in each of the two folded domains and five mutations in the hypervariable C-terminal disordered region (Fig. 4A). The M59T mutation, located on the surface of the macro-domain, is on

the opposite side of the active site and might have a stabilising effect due to the transition from a hydrophobic to a hydrophilic surface-exposed amino acid. The Y217H mutation is located at the base of a small, shallow pocket on the surface of the zinc-binding domain, which could be a binding pocket. Interestingly, a reversal to an ancestral Y can be observed at this position in 2019 (Node AA, Fig. 1). As the function of this domain is unclear, it is difficult to speculate on the effects of this specific mutation. However, it has been shown that the entire domain is crucial for RNA replication and viral assembly [87]. None of the five mutations found in the C-terminal domain interfere with the previously described short linear motifs found in this region, which interact with amphiphysin-SH3 (which is recruited by the virus to promote viral RNA replication) [88] and the G3BP-NTF2 domains (which are hijacked to block stress granule formation) [89]. Given the high density of short linear motifs in the disordered regions of viral proteins and the fact that the C-terminal disorder domain of Nsp3 acts as an interaction hub for host factors, it is possible that the T337I, T338M, D372E, L461P, and P471S mutations create new binding sites for host proteins or destroy existing ones [90]. Alternatively, these mutations might not have significant effects on the function of the protein and might be consequences of random drift.

The 6K protein is a poorly understood, highly hydrophobic protein that forms hexameric ion channels in the endoplasmic reticulum (ER) membrane [91, 92]. We have identified only one conservative mutation, V8I, which likely has limited or no effect on protein function (Fig. 4B).
Finally, the capsid protein (CP) is a multifunctional protein with an N-terminal, positively charged, intrinsically disordered region involved in RNA encapsidation [93], and a C-terminal, chymotrypsin-like protease domain that binds to the transmembrane helix of E2 and forms the inner lattice of the mature viral nucleocapsid [65, 94]. All four mutations found in the CP protein, P23S, V27I, K73R, and N79S, are located in the N-terminal disordered region (Fig. 4C). The mainly conservative nature makes it unlikely that these mutations have a major impact on the function of CP. This, and the fact that we found no emergent mutations in the C-terminal protease domain, suggests that the CP protein is under tight evolutionary constraints that allows only very limited variation in the amino acid sequence.

We speculated about the effect of the mutations based on their position in the structure, but further experimental validations should be performed to accurately determine the effect of individual mutations. This opens an interesting avenue for future work.

Discussion

CHIKV has left an indelible mark on the global landscape of infectious diseases, and its emergence and spread over time provide valuable insights into the complexity of vector-borne diseases. The current diversity of CHIKV is thought to have originated in sub-Saharan Africa in the 1920s [6]. This is consistent with our analysis, where the deepest split in our CHIKV tree is estimated to be in the 1950s. The emergence of the Asian genotype shortly thereafter in the 1930s marked the beginning of a series of events that eventually led to the formation of distinct lineages. These genotypes, including the West African, the Asian, and the ECSA genotype with the IOL, illustrate the intricate evolutionary history of CHIKV. Each genotype has played a unique role in the global spread of the virus.

Our analysis shows that there is considerable interregional transmission of CHIKV. The Southeast Asian region, particularly India, stands out as an important interregional transmission site linking CHIKV isolates from other regions. Sub-Saharan Africa, where CHIKV first emerged, also plays a central role in the global spread, albeit to a lesser extent than Southeast Asia. South America appears to be the primary source of intra-continental spread, rather than a source of transmission of CHIKV to other continents. This long-distance transmission emphasises the adaptability of CHIKV to different ecological and environmental conditions, as it can be transmitted in both urban and sylvatic cycles [100]. This is evidence of the resilience and adaptability of CHIKV as it navigates different regions and ecosystems. The possibility of transmission in both urban and sylvatic environments gives CHIKV the opportunity to spread through infected human travellers and cause new outbreaks, but also to circulate locally and establish an endemic occurrence of the virus.

The IOL represents an intriguing aspect of the global spread of CHIKV. It shows a multitude of bidirectional transmission events linking Southeast Asia, India, East Africa, the Arabian Peninsula, and Europe. This lineage emphasises the intricate network of CHIKV transmission in the Indian Ocean region and highlights the role of different regions in maintaining the presence of the virus.
Phylogenetic analysis of CHIKV Thai strains isolated between 2018 and 2020 during the large outbreak in Thailand revealed that they are mapped within the IOL to the ECSA genotype, the same genotype responsible for the massive Thailand outbreak in 2008–2009. The strains from the 2008–2009 outbreak however, possess the E1 A226V mutation, which is associated with enhanced transmission by *Ae. albopictus*, compared to strains circulating before 2008 [71, 75, 76, 96, 101-103]. A 2021 study by Khongwichit et. al. found that none of the ECSA strains isolated during the second massive outbreak in Thailand from late 2018 to early 2020 carried this E1 A226V mutation [24], nor did we find it in our Thai isolates (Fig. 1). Instead, the new Thai strains had the ancestral alanine at position 226 of the E1 envelope glycoprotein, showing similarities to previous outbreaks in Thailand in 1958 [23]. This leads us to hypothesise that there must have been other factors in the 2018 outbreak that lead to the rapid spread of the virus.

We and others found that the 2018–2020 Thai strains had additional mutations of interest, such as E1 K211E (Fig. 1, Node L, Fig. 3B) and E2 V264A (Fig. 1, Node K, Fig. 3C) [24]. It has been reported that positive selection had a dramatic effect on the alteration of the amino acid residue from lysine (K) to glutamic acid (E) at position 221 of the E1 protein and that mutations on the E1 and E2 envelope glycoproteins in general can affect the vector competence, transmission efficiency, and pathogenicity of the virus [71, 95-99]. The V264A substitution is located at the MXRA8 receptor-binding interface and the mutation could alter the interaction between the viral spike complex and the cell surface receptor. The E1 K211E mutation has been associated with enhanced viral infection in *Ae. aegypti* and has also been reported in other regions [104-108]. This adaptation to a different vector may have influenced the increased spread of CHIKV in Thailand in 2018–2019. Consistently, all sequences isolated and sequenced in this study contain alanine at position 226 of the E1 protein and carry the mutations E1 K211E and E2 V264A.

Two other notable mutations in the structural protein E2 are I211T (Fig. 1, Node A, Fig. 3C) and G205S (Fig. 1, Node V, Fig. 3C). The I211T mutation occurs at the IOL ancestral node while the G205S substitution occurs at node V, probably in early 2016, shortly before the progenitor of the 2018 Thai outbreak began to circulate in Bangladesh. Both mutations are located in the region that has been described as critical for antibody binding and recognition [78, 79]. Mutations at these

positions could therefore lead to evasion of the immune system, increased spread, and higher virulence in the population.

Another substitution that occurred later, around the end of 2017, in Bangladesh is the Nsp2 V793A reverse mutation (Fig. 1, Node X, Fig. 2C). As this mutation is located at the very end of the C-terminal disordered tail of Nsp2 (as predicted by Colabfold), which was not resolved in the crystallisation studies [46], we propose that the introduction of the hydrophobic moiety interferes with the optimal function of Nsp2, leading to the observed V793A back mutation. Importantly, the alanine at position 793 is also present in the rapidly expanding ECSA 2 lineage in South America, suggesting that it may have a beneficial effect on viral transmission (Fig. 1).

The comparison of the phylogenetic relationship of the CHIKV sequences from the 2018–2020 Thai outbreaks with other global sequences showed that the most recent outbreak in Thailand did not originate from the strain circulating in the country. It also belongs to the IOL, but probably originated from other countries in South Asia, most likely from Bangladesh via Myanmar in late 2017 or early 2018. The timing suggests a gradual overland introduction into Thailand from Bangladesh via Myanmar, e.g. through travelling and resettlement of people with subsequent spread within Thailand and spill over to China and Cambodia in mid-2019 and 2020. Due to the political situation in Myanmar, no information could be obtained on the number of positive CHIKV cases in the years between 2016 and 2018.

The emergence of CHIKV outbreaks in certain regions, such as the outbreak in Thailand in 2018, is an example of the ability of the virus to re-emerge and spread rapidly. This emphasises the importance of monitoring and understanding the dynamics of CHIKV transmission in order to take effective public health measures. It is not yet fully understood whether the re-emergence of the virus is caused by purely urban cycles with occasional re-introduction from other countries or whether there also is a sylvatic component. CHIKV could circulate in a sylvatic cycle of non-human primates and mosquitoes and remain undetected in the wild until a spill over event into the urban human mosquito cycle occurs, which in some cases could cause new local outbreaks [10, 11]. Both scenarios are possible and plausible. In some cases, very low genetic variation can be detected in strains occurring in the same geographical area years apart (Fig. 1, observed in Brazil in ECSA 2 or after Node M in India). In these cases, undetected sylvatic transmission could be suggested as a silent reservoir. However, in cases such as the 2018 Thai outbreak, introduction from an urban cycle in a neighbouring country is more likely considering the genetic relatedness and timing of virus spread.

A global increase in CHIKV circulation was detected in 2023. By 30 November, more than 460 000 cases had been reported accompanied by 360 deaths [109]. South America was particularly affected. Argentina and Uruguay reported local transmission for the first time in 2023 [110]. Contributing factors include climate change, which leads to changes in vector activity and distribution, and increased human travel, which plays an important role in the spread of CHIKV in South America and globally [111-115]. Unusual temperature spikes, prolonged warm spells, and altered rainfall patterns combined with increased humidity have created conditions that favour the survival and proliferation of Ae. aegypti and Ae. albopictus mosquitoes in regions where they were previously absent [116-121]. The emergence of CHIKV in the Caribbean islands, a favourite destination for tourists from North America and Europe, creates additional new opportunities for

intercontinental transmission of the infection [115]. Unanticipated and rapid urbanization further promotes the spread of the virus, as Ae. aegypti and Ae. albopictus are particularly attracted to urban areas and warm environments. These mosquitoes utilise water-containers in or near households, such as plant pots and vases, for breeding, which further increases the rate of transmission [122, 123].

