GLOBAL ANALYSIS OF NON-SYNONYMOUS SNPs IN MONKEYPOX VIRUS

Oluwaseyi Abraham Olawale, Oguntomi Ayomide Samuel, Zaimat Sanni, Precious Oyeboade, Shabana Bibi, Maryam Idrees, Meenakshi Agarwal

Oluwaseyi Abraham Olawale, Gen'Omics Research Hub (GENOMAC HUB), Ogbomoso, Nigeria
Oguntomi Ayomide Samuel, Gen'Omics Research Hub (GENOMAC HUB), Ogbomoso, Nigeria
Zaimat Sanni, Department of Medical Laboratory Science, Ladoke Akintola University of Technology, Ogbomoso, Nigeria.
Precious Oyeboade, Department of Physiology, Ladoke Akintola University of Technology, Ogbomoso, Nigeria.
Shabana Bibi, Department of Biosciences, Shifa Tameer-e-Millat University, Islamabad 44000, Pakistan
Maryam Idrees, Department of Microbiology, Quaid-i-Azam University, Islamabad, Pakistan.
Meenakshi Agarwal, Florida A&M Universit, Tallahassee, Fl, USA.
Correspondence Author: Oluwaseyi Olawale.

Keywords: Non-synonymous SNPs, Whole genome sequence SNP analysis, Monkeypox Virus (MPXV).

ABSTRACT

In May of 2022, it was determined that a virus known as monkeypox (MPX) had spread over the world. The MPXV strain responsible for 87000 illnesses and 112 fatalities has been identified. This strain had a global average of 38.35 mutations per virus sample, with the 2022 epidemic seeing the highest MPXV mutation rate. Our aim was to use publicly accessible NCBI (National Center for Biotechnology Information) data to explore the genetic diversity and evolutionary history of monkeypox viruses (MPXVs).

In-depth comparisons of the genomic characteristics of 376 publicly available complete MPXV genome sequences were performed. Nucleotide sequence variation, gene order, mapping, alignment, and mapping accuracy were among these features. A large amount of sequence conservation was discovered in the MPXV genomes by the meta-analysis, with only a few minor changes between isolates. Multiple genomic regions were discovered to be highly conserved among all MPXV strains, suggesting an essential function in viral replication and persistence. The use of phylogenetic analyses to trace the evolutionary relationships of MPXV strains revealed that the virus has diverged into several subgroups with distinct genetic makeups.

Conclusion: Our findings showed that non-synonymous single nucleotide polymorphisms may be impacting the evolution of MPXV isolates that are dispersed globally and promoting improved adaptability and human-to-human transmissibility.

1. INTRODUCTION

Poxviridae viruses are oval-shaped viruses with large double-stranded DNA genomes and the ability to replicate in the cytoplasm of host cells rather than the nucleus (1). Poxvirus infections typically result in skin nodules, lesions, or a widespread rash (2). The orthopoxvirus genus includes closely related poxviruses such as monkeypox virus (MPXV), variola virus (VARV), vaccinia virus (VACV), and cowpox virus (CPV) (3). As a result, MPXV is defined as a zoonotic orthopox DNA virus that is the causative agent of monkeypox illness (MPX), which was first described in the Democratic Republic of the Congo (DRC) (previously known as Zaire) in 1970 (5).

MPX causes a variety of signs and symptoms, ranging from dermatologic to systemic issues. Fever is usually the initial sign of infection, followed by the growth of many papular, vesiculopustular, and ulcerative lesions on the face and body, as well as a detectable lymphadenopathy. Encephalitis, pneumonitis, keratitis, and subsequent bacterial infections are the most prevalent problems linked with MPX (6, 7). Close contact with an infected animal (rodents are considered to be the principal animal reservoir for human transmission), large respiratory droplets, close or direct touch with skin lesions, or possibly contaminated fomites can all result in MPXV transmission (8). Because orthopoxviruses are serologically cross-reactive, polymerase chain reaction (PCR) is favored over antigen and antibody detection methods. The European Medicines Agency (EMA) approved tecovirimat, a smallpox (VARV) antiviral drug, for the treatment of MPX in 2022, although it is not yet widely available. Smallpox immunizations provide 85% protection against MPXV, and previously vaccinated persons may experience milder symptoms. An improved two-dose vaccination based on vaccinia virus was approved for the prevention of MPX in 2019, but it is not yet widely available (6).