In summary, CHIKV has a rich evolutionary history, originating in sub-Saharan Africa and spreading worldwide through complex genotypes, lineages, and transmission centres. A detailed analysis of Thai strains from 2018 to 2020 shows that unique mutations associated with virus replication, receptor binding and transmission occur throughout the genome, suggesting alternative factors for the rapid spread in the 2018 outbreak. The ability of the virus to re-emerge and spread rapidly, combined with climate change and urbanisation, poses an ongoing public health challenge. Monitoring and understanding CHIKV dynamics remain critical to an effective response to the unpredictable outbreaks of the virus.

Acknowledgements

JHOP is funded by the Swedish Research Council VR (grant no.: 2020-02593). This study was partially funded by the Academic Promotion Programme of Shandong First Medical University (grant no.: 2019QL006). JK was funded by the European Union's Horizon 2020 research innovation program (grant no.: 874735 (VEO)), and the SciLifeLab Pandemic Preparedness projects (grant no.: LPP1-007 and REPLP1:005).

References

1. Strauss, J. H.; Strauss, E. G., The alphaviruses: gene expression, replication, and evolution. *Microbiological reviews* **1994,** 58, (3), 491-562.
2. Cook, G. C.; Zumla, A.; Manson, P. S., *Manson's tropical diseases*. 22nd;22; ed.; Saunders/Elsevier: Edinburgh, 2009.
3. Robinson, M.; Conan, A.; Duong, V.; Ly, S.; Ngan, C.; Buchy, P.; Tarantola, A.; Rodó, X., A model for a chikungunya outbreak in a rural Cambodian setting: implications for disease control in uninfected areas. *PLoS Negl Trop Dis* **2014,** 8, (9), e3120.
4. Yakob, L.; Clements, A. C., A mathematical model of chikungunya dynamics and control: the major epidemic on Réunion Island. *PLoS One* **2013,** 8, (3), e57448.
5. Powers, A. M.; Logue, C. H., Changing patterns of chikungunya virus: re-emergence of a zoonotic arbovirus. *J Gen Virol* **2007,** 88, (Pt 9), 2363-2377.
6. Staples, J. E.; Breiman, R. F.; Powers, A. M., Chikungunya fever: an epidemiological review of a re-emerging infectious disease. *Clin Infect Dis* **2009,** 49, (6), 942-8.
7. Gould, E.; Pettersson, J.; Higgs, S.; Charrel, R.; de Lamballerie, X., Emerging arboviruses: Why today? *One Health* **2017,** 4, 1-13.
8. CDC Chikungunya Virus.
9. Vairo, F.; Haider, N.; Kock, R.; Ntoumi, F.; Ippolito, G.; Zumla, A., Chikungunya: Epidemiology, Pathogenesis, Clinical Features, Management, and Prevention. *Infect Dis Clin North Am* **2019,** 33, (4), 1003-1025.
10. Althouse, B. M.; Guerbois, M.; Cummings, D. A. T.; Diop, O. M.; Faye, O.; Faye, A.; Diallo, D.; Sadio, B. D.; Sow, A.; Faye, O.; Sall, A. A.; Diallo, M.; Benefit, B.; Simons, E.; Watts, D. M.; Weaver, S. C.; Hanley, K. A., Role of monkeys in the sylvatic cycle of chikungunya virus in Senegal. *Nat Commun* **2018,** 9, (1), 1046.
11. Thiboutot, M. M.; Kannan, S.; Kawalekar, O. U.; Shedlock, D. J.; Khan, A. S.; Sarangan, G.; Srikanth, P.; Weiner, D. B.; Muthumani, K., Chikungunya: A Potentially Emerging Epidemic? *PLOS Neglected Tropical Diseases* **2010,** 4, (4), e623.

12. Tsetsarkin, K. A.; Chen, R.; Weaver, S. C., Interspecies transmission and chikungunya virus emergence. *Curr Opin Virol* **2016,** 16, 143-150.
13. Powers, A. M., Vaccine and Therapeutic Options To Control Chikungunya Virus. *Clin Microbiol Rev* **2018,** 31, (1), 10.1128/cmr.00104-16.
14. Powers, A.; Huang, H.; Roehrig, J.; Strauss, E.; Weaver, S., Family - Togaviridae. In *Virus Taxonomy - Ninth Report of the International Committee on Taxonomy of Viruses*, King, A. M. Q.; Adams, M. J.; Carstens, E. B.; Lefkowitz, E. J., Eds. Elsevier: San Diego, 2012; pp 1103-1110.
15. Rodríguez-Morales, A. J.; Cardona-Ospina, J. A.; Fernanda Urbano-Garzón, S.; Sebastian Hurtado-Zapata, J., Prevalence of Post-Chikungunya Infection Chronic Inflammatory Arthritis: A Systematic Review and Meta-Analysis. *Arthritis Care Res (Hoboken)* **2016,** 68, (12), 1849-1858.
16. Schneider, M.; Narciso-Abraham, M.; Hadl, S.; McMahon, R.; Toepfer, S.; Fuchs, U.; Hochreiter, R.; Bitzer, A.; Kosulin, K.; Larcher-Senn, J.; Mader, R.; Dubischar, K.; Zoihsl, O.; Jaramillo, J.-C.; Eder-Lingelbach, S.; Buerger, V.; Wressnigg, N., Safety and immunogenicity of a single-shot live-attenuated chikungunya vaccine: a double-blind, multicentre, randomised, placebo-controlled, phase 3 trial. *The Lancet* **2023,** 401, (10394), 2138-2147.
17. Valneva SE, Valneva Announces U.S. FDA Approval of World's First Chikungunya Vaccine, IXCHIQ®. In SE, V., Ed. Saint-Herblain (France), 2023.
18. Campos, R. K.; Preciado-Llanes, L.; Azar, S. R.; Kim, Y. C.; Brandon, O.; López-Camacho, C.; Reyes-Sandoval, A.; Rossi, S. L., Adenoviral-Vectored Mayaro and Chikungunya Virus Vaccine Candidates Afford Partial Cross-Protection From Lethal Challenge in A129 Mouse Model. *Frontiers in Immunology* **2020,** 11.
19. de Lima Cavalcanti, T. Y. V.; Pereira, M. R.; de Paula, S. O.; Franca, R. F. d. O., A Review on Chikungunya Virus Epidemiology, Pathogenesis and Current Vaccine Development. *Viruses* **2022,** 14, (5), 969.
20. Volk, S. M.; Chen, R.; Tsetsarkin, K. A.; Adams, A. P.; Garcia, T. I.; Sall, A. A.; Nasar, F.; Schuh, A. J.; Holmes, E. C.; Higgs, S.; Maharaj, P. D.; Brault, A. C.; Weaver, S. C., Genome-Scale Phylogenetic Analyses of Chikungunya Virus Reveal Independent Emergences of Recent Epidemics and Various Evolutionary Rates. *Journal of Virology* **2010,** 84, (13), 6497-6504.
21. Weaver, S. C.; Lecuit, M., Chikungunya Virus and the Global Spread of a Mosquito-Borne Disease. *New England Journal of Medicine* **2015,** 372, (13), 1231-1239.
22. Suhrbier, A., Rheumatic manifestations of chikungunya: emerging concepts and interventions. *Nat Rev Rheumatol* **2019,** 15, (10), 597-611.
23. Hammon, W. M.; Rundnick, A.; Sather, G. E., Viruses Associated with Epidemic Hemorrhagic Fevers of the Philippines and Thailand. *Science* **1960,** 131, (3407), 1102-1103.
24. Khongwichit, S.; Chansaenroj, J.; Thongmee, T.; Benjamanukul, S.; Wanlapakorn, N.; Chirathaworn, C.; Poovorawan, Y., Large-scale outbreak of Chikungunya virus infection in Thailand, 2018–2019. *PLOS ONE* **2021,** 16, (3), e0247314.
25. Wanlapakorn, N.; Thongmee, T.; Linsuwanon, P.; Chattakul, P.; Vongpunsawad, S.; Payungporn, S.; Poovorawan, Y., Chikungunya outbreak in Bueng Kan Province, Thailand, 2013. *Emerg Infect Dis* **2014,** 20, (8), 1404-6.
26. Division of Epidemiology, D. o. D. C., Ministry of Public Health Thailand, Annual Epidemiological Surveillance Report. In Department of Disease Control, T., Ed. Division of Epidemiology, Department of Disease Control, Ministry of Public Health Thailand: 2008 to 2022.
27. Lindsey, N. P.; Prince, H. E.; Kosoy, O.; Laven, J.; Messenger, S.; Staples, J. E.; Fischer, M., Chikungunya virus infections among travelers-United States, 2010-2013. *Am J Trop Med Hyg* **2015,** 92, (1), 82-7.