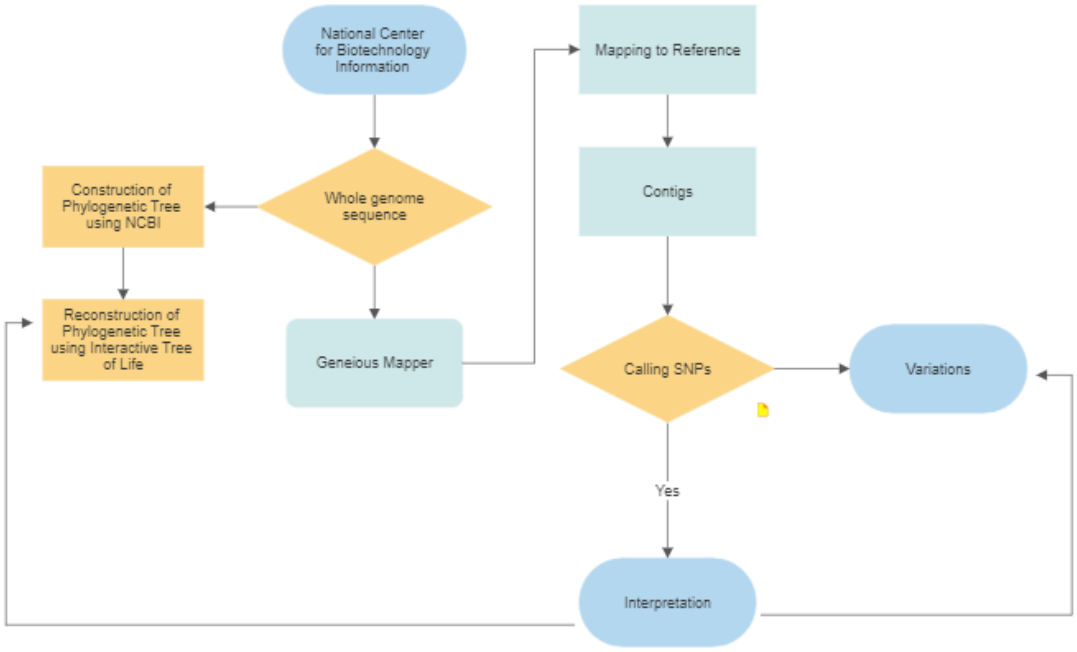
MPXV is becoming more prevalent in the human population, which could lead to genetic adaptation for effective transmissibility and pathogenicity (5). From January 1 to June 22, 2022, 3413 laboratories confirmed MPXV cases and one fatality were reported from five WHO Regions comprising 50 countries/territories. WHO The European Region had the highest number of laboratory confirmed cases (86%; 2933/3413), followed by the Americas Region (11%; 381/3413), Africa (2%; 73/3413), the Eastern Mediterranean Region (1%; 15/3413), and the Western Pacific Region (1%; 11/3413).In Nigeria, one fatality was recorded in the second quarter of 2022 (9).

Several variables have contributed to the success of the MPXV 2022 outbreak. The successful eradication of VARV (smallpox virus) under the auspices of WHO resulted in the discontinuation of VARV vaccine, resulting in an increase in the population vulnerable to VARV and other phylogenetically related orthopoxviruses (10). In addition to the immunological and ecological niche vacated by VARV (11), suspected MPXV genome variations, relaxation in preventive measures against corona virus disease 2019 (covid-2019), restoration of international travel, and sexual interactions associated with large gatherings (12) may have aided MPXV reemergence and spread. Thornhill et al. investigated a case series of MPXV infection across 5 continents and 16 countries from April to June 2022 and observed that 98% of the infected people were gay or bisexual men, 75% were Caucasians, and 41% had an accompanying human immunodeficiency virus (HIV) infection. The average age of those afflicted was 38 years (13).