28. Gibney, K. B.; Fischer, M.; Prince, H. E.; Kramer, L. D.; St George, K.; Kosoy, O. L.; Laven, J. J.; Staples, J. E., Chikungunya fever in the United States: a fifteen year review of cases. *Clin Infect Dis* **2011,** 52, (5), e121-6.
29. Lindsey, N. P.; Staples, J. E.; Fischer, M., Chikungunya Virus Disease among Travelers-United States, 2014-2016. *Am J Trop Med Hyg* **2018,** 98, (1), 192-197.
30. Adams, L. E.; Martin, S. W.; Lindsey, N. P.; Lehman, J. A.; Rivera, A.; Kolsin, J.; Landry, K.; Staples, J. E.; Sharp, T. M.; Paz-Bailey, G.; Fischer, M., Epidemiology of Dengue, Chikungunya, and Zika Virus Disease in U.S. States and Territories, 2017. *Am J Trop Med Hyg* **2019,** 101, (4), 884-890.
31. Centers for Disease Control and Prevention, N. C. f. E. a. Z. I. D. N., Division of Vector-Borne Diseases (DVBD) Chikungunya in the US. https://www.cdc.gov/chikungunya/geo/chikungunya-in-the-us.html (May 25),
32. European Centre for Disease Prevention and Control *Annual epidemiological report. Chikungunya fever*; ECDC: Stockholm, 2008 to 2021.
33. Bolger, A. M.; Lohse, M.; Usadel, B., Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014,** 30, (15), 2114-2120.
34. Katoh, K.; Misawa, K.; Kuma, K. i.; Miyata, T., MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **2002,** 30, (14), 3059-3066.
35. Rozewicki, J.; Li, S.; Amada, K. M.; Standley, D. M.; Katoh, K., MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Research* **2019,** 47, (W1), W5-W10.
36. Minh, B. Q.; Schmidt, H. A.; Chernomor, O.; Schrempf, D.; Woodhams, M. D.; von Haeseler, A.; Lanfear, R., IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **2020,** 37, (5), 1530-1534.
37. Kalyaanamoorthy, S.; Minh, B. Q.; Wong, T. K. F.; von Haeseler, A.; Jermiin, L. S., ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **2017,** 14, (6), 587-589.
38. Rambaut, A.; Lam, T. T.; Max Carvalho, L.; Pybus, O. G., Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution* **2016,** 2, (1), vew007.
39. Drummond, A. J.; Suchard, M. A.; Xie, D.; Rambaut, A., Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **2012,** 29, (8), 1969-1973.
40. Suchard, M. A.; Lemey, P.; Baele, G.; Ayres, D. L.; Drummond, A. J.; Rambaut, A., Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* **2018,** 4, (1).
41. Minin, V. N.; Suchard, M. A., Counting labeled transitions in continuous-time Markov models of evolution. *J Math Biol* **2008,** 56, (3), 391-412.
42. Rambaut, A.; Drummond, A. J.; Xie, D.; Baele, G.; Suchard, M. A., Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology* **2018,** 67, (5), 901-904.
43. Rambaut, A. FigTree - Tree Figure Drawing Tool Version 1.4.4. http://tree.bio.ed.ac.uk/ (May 30),
44. Mirdita, M.; Schutze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M., ColabFold: making protein folding accessible to all. *Nat Methods* **2022,** 19, (6), 679-682.
45. Tan, Y. B.; Chmielewski, D.; Law, M. C. Y.; Zhang, K.; He, Y.; Chen, M.; Jin, J.; Luo, D., Molecular architecture of the Chikungunya virus replication complex. *Sci Adv* **2022,** 8, (48), eadd2536.
46. Narwal, M.; Singh, H.; Pratap, S.; Malik, A.; Kuhn, R. J.; Kumar, P.; Tomar, S., Crystal structure of chikungunya virus nsP2 cysteine protease reveals a putative flexible loop blocking its active site. *Int J Biol Macromol* **2018,** 116, 451-462.

47. Song, H.; Zhao, Z.; Chai, Y.; Jin, X.; Li, C.; Yuan, F.; Liu, S.; Gao, Z.; Wang, H.; Song, J.; Vazquez, L.; Zhang, Y.; Tan, S.; Morel, C. M.; Yan, J.; Shi, Y.; Qi, J.; Gao, F.; Gao, G. F., Molecular Basis of Arthritogenic Alphavirus Receptor MXRA8 Binding to Chikungunya Virus Envelope Protein. *Cell* **2019,** 177, (7), 1714-1724 e12.
48. Chen, R.; Puri, V.; Fedorova, N.; Lin, D.; Hari, K. L.; Jain, R.; Rodas, J. D.; Das, S. R.; Shabman, R. S.; Weaver, S. C., Comprehensive Genome Scale Phylogenetic Study Provides New Insights on the Global Expansion of Chikungunya Virus. *J Virol* **2016,** 90, (23), 10600-10611.
49. Giovanetti, M.; Vazquez, C.; Lima, M.; Castro, E.; Rojas, A.; de la Fuente, A. G.; Aquino, C.; Cantero, C.; Fleitas, F.; Torales, J.; Barrios, J.; Ortega, M. J.; Gamarra, M. L.; Villalba, S.; Alfonzo, T.; Xavier, J.; Adelino, T.; Fritsch, H.; Iani, F. C. M.; Pereira, G. C.; de Oliveira, C.; Schuab, G.; Rodrigues, E. S.; Kashima, S.; Leite, J.; Gresh, L.; Franco, L.; Tegally, H.; Van Voorhis, W. C.; Lessels, R.; de Filippis, A. M. B.; Ojeda, A.; Sequera, G.; Montoya, R.; Holmes, E. C.; de Oliveira, T.; Rico, J. M.; Lourenço, J.; Fonseca, V.; Alcantara, L. C. J., Rapid epidemic expansion of chikungunya virus-ECSA lineage in Paraguay. *medRxiv* **2023**.
50. Phadungsombat, J.; Imad, H. A.; Nakayama, E. E.; Leaungwutiwong, P.; Ramasoota, P.; Nguitragool, W.; Matsee, W.; Piyaphanee, W.; Shioda, T., Spread of a Novel Indian Ocean Lineage Carrying E1-K211E/E2-V264A of Chikungunya Virus East/Central/South African Genotype across the Indian Subcontinent, Southeast Asia, and Eastern Africa. *Microorganisms* **2022,** 10, (2).
51. Parola, P.; de Lamballerie, X.; Jourdan, J.; Rovery, C.; Vaillant, V.; Minodier, P.; Brouqui, P.; Flahault, A.; Raoult, D.; Charrel, R. N., Novel chikungunya virus variant in travelers returning from Indian Ocean islands. *Emerg Infect Dis* **2006,** 12, (10), 1493-9.
52. Dellagi, K.; Salez, N.; Maquart, M.; Larrieu, S.; Yssouf, A.; Silaï, R.; Leparc-Goffart, I.; Tortosa, P.; de Lamballerie, X., Serological Evidence of Contrasted Exposure to Arboviral Infections between Islands of the Union of Comoros (Indian Ocean). *PLoS Negl Trop Dis* **2016,** 10, (12), e0004840.
53. Pulmanausahakul, R.; Roytrakul, S.; Auewarakul, P.; Smith, D. R., Chikungunya in Southeast Asia: understanding the emergence and finding solutions. *International Journal of Infectious Diseases* **2011,** 15, (10), e671-e676.
54. Wahid, B.; Ali, A.; Rafique, S.; Idrees, M., Global expansion of chikungunya virus: mapping the 64-year history. *International Journal of Infectious Diseases* **2017,** 58, 69-76.
55. Sahadeo, N. S. D.; Allicock, O. M.; De Salazar, P. M.; Auguste, A. J.; Widen, S.; Olowokure, B.; Gutierrez, C.; Valadere, A. M.; Polson-Edwards, K.; Weaver, S. C.; Carrington, C. V. F., Understanding the evolution and spread of chikungunya virus in the Americas using complete genome sequences. *Virus Evolution* **2017,** 3, (1).
56. Phadungsombat, J.; Imad, H.; Rahman, M.; Nakayama, E. E.; Kludkleeb, S.; Ponam, T.; Rahim, R.; Hasan, A.; Poltep, K.; Yamanaka, A.; Matsee, W.; Piyaphanee, W.; Phumratanaprapin, W.; Shioda, T., A Novel Sub-Lineage of Chikungunya Virus East/Central/South African Genotype Indian Ocean Lineage Caused Sequential Outbreaks in Bangladesh and Thailand. *Viruses* **2020,** 12, (11).
57. Jones, R.; Bragagnolo, G.; Arranz, R.; Reguera, J., Capping pores of alphavirus nsP1 gate membranous viral replication factories. *Nature* **2021,** 589, (7843), 615-619.
58. Cheng, R. H.; Kuhn, R. J.; Olson, N. H.; Rossmann, M. G.; Choi, H. K.; Smith, T. J.; Baker, T. S., Nucleocapsid and glycoprotein organization in an enveloped virus. *Cell* **1995,** 80, (4), 621-30.
59. Li, L.; Jose, J.; Xiang, Y.; Kuhn, R. J.; Rossmann, M. G., Structural changes of envelope proteins during alphavirus fusion. *Nature* **2010,** 468, (7324), 705-8.
60. Lescar, J.; Roussel, A.; Wien, M. W.; Navaza, J.; Fuller, S. D.; Wengler, G.; Wengler, G.; Rey, F. A., The Fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. *Cell* **2001,** 105, (1), 137-48.