MPXV is genetically divided into two clades (14): West Africa (WA) and Congo Basin (CB); whole genome study found that the WA and CB lineages are separated by a 900 bp difference. MPXV has been shown to gain or lose genetic material in studies (11); nevertheless, based on the stability of the double-stranded DNA genome, MPXV should have a low mutation rate (15). Wherever possible, MPXV DNA from the present outbreak is being sequenced; preliminary data from polymerase chain reaction (PCR) assays of MPXV isolates from the current outbreak suggest the presence of genes associated with the WA clade (9). For the first time, the spread of MPXV has spread outside of West and Central Africa; it has been proposed that the successive loss of genetic factors unimportant for human pathogenesis served as a mechanism that resulted in the emergence of highly adapted MPXV with increased virulence and the ability to transmit efficiently and rapidly from person to person. As a result, the genetic diversity between the freshly sequenced MPXV isolates and the Zaire-96-I-6 strain is reported in this paper. The goal of this study is to investigate potential variants, anticipate impacts, and conduct phylogenetic analysis.

2. MATERIALS AND METHODS:

2.1 WORKFLOW



2.2 DATA

The SNP analysis was performed using whole genome sequences of MPXV isolates. WGS helps to detect copy number variations, rearrangement and structural variants. For variant analysis, Whole genome sequences of 376 MPXV isolates that were submitted to the NCBI (National Center for Biotechnology Information) from May to July 2022 were retrieved [16].

2.3 SNPs Analysis, Mapping and Alignment

In this study, SNPs analysis was carried out on the mutation of Monkeypox virus in relation to the current Global Pandemic of Monkeypox virus. Using Clade 1 as isolate which is monkeypox virus gotten from Zaire-96-I-16[15]. Whole genome sequences were mapped against Zaire-96-I-16 complete genome as reference sequence using geneious mapper v9.0.5 because it produces superior results and can correctly align structural variants [17], with medium sensitivity, specifying the time it takes to assemble and the accuracy of the assembly to create contig. Protein coding prediction was carried out using EMBOSS 6.5.7, tool, tcode on geneious software v9.0.5

2.4 SNPs Calling and Effect Prediction

SNPs are possible changes that, occurs between the sample and reference sequence. Using the Geneious Software v9.0.5 [17], we imported the downloaded WGS Sequences and conducted the alignment of 376 sequences using clade 1 as the reference (Zaire-96-I-16). The variant

calling generated for all samples used the geneious mapper v9.0.5 with a Minimum variant frequency of 0.6 and a Maximum Variant P-value of 10,000,000,000, a Minimum Strand-Bias P-value of 0.000001, when exceeding 65% bias value to determine the certainty of each association to the reference sequence.

2.5 PHYLOGENETIC ANALYSIS

Study of geographical pattern to determine the pattern of mutation and the common predecessor of each strain of all 2022 monkeypox virus strains reveals various genetic profiles and that random mutations occurred during the evolution process. Active surveillance has obtained monkeypox disease prevalence and topographical reach but there is no available information about virus diversity. Comparative analysis of strains was performed on NCBI based on statistical analysis to determine positions of important differentiation between the strains. We retrieved 376 whole genome sequences of humans from May to July of 2022 on the National Center of Biotechnology Information (NCBI) and edited using Interactive Tree Of Life (ITOL) We obtained 13 clades to discover the common predecessor of each strain of all 2022 monkeypox virus strains.

3 RESULTS

3.1 GENES

According to data on the monkeypox virus, the B17R mutation has been shown to occur multiple times, while the N4R, B11R, J2L, F8L, and B18R mutations have each been found to occur only once. Mutations point to the virus's advanced evolutionary state and host adhesion. J2L, B17R, B18R, B11R, F8L, A27L, and N4R are only some of the genes that can be altered. There are seven separate instances of variance in the B17R gene, but only one instance in each of the other six genes. that's a total of 13 distinct permutations. None of the 33 sequences had the same length. There are 27 one-sequences, 3 two-sequences, and 2 four-sequences. The CDS opinions diverge. The CDS locations of 20 sequences are absent, while those of the remaining 13 sequences range widely. This study aimed to determine the structural and functional impact of missense mutations on the N4R gene.

3.2 SNPs Analysis

On the basis of result similarity, all sequences are 400,483 in length, with an elevated standard of resemblance of 71.4% pair-wise after map to reference. SNPs calling was conducted using the geneious software v9.0.5. While 33 SNPs were obtained; however, 26 SNPs which were synonymous (not causing a change in amino acid) were not considered as it was assumed that synonymous SNPs were not consequential because the primary sequence of the protein is preserved. SNPs were allocated identification based on their positions in the genome of the reference strain. For instance, SNP which locates at position 4567 is named after SNP-4567. Ultimately, each variant at the interval of 1 is separated on the premise of the different positions of sequence nucleotide polymorphism (Transversion, Transition), genes and the effect on protein structure. Variant analysis of MPXV isolates revealed missense mutations; a missense mutation involves a point mutation in which a single nucleotide change results in a codon that encodes a different amino acid, as shown in Figure 1a and 1b.

3.3 Identification of nsSNPs on a Coding Region of Protein

We utilized the geneious software to conduct joint SNPs analysis to explore variants. The SNPs which have no effect on amino acid were filtered out and 7 non synonymous SNPs (alteration of amino acid) were found in the coding region. The nonsynonymous single nucleotide polymorphisms hereby referred to as nsSNPs includes, SNP191565, SNP2723, SNP175877, SNP169783, SNP54359, SNP176386, SNP177993. The nsSNPs resulted in amino acid substitution at the coding region.

3.4 SNPs With Effect on Protein Structure

SNP191565, with variant A as reference nucleotide and G as the variant nucleotide (Adenine-Guanine) encodes a codon change (ACA-GCA) resulting to a substitution of the amino acid from Threonine(T) to Alanine(A), with protein id (NP_536615.1), a length of 1 and locus tag (MPXVgp188) at a variant frequency of 61.2% and a variant P-Value 1.6E-47. SNP2723, with variant A as reference nucleotide and T as the variant nucleotide (Adenine-Thymine) encodes a codon change (TTA-ATA) resulting to a substitution of the amino acid from Leucine(L) to Isoleucine(I), with protein id (NP_536429.1), a length of 1 and locus tag (MPXVgp002) at a variant frequency of 75.5% and a variant P-Value 6.9E-61. SNP175877, with variant G, as

reference nucleotide and A as the variant nucleotide (Guanine-Adenine) encodes a codon change (GGA-GAA) which results to a substitution of the amino acid from Glycine(G) to Glutamate(E), with protein id (NP_536605.1), a length of 1 and locus tag (MPXVgp178) at a variant frequency of 61.1% and a variant P-Value 4.2E-52. SNP169783, with variant G as reference nucleotide and A as the variant nucleotide (G-A) encodes a codon change (GAA-AAA) resulting to substitution of the amino acid from Glutamate(E) to Lysine (K), with protein id (NP_536599.1), a length of 1 and locus tag (MPXVgp172) at a variant frequency of 79.6% and a variant P-Value 8.6E-76. SNP54359, with variant C as reference nucleotide and T as the variant nucleotide (Cytosine-Thymine) encodes a codon change (GAT-AAT) resulting to a substitution of the amino acid from Aspartate(D) to Asparagine(N), with protein id (NP_536484.1), a length of 1 and locus tag (MPXVgp057) at a variant frequency of 84.5% and a variant P-Value 9.7E89. SNP176386, with a variant G, as reference nucleotide and A as variant nucleotide (Guanine-Adenine) encodes a codon change (GAC-AAC) resulting to a substitution of the amino acid from Aspartate (D) to Asparagine (N), with protein id (NP_536605.1), a length of 1 and locus tag (MPXVgp178) at a variant frequency of 61.1% and a variant P-Value 4.2E-52. SNP177993, with a variant C as reference nucleotide and T as variant nucleotide (Cytosine-Thymine) encodes a codon change (GCA-GTA) which results to a substitution of amino acid from Alanine (A) to Valine (V), with protein id (NP_536606.1), a length of 1 and locus tag (MPXVgp179) at a variant frequency of 61.1% and a variant P-Value 4.2E-52.