61. Lee, R. C.; Hapuarachchi, H. C.; Chen, K. C.; Hussain, K. M.; Chen, H.; Low, S. L.; Ng, L. C.; Lin, R.; Ng, M. M.; Chu, J. J., Mosquito cellular factors and functions in mediating the infectious entry of chikungunya virus. *PLoS Negl Trop Dis* **2013,** 7, (2), e2050.
62. Zhang, R.; Kim, A. S.; Fox, J. M.; Nair, S.; Basore, K.; Klimstra, W. B.; Rimkunas, R.; Fong, R. H.; Lin, H.; Poddar, S.; Crowe, J. E., Jr.; Doranz, B. J.; Fremont, D. H.; Diamond, M. S., Mxra8 is a receptor for multiple arthritogenic alphaviruses. *Nature* **2018,** 557, (7706), 570-574.
63. Wahlberg, J. M.; Garoff, H., Membrane fusion process of Semliki Forest virus. I: Low pH-induced rearrangement in spike protein quaternary structure precedes virus penetration into cells. *J Cell Biol* **1992,** 116, (2), 339-48.
64. Voss, J. E.; Vaney, M. C.; Duquerroy, S.; Vonrhein, C.; Girard-Blanc, C.; Crublet, E.; Thompson, A.; Bricogne, G.; Rey, F. A., Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. *Nature* **2010,** 468, (7324), 709-12.
65. Yap, M. L.; Klose, T.; Urakami, A.; Hasan, S. S.; Akahata, W.; Rossmann, M. G., Structural studies of Chikungunya virus maturation. *Proc Natl Acad Sci U S A* **2017,** 114, (52), 13703-13707.
66. Mangala Prasad, V.; Blijleven, J. S.; Smit, J. M.; Lee, K. K., Visualization of conformational changes and membrane remodeling leading to genome delivery by viral class-II fusion machinery. *Nat Commun* **2022,** 13, (1), 4772.
67. Sahoo, B.; Gudigamolla, N. K.; Chowdary, T. K., Acidic pH-Induced Conformational Changes in Chikungunya Virus Fusion Protein E1: a Spring-Twisted Region in the Domain I-III Linker Acts as a Hinge Point for Swiveling Motion of Domains. *J Virol* **2020,** 94, (23).
68. Su, L.; Lou, X.; Yan, H.; Yang, Z.; Mao, H.; Yao, W.; Sun, Y.; Pan, J.; Zhang, Y., Importation of a novel Indian Ocean lineage carrying E1-K211E and E2-V264A of Chikungunya Virus in Zhejiang Province, China, in 2019. *Virus Genes* **2023,** 59, (5), 693-702.
69. Agarwal, A.; Sharma, A. K.; Sukumaran, D.; Parida, M.; Dash, P. K., Two novel epistatic mutations (E1:K211E and E2:V264A) in structural proteins of Chikungunya virus enhance fitness in Aedes aegypti. *Virology* **2016,** 497, 59-68.
70. Rangel, M. V.; McAllister, N.; Dancel-Manning, K.; Noval, M. G.; Silva, L. A.; Stapleford, K. A., Emerging Chikungunya Virus Variants at the E1-E1 Interglycoprotein Spike Interface Impact Virus Attachment and Inflammation. *J Virol* **2022,** 96, (4), e0158621.
71. Tsetsarkin, K. A.; Vanlandingham, D. L.; McGee, C. E.; Higgs, S., A single mutation in chikungunya virus affects vector specificity and epidemic potential. *PLoS Pathog* **2007,** 3, (12), e201.
72. Tsetsarkin, K. A.; Chen, R.; Yun, R.; Rossi, S. L.; Plante, K. S.; Guerbois, M.; Forrester, N.; Perng, G. C.; Sreekumar, E.; Leal, G.; Huang, J.; Mukhopadhyay, S.; Weaver, S. C., Multi-peaked adaptive landscape for chikungunya virus evolution predicts continued fitness optimization in Aedes albopictus mosquitoes. *Nat Commun* **2014,** 5, 4084.
73. Fortuna, C.; Toma, L.; Remoli, M. E.; Amendola, A.; Severini, F.; Boccolini, D.; Romi, R.; Venturi, G.; Rezza, G.; Di Luca, M., Vector competence of Aedes albopictus for the Indian Ocean lineage (IOL) chikungunya viruses of the 2007 and 2017 outbreaks in Italy: a comparison between strains with and without the E1:A226V mutation. *Eurosurveillance* **2018,** 23, (22), 1800246.
74. Kumar, A.; Mamidi, P.; Das, I.; Nayak, T. K.; Kumar, S.; Chhatai, J.; Chattopadhyay, S.; Suryawanshi, A. R.; Chattopadhyay, S., A novel 2006 Indian outbreak strain of Chikungunya virus exhibits different pattern of infection as compared to prototype strain. *PLoS One* **2014,** 9, (1), e85714.
75. Arias-Goeta, C.; Mousson, L.; Rougeon, F.; Failloux, A. B., Dissemination and transmission of the E1-226V variant of chikungunya virus in Aedes albopictus are controlled at the midgut barrier level. *PLoS One* **2013,** 8, (2), e57548.

76. Vazeille, M.; Moutailler, S.; Coudrier, D.; Rousseaux, C.; Khun, H.; Huerre, M.; Thiria, J.; Dehecq, J. S.; Fontenille, D.; Schuffenecker, I.; Despres, P.; Failloux, A. B., Two Chikungunya isolates from the outbreak of La Reunion (Indian Ocean) exhibit different patterns of infection in the mosquito, Aedes albopictus. *PLoS One* **2007,** 2, (11), e1168.
77. Kim, A. S.; Zimmerman, O.; Fox, J. M.; Nelson, C. A.; Basore, K.; Zhang, R.; Durnell, L.; Desai, C.; Bullock, C.; Deem, S. L.; Oppenheimer, J.; Shapiro, B.; Wang, T.; Cherry, S.; Coyne, C. B.; Handley, S. A.; Landis, M. J.; Fremont, D. H.; Diamond, M. S., An Evolutionary Insertion in the Mxra8 Receptor-Binding Site Confers Resistance to Alphavirus Infection and Pathogenesis. *Cell Host & Microbe* **2020,** 27, (3), 428-440.e9.
78. Porta, J.; Mangala Prasad, V.; Wang, C. I.; Akahata, W.; Ng, L. F.; Rossmann, M. G., Structural Studies of Chikungunya Virus-Like Particles Complexed with Human Antibodies: Neutralization and Cell-to-Cell Transmission. *J Virol* **2016,** 90, (3), 1169-77.
79. Zhou, Q. F.; Fox, J. M.; Earnest, J. T.; Ng, T. S.; Kim, A. S.; Fibriansah, G.; Kostyuchenko, V. A.; Shi, J.; Shu, B.; Diamond, M. S.; Lok, S. M., Structural basis of Chikungunya virus inhibition by monoclonal antibodies. *Proc Natl Acad Sci U S A* **2020,** 117, (44), 27637-27645.
80. Pal, P.; Dowd, K. A.; Brien, J. D.; Edeling, M. A.; Gorlatov, S.; Johnson, S.; Lee, I.; Akahata, W.; Nabel, G. J.; Richter, M. K.; Smit, J. M.; Fremont, D. H.; Pierson, T. C.; Heise, M. T.; Diamond, M. S., Development of a highly protective combination monoclonal antibody therapy against Chikungunya virus. *PLoS Pathog* **2013,** 9, (4), e1003312.
81. Gaunt, M. W.; Gubler, D. J.; Pettersson, J. H. O.; Kuno, G.; Wilder-Smith, A.; de Lamballerie, X.; Gould, E. A.; Falconar, A. K., Recombination of B- and T-cell epitope-rich loci from Aedes- and Culex-borne flaviviruses shapes Zika virus epidemiology. *Antiviral Research* **2020,** 174, 104676.
82. Götte, B.; Liu, L.; McInerney, G. M., The Enigmatic Alphavirus Non-Structural Protein 3 (nsP3) Revealing Its Secrets at Last. *Viruses* **2018,** 10, (3).
83. McPherson, R. L.; Abraham, R.; Sreekumar, E.; Ong, S.-E.; Cheng, S.-J.; Baxter, V. K.; Kistemaker, H. A.; Filippov, D. V.; Griffin, D. E.; Leung, A. K., ADP-ribosylhydrolase activity of Chikungunya virus macrodomain is critical for virus replication and virulence. *Proceedings of the National Academy of Sciences* **2017,** 114, (7), 1666-1671.
84. Gao, Y.; Goonawardane, N.; Ward, J.; Tuplin, A.; Harris, M., Multiple roles of the non-structural protein 3 (nsP3) alphavirus unique domain (AUD) during Chikungunya virus genome replication and transcription. *PLoS pathogens* **2019,** 15, (1), e1007239.
85. Meshram, C. D.; Agback, P.; Shiliaev, N.; Urakova, N.; Mobley, J. A.; Agback, T.; Frolova, E. I.; Frolov, I., Multiple host factors interact with the hypervariable domain of chikungunya virus nsP3 and determine viral replication in cell-specific mode. *Journal of virology* **2018,** 92, (16), 10.1128/jvi. 00838-18.
86. Mutso, M.; Morro, A. M.; Smedberg, C.; Kasvandik, S.; Aquilimeba, M.; Teppor, M.; Tarve, L.; Lulla, A.; Lulla, V.; Saul, S., Mutation of CD2AP and SH3KBP1 binding motif in alphavirus nsP3 hypervariable domain results in attenuated virus. *Viruses* **2018,** 10, (5), 226.
87. Gao, Y.; Goonawardane, N.; Ward, J.; Tuplin, A.; Harris, M., Multiple roles of the non-structural protein 3 (nsP3) alphavirus unique domain (AUD) during Chikungunya virus genome replication and transcription. *PLoS Pathog* **2019,** 15, (1), e1007239.
88. Neuvonen, M.; Kazlauskas, A.; Martikainen, M.; Hinkkanen, A.; Ahola, T.; Saksela, K., SH3 domain-mediated recruitment of host cell amphiphysins by alphavirus nsP3 promotes viral RNA replication. *PLoS Pathog* **2011,** 7, (11), e1002383.
89. Panas, M. D.; Schulte, T.; Thaa, B.; Sandalova, T.; Kedersha, N.; Achour, A.; McInerney, G. M., Viral and cellular proteins containing FGDF motifs bind G3BP to block stress granule formation. *PLoS Pathog* **2015,** 11, (2), e1004659.
90. Mihalic, F.; Simonetti, L.; Giudice, G.; Sander, M. R.; Lindqvist, R.; Peters, M. B. A.; Benz, C.; Kassa, E.; Badgujar, D.; Inturi, R.; Ali, M.; Krystkowiak, I.; Sayadi, A.; Andersson, E.; Aronsson, H.; Soderberg, O.; Dobritzsch, D.; Petsalaki, E.; Overby, A. K.; Jemth, P.;

Davey, N. E.; Ivarsson, Y., Large-scale phage-based screening reveals extensive pan-viral mimicry of host short linear motifs. *Nat Commun* **2023,** 14, (1), 2409.

91. Melton, J. V.; Ewart, G. D.; Weir, R. C.; Board, P. G.; Lee, E.; Gage, P. W., Alphavirus 6K proteins form ion channels. *J Biol Chem* **2002,** 277, (49), 46923-31.

92. Dey, D.; Siddiqui, S. I.; Mamidi, P.; Ghosh, S.; Kumar, C. S.; Chattopadhyay, S.; Ghosh, S.; Banerjee, M., The effect of amantadine on an ion channel protein from Chikungunya virus. *PLoS Negl Trop Dis* **2019,** 13, (7), e0007548.

93. Lulla, V.; Kim, D. Y.; Frolova, E. I.; Frolov, I., The amino-terminal domain of alphavirus capsid protein is dispensable for viral particle assembly but regulates RNA encapsidation through cooperative functions of its subdomains. *J Virol* **2013,** 87, (22), 12003-19.

94. Choi, H. K.; Tong, L.; Minor, W.; Dumas, P.; Boege, U.; Rossmann, M. G.; Wengler, G., Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. *Nature* **1991,** 354, (6348), 37-43.

95. Maljkovic Berry, I.; Eyase, F.; Pollett, S.; Konongoi, S. L.; Joyce, M. G.; Figueroa, K.; Ofula, V.; Koka, H.; Koskei, E.; Nyunja, A.; Mancuso, J. D.; Jarman, R. G.; Sang, R., Global Outbreaks and Origins of a Chikungunya Virus Variant Carrying Mutations Which May Increase Fitness for Aedes aegypti: Revelations from the 2016 Mandera, Kenya Outbreak. *Am J Trop Med Hyg* **2019,** 100, (5), 1249-1257.

96. Tsetsarkin, K. A.; Chen, R.; Leal, G.; Forrester, N.; Higgs, S.; Huang, J.; Weaver, S. C., Chikungunya virus emergence is constrained in Asia by lineage-specific adaptive landscapes. *Proc Natl Acad Sci U S A* **2011,** 108, (19), 7872-7.

97. Tsetsarkin, K. A.; McGee, C. E.; Volk, S. M.; Vanlandingham, D. L.; Weaver, S. C.; Higgs, S., Epistatic roles of E2 glycoprotein mutations in adaption of chikungunya virus to Aedes albopictus and Ae. aegypti mosquitoes. *PLoS One* **2009,** 4, (8), e6835.

98. Hawman, D. W.; Carpentier, K. S.; Fox, J. M.; May, N. A.; Sanders, W.; Montgomery, S. A.; Moorman, N. J.; Diamond, M. S.; Morrison, T. E., Mutations in the E2 Glycoprotein and the 3' Untranslated Region Enhance Chikungunya Virus Virulence in Mice. *J Virol* **2017,** 91, (20).

99. Singh, R. K.; Tiwari, S.; Mishra, V. K.; Tiwari, R.; Dhole, T. N., Molecular epidemiology of Chikungunya virus: mutation in E1 gene region. *J Virol Methods* **2012,** 185, (2), 213-20.

100. Silva, J. V. J., Jr.; Ludwig-Begall, L. F.; Oliveira-Filho, E. F.; Oliveira, R. A. S.; Durães-Carvalho, R.; Lopes, T. R. R.; Silva, D. E. A.; Gil, L., A scoping review of Chikungunya virus infection: epidemiology, clinical characteristics, viral co-circulation complications, and control. *Acta Trop* **2018,** 188, 213-224.

101. Theamboonlers, A.; Rianthavorn, P.; Praianantathavorn, K.; Wuttirattanakowit, N.; Poovorawan, Y., Clinical and molecular characterization of chikungunya virus in South Thailand. *Jpn J Infect Dis* **2009,** 62, (4), 303-5.

102. Pulmanausahakul, R.; Roytrakul, S.; Auewarakul, P.; Smith, D. R., Chikungunya in Southeast Asia: understanding the emergence and finding solutions. *Int J Infect Dis* **2011,** 15, (10), e671-6.

103. Powers, A. M.; Brault, A. C.; Tesh, R. B.; Weaver, S. C., Re-emergence of Chikungunya and O'nyong-nyong viruses: evidence for distinct geographical lineages and distant evolutionary relationships. *J Gen Virol* **2000,** 81, (Pt 2), 471-9.

104. Shrinet, J.; Jain, S.; Sharma, A.; Singh, S. S.; Mathur, K.; Rana, V.; Bhatnagar, R. K.; Gupta, B.; Gaind, R.; Deb, M.; Sunil, S., Genetic characterization of Chikungunya virus from New Delhi reveal emergence of a new molecular signature in Indian isolates. *Virol J* **2012,** 9, 100.

105. Patil, J.; More, A.; Patil, P.; Jadhav, S.; Newase, P.; Agarwal, M.; Amdekar, S.; Raut, C. G.; Parashar, D.; Cherian, S. S., Genetic characterization of chikungunya viruses isolated during the 2015-2017 outbreaks in different states of India, based on their E1 and E2 genes. *Arch Virol* **2018,** 163, (11), 3135-3140.

106. Badar, N.; Salman, M.; Ansari, J.; Aamir, U.; Alam, M. M.; Arshad, Y.; Mushtaq, N.; Ikram, A.; Qazi, J., Emergence of Chikungunya Virus, Pakistan, 2016-2017. *Emerg Infect Dis* **2020,** 26, (2), 307-310.
107. Lindh, E.; Argentini, C.; Remoli, M. E.; Fortuna, C.; Faggioni, G.; Benedetti, E.; Amendola, A.; Marsili, G.; Lista, F.; Rezza, G.; Venturi, G., The Italian 2017 Outbreak Chikungunya Virus Belongs to an Emerging Aedes albopictus-Adapted Virus Cluster Introduced From the Indian Subcontinent. *Open Forum Infect Dis* **2019,** 6, (1), ofy321.
108. Melan, A.; Aung, M. S.; Khanam, F.; Paul, S. K.; Riaz, B. K.; Tahmina, S.; Kabir, M. I.; Hossain, M. A.; Kobayashi, N., Molecular characterization of chikungunya virus causing the 2017 outbreak in Dhaka, Bangladesh. *New Microbes New Infect* **2018,** 24, 14-16.
109. ECDC Chikungunya worldwide overview. https://www.ecdc.europa.eu/en/chikungunya-monthly (20. October),
110. PAHO With rising cases, experts discuss Chikungunya spread in the Americas. https://www.paho.org/en/news/4-5-2023-rising-cases-experts-discuss-chikungunya-spread-americas#:~:text=Situation%20in%20the%20Americas&text=Over%20a%20million%20cases%20were,over%20214%2C000%20cases%20reported%20cases. (20. October),
111. Mercier, A.; Obadia, T.; Carraretto, D.; Velo, E.; Gabiane, G.; Bino, S.; Vazeille, M.; Gasperi, G.; Dauga, C.; Malacrida, A. R.; Reiter, P.; Failloux, A. B., Impact of temperature on dengue and chikungunya transmission by the mosquito Aedes albopictus. *Sci Rep* **2022,** 12, (1), 6973.
112. Jourdain, F.; Roiz, D.; de Valk, H.; Noel, H.; L'Ambert, G.; Franke, F.; Paty, M. C.; Guinard, A.; Desenclos, J. C.; Roche, B., From importation to autochthonous transmission: Drivers of chikungunya and dengue emergence in a temperate area. *PLoS Negl Trop Dis* **2020,** 14, (5), e0008320.
113. Fischer, D.; Thomas, S. M.; Suk, J. E.; Sudre, B.; Hess, A.; Tjaden, N. B.; Beierkuhnlein, C.; Semenza, J. C., Climate change effects on Chikungunya transmission in Europe: geospatial analysis of vector's climatic suitability and virus' temperature requirements. *International Journal of Health Geographics* **2013,** 12, (1), 51.
114. Tjaden, N. B.; Suk, J. E.; Fischer, D.; Thomas, S. M.; Beierkuhnlein, C.; Semenza, J. C., Modelling the effects of global climate change on Chikungunya transmission in the 21st century. *Scientific Reports* **2017,** 7, (1), 3813.
115. Leparc-Goffart, I.; Nougairede, A.; Cassadou, S.; Prat, C.; de Lamballerie, X., Chikungunya in the Americas. *The Lancet* **2014,** 383, (9916), 514.
116. Marengo, J. A.; Ambrizzi, T.; Barreto, N.; Cunha, A. P.; Ramos, A. M.; Skansi, M.; Molina Carpio, J.; Salinas, R., The heat wave of October 2020 in central South America. *International Journal of Climatology* **2022,** 42, (4), 2281-2298.
117. Cai, W.; McPhaden, M. J.; Grimm, A. M.; Rodrigues, R. R.; Taschetto, A. S.; Garreaud, R. D.; Dewitte, B.; Poveda, G.; Ham, Y.-G.; Santoso, A., Climate impacts of the El Niño–southern oscillation on South America. *Nature Reviews Earth & Environment* **2020,** 1, (4), 215-231.
118. Ellwanger, J. H.; Kulmann-Leal, B.; Kaminski, V. L.; Valverde-Villegas, J. M.; Veiga, A.; Spilki, F. R.; Fearnside, P. M.; Caesar, L.; Giatti, L. L.; Wallau, G. L.; Almeida, S. E. M.; Borba, M. R.; Hora, V. P. D.; Chies, J. A. B., Beyond diversity loss and climate change: Impacts of Amazon deforestation on infectious diseases and public health. *An Acad Bras Cienc* **2020,** 92, (1), e20191375.
119. Olmos, M. B.; Bostik, V., Climate Change and Human Security-The Proliferation of Vector-Borne Diseases due to Climate Change. *Military Medical Science Letters/Vojenské Zdravotnické Listy* **2021,** 90, (2).
120. Sadeghieh, T.; Sargeant, J. M.; Greer, A. L.; Berke, O.; Dueymes, G.; Gachon, P.; Ogden, N. H.; Ng, V., Zika virus outbreak in Brazil under current and future climate. *Epidemics* **2021,** 37, 100491.