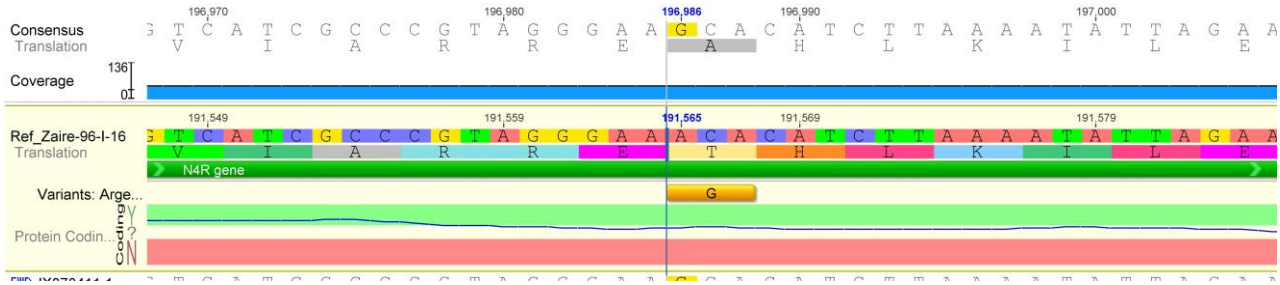
More virulent strains of the virus have been found in isolates from Central Africa (Clade I) compared to Western Africa (Clade II). The differences in amino acids expressed have prompted research into their potential for pathogenicity. The N4R gene has a snp (191565) in which Alanine replaces Threonine. In typical usage, threonine is categorized as a polar hydrophilic (water-loving) amino acid. Protein structure and development rely on its ability to establish hydrogen bonds with a wide range of polar substrates (other molecules). Threonine is a versatile amino acid that can be found either within or on the outside of a protein. It is rather neutral in respect to snp 191565 on N4R gene, though it has a substitution preference with different small or polar amino acids. Alanine is non polar and not hydrophobic.it consists of a normal C-beta carbon and it is present in almost all non-critical protein contexts. Due to the differences in function on the protein structure, it is presumed that the change in amino acid causes snp 191565 on N4R gene to be less virulent. Isoleucine is substituted for Leucine on snp 2723 on J2L gene. Leucine is aliphatic and hydrophobic. It prefers to be entrenched in

protein hydrophobic cores. It also exhibits a fondness for being inside alpha helices more so than in beta strands. Isoleucine prefers to be entrenched in protein hydrophobic cores and it is more laborious for this amino acid to acquire an alpha helical configuration, although it is easy and even preferred for it to stay within beta sheets. Due to the similarity in function of both amino acids on the protein structure, it is presumed that the effect on amino acid causes monkeypox virus to remain the same. Lysine is substituted for glutamate on snp 175877 on BI7R gene Glutamate is charged and polar. It generally prefers to be on the protein's surface, bared to aqueous surroundings. When immersed in the protein, glutamates are usually associated in salt bridges, where they match with a positively charged amino acid to produce stabilizing hydrogen bonds, that can be crucial for stability of protein. Lysine is considered a bit amphipathic (part hydrophilic, part hydrophobic) and are usually associated in salt bridges, where they match with a negatively charged amino acid to produce stabilizing hydrogen bonds that can be crucial for protein stability. Due to the similar function of these amino acids on protein structure, it is presumed that the effect on amin acid remains the same. Asparagine is substituted for Aspartate on snp 176386 on B17R gene. Aspartate is charged and polar and favors being on the protein surface, bared to aqueous surroundings. When immersed within the protein. Aspartates are usually involved in salt bridges, where they match with a positively charged amino acid to produce stabilizing hydrogen bonds, that can be crucial for stability of protein. Asparagine is a polar amino acid which would rather be on the surface of proteins bared to an aqueous environment. Due to the similarity in function of these amino acids on protein structure, it is presumed that the effect on amin acid remains the same. Alanine is non-polar and not hydrophobic. Nonetheless, it has a normal C-beta carbon, which means it is as hampered as other amino acids with consideration to the conformations that the backbone can acquire. Alanine is substituted by Valine on snp 177993 on B18R gene. Valine is hydrophobic, and prefers to be entrenched in protein hydrophobic cores. Valine contains two non-hydrogen substituents affixed to their C-beta carbon. it is more strenuous for valine to acquire an alpha helical conformation, although it is uncomplicated and even preferred for it to stay within beta sheets. While Alanine and Valine differs in function on protein structure, there is no change to the effect on the amino acid and it remains the same. Glutamate is substituted for glycine on snp 175877 on B17R gene. Glycine is a special amino acid because it contains a hydrogen as its side chain in place of carbon as is the case in other amino acids. This shows that there is considerably more conformational flexibility in glycine. This entails that glycine can inhabit parts of protein structures that are prohibited to all other amino acids. Glutamate is charged and polar. It generally prefers to be on the protein's surface, bared to aqueous surroundings. When