121. Lippi, C. A.; Stewart-Ibarra, A. M.; Loor, M. F. B.; Zambrano, J. E. D.; Lopez, N. A. E.; Blackburn, J. K.; Ryan, S. J., Geographic shifts in Aedes aegypti habitat suitability in Ecuador using larval surveillance data and ecological niche modeling: Implications of climate change for public health vector control. *PLoS neglected tropical diseases* **2019,** 13, (4), e0007322.
122. Colón-González, F. J.; Sewe, M. O.; Tompkins, A. M.; Sjödin, H.; Casallas, A.; Rocklöv, J.; Caminade, C.; Lowe, R., Projecting the risk of mosquito-borne diseases in a warmer and more populated world: a multi-model, multi-scenario intercomparison modelling study. *The Lancet Planetary Health* **2021,** 5, (7), e404-e414.
123. Kolimenakis, A.; Heinz, S.; Wilson, M. L.; Winkler, V.; Yakob, L.; Michaelakis, A.; Papachristos, D.; Richardson, C.; Horstick, O., The role of urbanisation in the spread of Aedes mosquitoes and the diseases they transmit—A systematic review. *PLOS Neglected Tropical Diseases* **2021,** 15, (9), e0009631.

Figure legends

Fig. 1. Phylogenetic and Phylogeographic Analysis of the ECSA CHIKV samples included in the present study. Phylogenetic tree of the CHIKV ECSA genotype. Strains are colour-coded by CHIKV lineage (where light green represents ECSA 1, red ECSA 2, and blue IOL). Strains sequenced in this study are indicated by an ^{x}. Black values at red circles indicate branching times, grey values display the mutational speed. Major (basal) branches with posterior probabilities of ≥0.95 are indicated by an *. Amino acid substitutions at different alphabetically characterized nodes are indicated in green boxes if affecting the entire IOL and in orange boxes if affecting only certain clades. Countries are colour-coded according to the phylogenetic tree, and suggested transmission events are shown with arrows on the global map. Uninterrupted lines on the map show data from our investigation, while dashed lines display previous assumptions of spread.

Fig. 2: Emergent mutations in the Nsp1-Nsp2-Nsp4 replication complex. A) Representative view of the CHIKV replication complex. Nsp1 is in grey, Nsp2 in yellow and Nsp4 in pink colour. Structural models for visualization were obtained from PDBid 7y38 [45] and 4ztb [46]. In all panels the mutations arising in the IOL are shown as purple sticks. Note that highlighted amino acids correspond to the position of the indicated mutations but not always to the actual amino acid involved in the mutation process. (see methods for more information) **B)** Close-up view, highlighting the mutations on Nsp1. The GTP and ATP which are cofactors for the Nsp1 are also shown as orange sticks. **C)** Model of Nsp2 with indicated mutations. The catalytic residues are shown as green sticks and the RNA fragment bound to the helicase domain is shown as orange sticks. **D)** Detailed view showcasing the mutations on Nsp4. **E)** Summary of all mutations found in the replication complex.

Fig. 3: Emergent mutations in the E1-E2-E3 spike complex. A) Representative view of the Spike trimer. Structural models for visualization were obtained from PDBid 6jo8 [47]. E1 is coloured cyan, E2 blue and E3 is green. The receptor Mxra8 is included to aid the visualization of the receptor binding interface and is coloured brown. The position of the membrane at the base of the spike is indicated. All emergent mutations are shown as sticks and coloured purple as in Fig. 2. **B)** Close-up view, highlighting the mutations on E1. The domains I, II and III are indicated. **C)** enlarged model of E2 highlighting emergent mutations. Domains A, B and C are indicated. **D)** Close-up view, showcasing the mutations on E3. **E)** Summary of all mutations found in the spike

complex. Asterix denotes mutations that are not part of the analysed structure and are therefore not visualized.

Fig. 4: Emergent mutations in the Nsp3, 6K and CP proteins. All mutations were mapped onto the Colabfold-predicted structural models. Positions of mutations are shown as purple sticks, and the overall colouring of the proteins is according to the pIDDT score indicating the confidence of the prediction. Blue colour signifies highest prediction confidence and orange lowest as shown by the legend. **A)** Emergent mutations in Nsp3 protein. Macro and Zinc-binding domains are indicated. **B)** Predicted structure of 6K protein with highlighted V6I mutation. **C)** Capsid protein model with highlighted mutations in the disordered N-terminal tail of the protein. **D)** Summary of all emergent mutations found in the Nsp3, 6K and capsid proteins.

Supplementary information

Supplementary Fig. S1. A maximum likelihood tree of all 2,564 complete or near complete CHIKV genome sequences. Sequences that were included in the subsampled dataset were colour coded according colours in Fig. 1.

Supplementary Fig. S2. TempEst regression for the ECSA genotype. Data points were colour coded according colours in Fig. 1.

Supplementary Fig. S3. Maximum clade credibility tree from Fig. 1, here depicting all branches with posterior probability of ≥ 0.95.

Supplementary Table S1. Includes strain name, sampling date, geographical location and NCBI GenBank accession number for all twelve patient samples sequenced in the present study.

Author contribution

Conceptualization Ideas: JHOP, XdL, WS; Funding Acquisition: JHOP, ÅL, WS; Investigation: JK, FM; Formal Analysis: MWG, JB, JH, CL; Visualization: JK, FM; Writing – Original Draft Preparation: JK, FM; Writing – Editing: JK, FM, JHOP; Review: JK, MG, JB, ÅL, JH, XdL, JHOP.

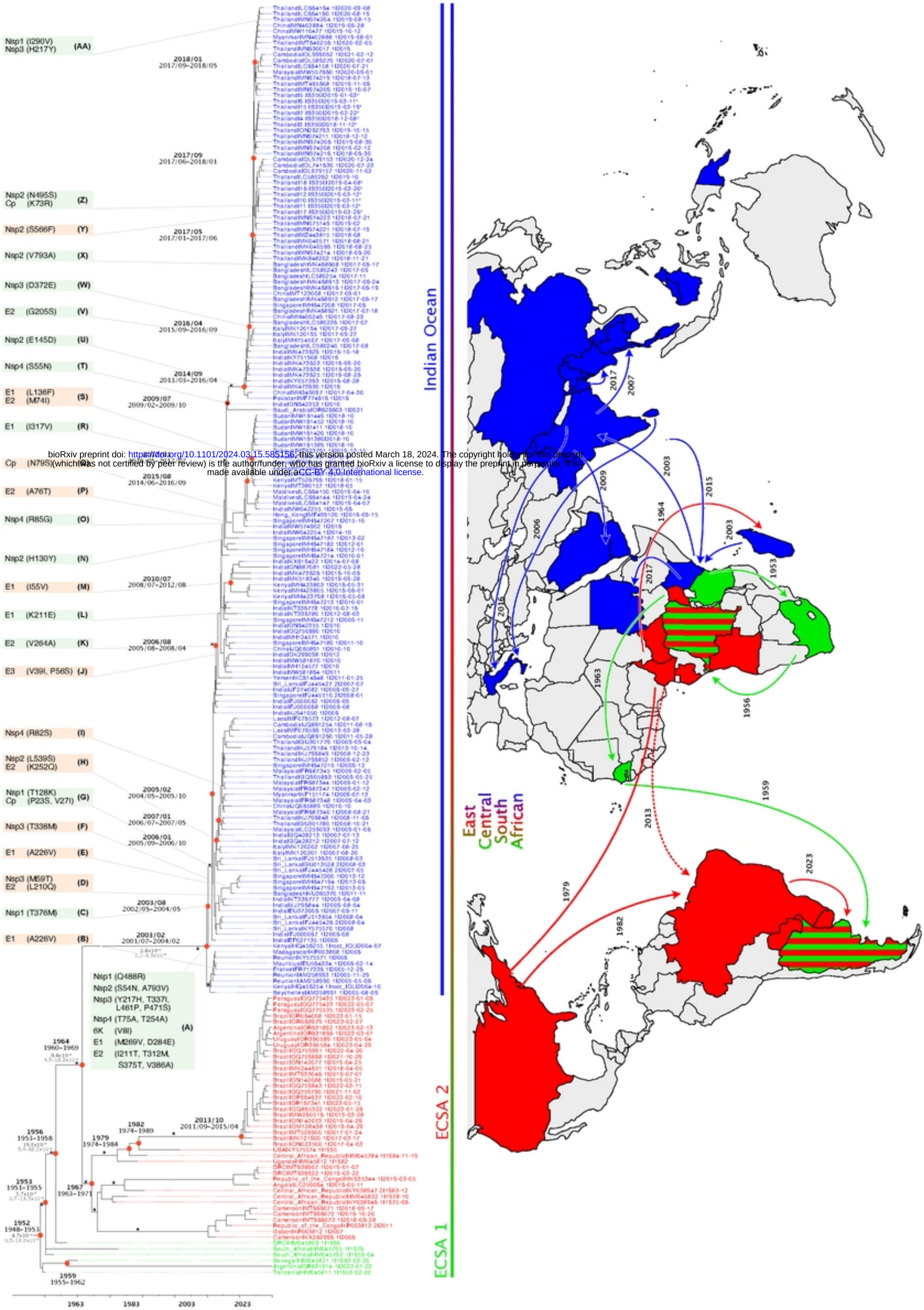
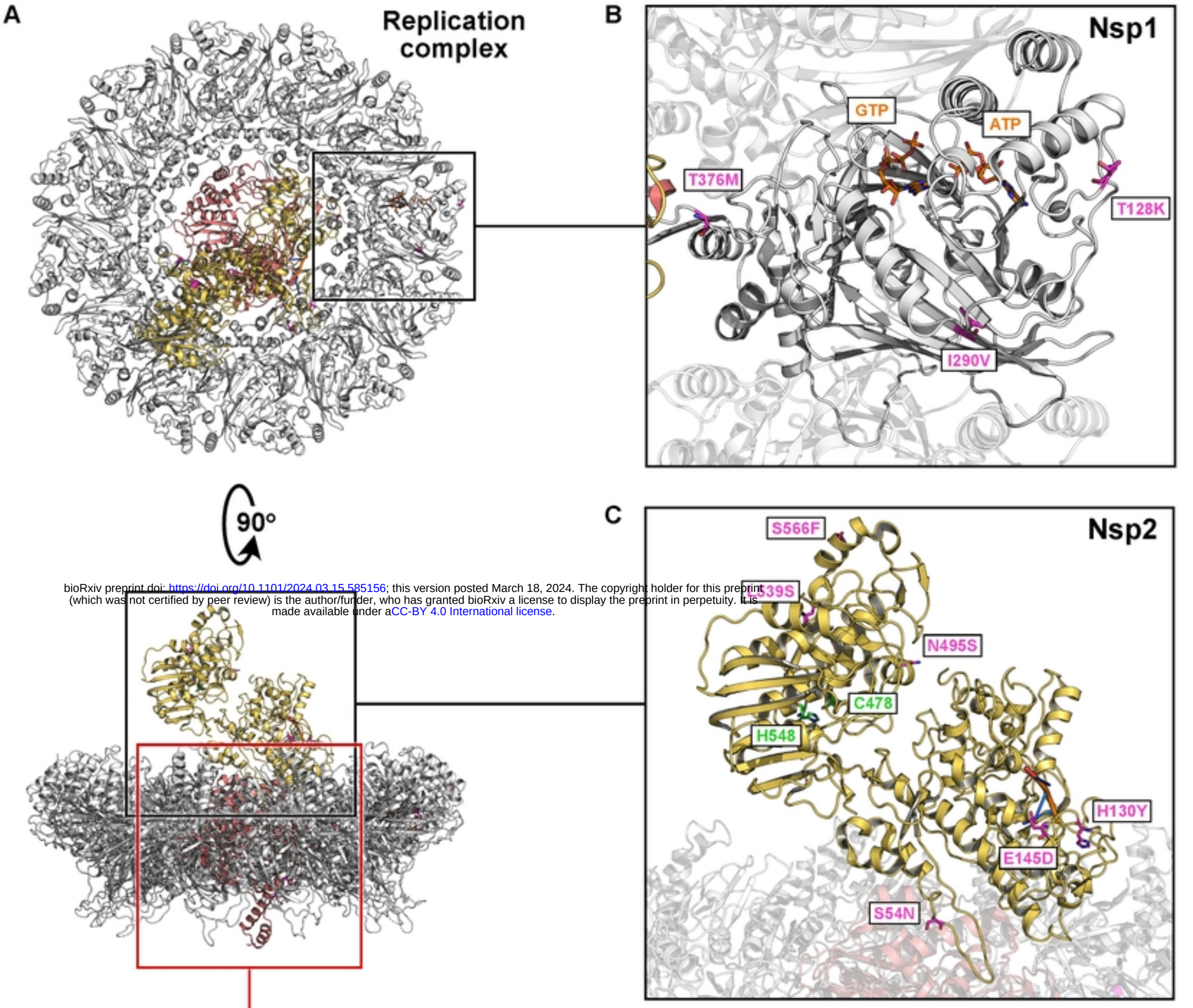


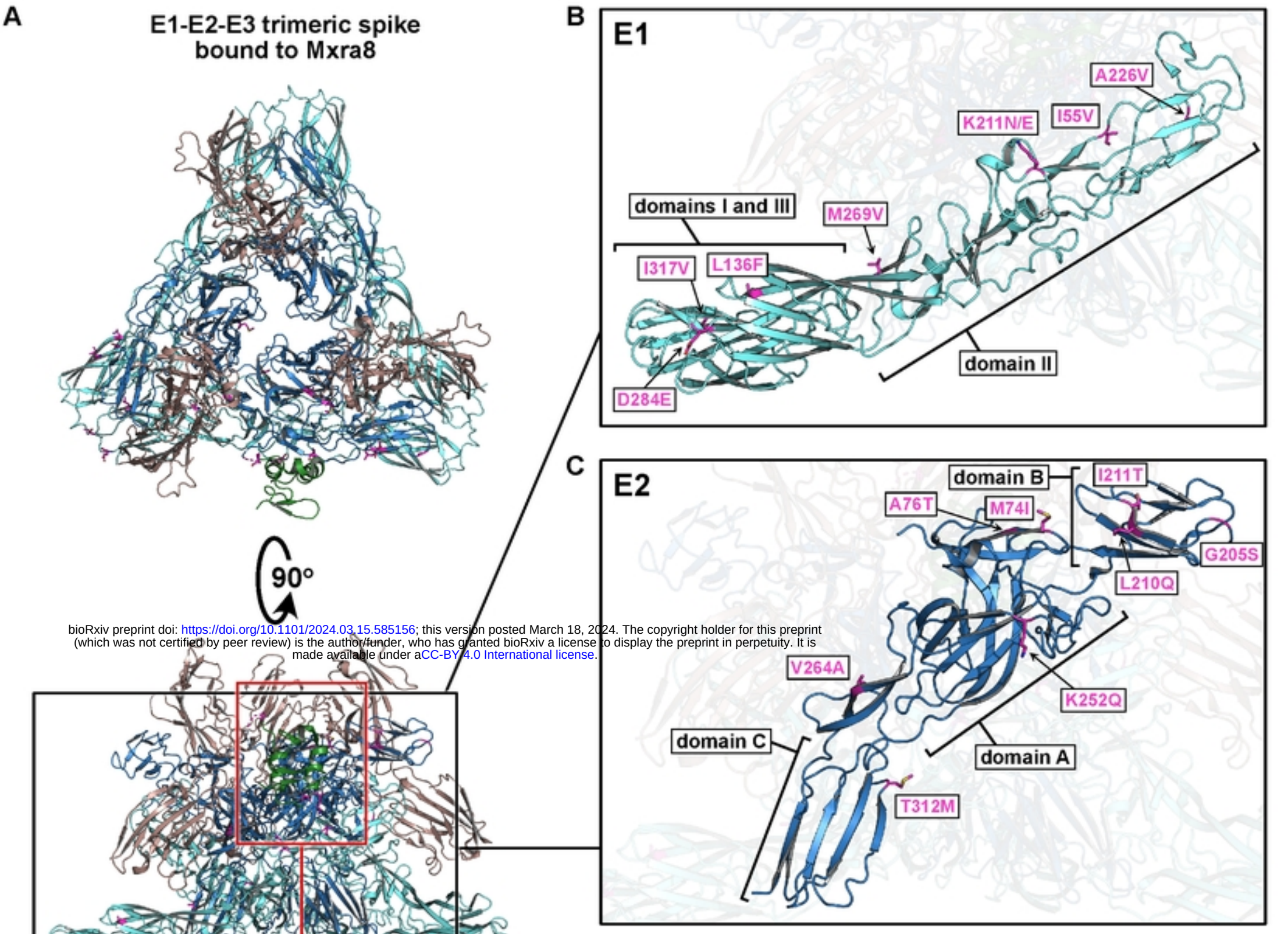
Figure 1



E

protein	mutation	conservation	location/contacts
Nsp1	T128K	polar to charged	surface exposed
	I290V	conservative	hydrophobic core
	T376M	polar to hydrophobic	ambiguous
	Q488R *	polar to charged	surface exposed
Nsp2	S54N	conservative	Nsp4 interface
	H130Y	semi-conservative	surface exposed
	E145D	conservative	surface exposed
	N495S	conservative	surface exposed
	L539S	hydrophobic to polar	surface exposed
	S566F	polar to aromatic	surface exposed
	A793V *	semi-conservative	surface exposed
Nsp4	S55N	conservative	surface exposed
	T75A	semi-conservative	surface exposed
	R82S	charged to polar	surface exposed
	R85G	charged to small flexible	surface exposed
	T254A	semi-conservative	surface exposed

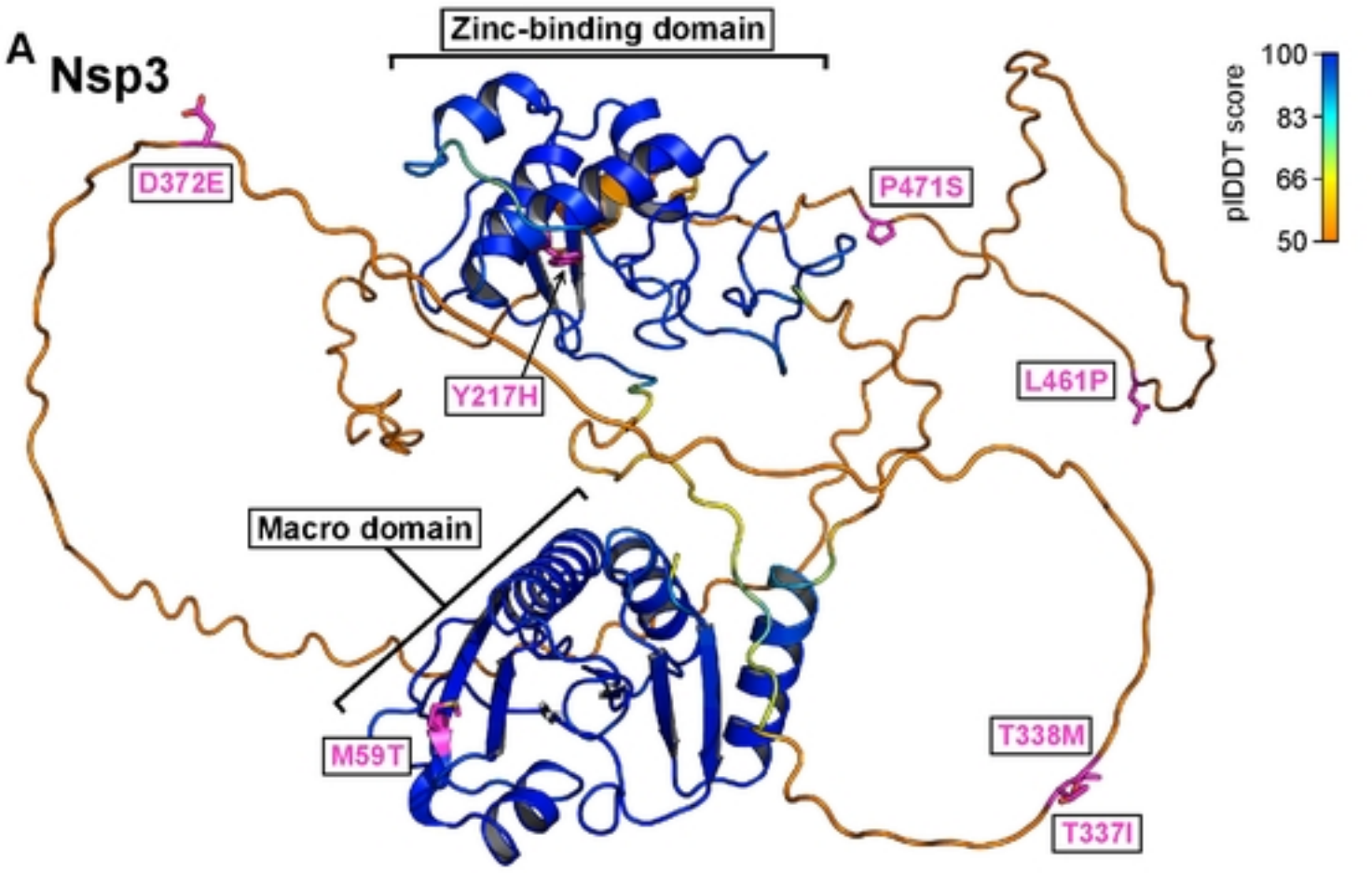
Figure 2

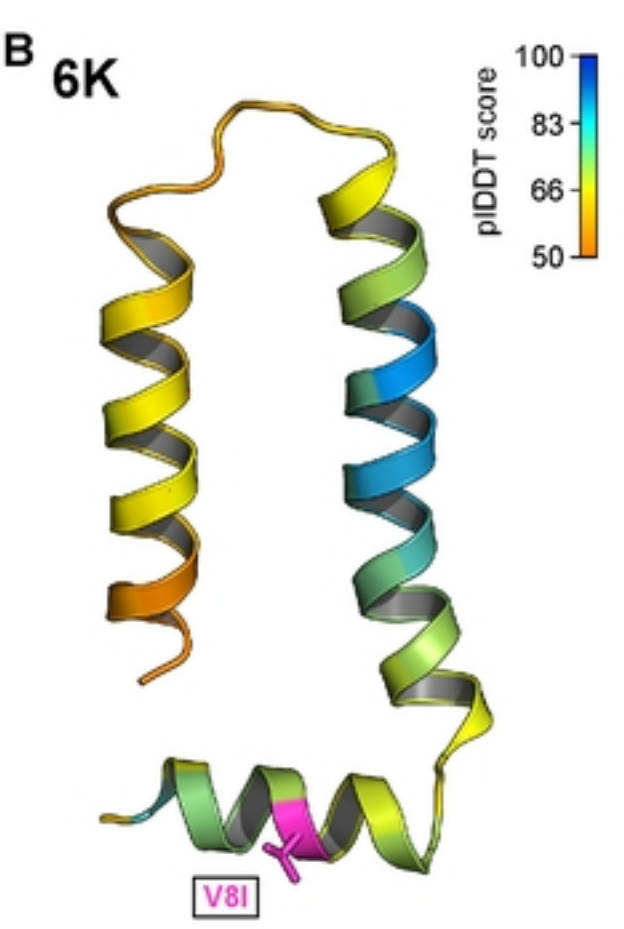


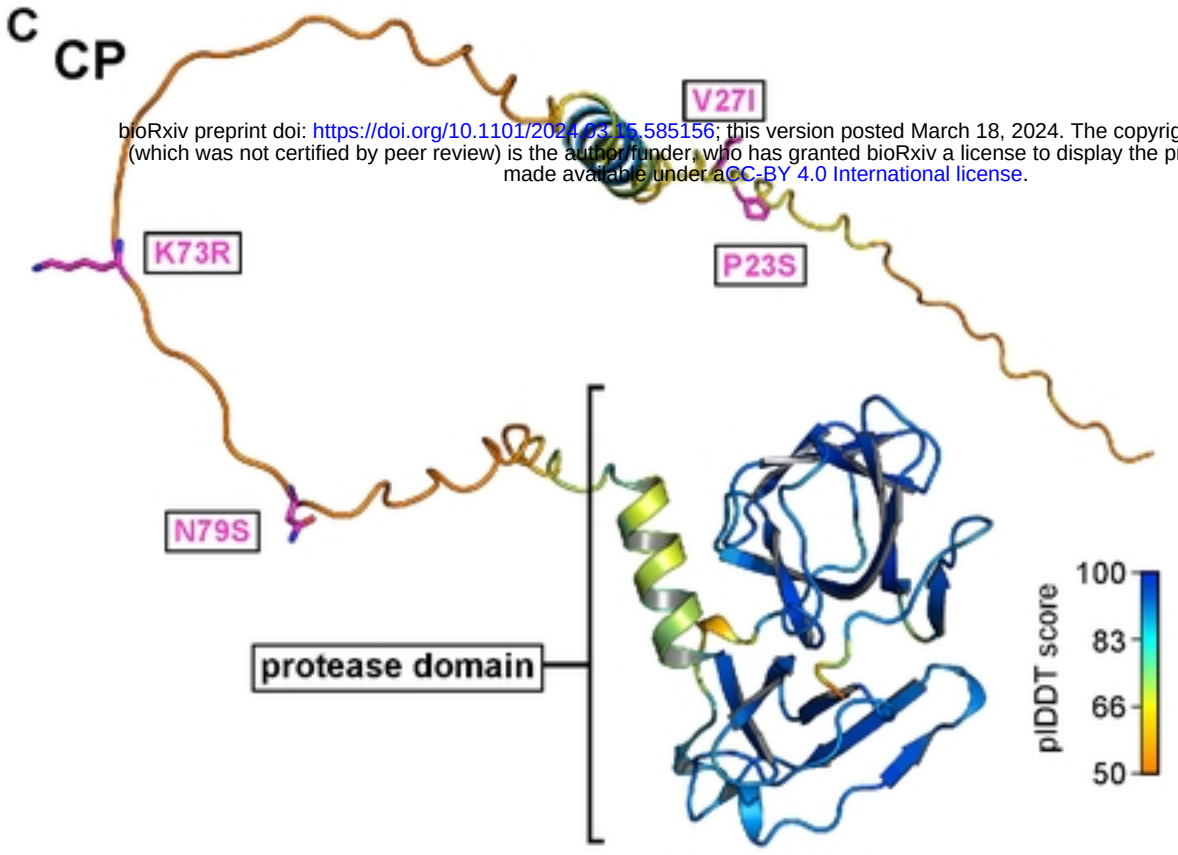
E

protein	mutation	conservation	location/contacts
E1	I55V	conservative	hydrophobic core
	L136F	semi-conservative	surface exposed
	K211E	polar to charged	surface exposed
	A226V	semi-conservative	Mxra8 interface
	M269V	conservative	surface exposed
	D284E	conservative	surface exposed
	I317V	conservative	surface exposed
E2	M74I	conservative	Mxra8 interface
	A76T	semi-conservative	Mxra8 interface
	G205S	small flexible to polar	surface exposed
	L210Q	hydrophobic to polar	surface exposed
	I211T	hydrophobic to polar	surface exposed
	K252Q	charged to polar	E3 interface
	V264A	semi-conservative	Mxra8 interface
	T312M	polar to hydrophobic	surface exposed
	S375T*	conservative	surface exposed
	V386A*	semi-conservative	surface exposed
E3	V39I	conservative	E2 interface
	P56S	cyclic to polar	surface exposed

Figure 3







D

protein	mutation	conservation	location/contacts
Nsp3	M59T	hydrophobic to polar	surface exposed
	Y217H	semi-conservative	surface exposed
	T337I	polar to hydrophobic	surface exposed
	T338M	polar to hydrophobic	surface exposed
	D372E	conservative	surface exposed
	L461P	hydrophobic to cyclic	surface exposed
	P471S	cyclic to polar	surface exposed
6K	V8I	conservative	surface exposed
CP	P23S	cyclic to polar	surface exposed
	V27I	conservative	surface exposed
	K73R	conservative	surface exposed
	N79S	conservative	surface exposed

Figure 4