© GSJ

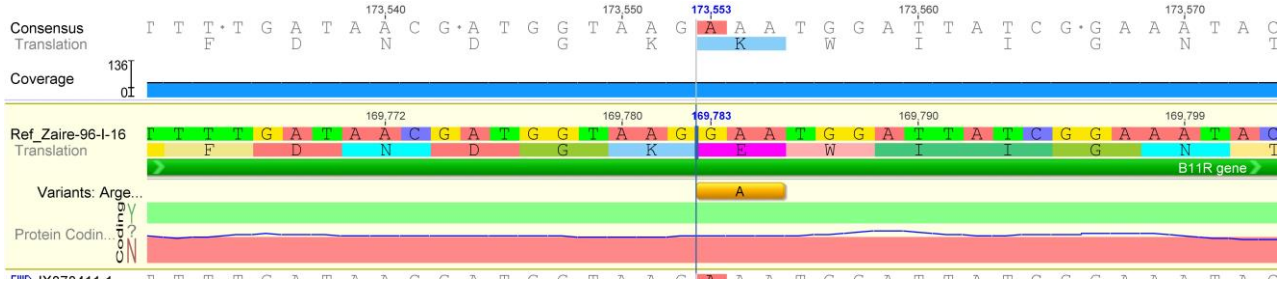
immersed in the protein, glutamates are usually associated in salt bridges, where they match with a positively charged amino acid to produce stabilizing hydrogen bonds, that can be crucial for stability of protein. Due to the similar function of these amino acids on protein structure, it is presumed that the effect on amin acid remains the same.

1a.



SNP 191,565-SNP 191,567 resulted in amino acid substitution from Threonine to Alanine (T- A) on the coding region of the N4R gene.

1b.



SNP 169783-SNP 169785 resulted in amino acid substitution from Glutamate to Lysine (E-K) on the coding region of the B11R gene.

MISSENSE MUTATION

Position	Reference	Outbreak2022	Type	Ftype	Nucleotide Change	Amino Acid Change	Effect	Protein id	Variant Value	P-Gene	Locus tag	Product
2723	A	T	snp	CDS	A-T	L-I	Non-synonymous Variant	NP_536429.1	6.90E-61	J2L	MPXVgp002	J2L
191565	A	G	snp	CDS	A-G	T-A	Non-synonymous Variant	NP_536615.1	1.60E-47	N4R	MPXVgp188	Ankyrin-like protein
177993	C	T	snp	CDS	C-T	A-V	Non-synonymous Variant	NP_536606.1	4.20E-52	B18R	MPXVgp179	B18R
176386	G	A	snp	CDS	G-A	D-N	Non-synonymous Variant	NP_536605.1	4.20E-52	B17R	MPXVgp178	B17R
175877	G	A	snp	CDS	G-A	G-E	Missense Mutation	NP_536605.1	4.20E-52	B17R	MPXVgp178	B17R
169783	G	A	snp	CDS	G-A	E-K	Missense Mutation	NP_536599.1	8.60E-76	B11R	MPXVgp172	B11R
54359	C	T	snp	CDS	C-T	D-N	Missense Mutation	NP_536484.1	9.70E-89	F8L	MPXVgp057	DNA Polymerase
177696	T	A	snp	CDS	T-A		Synonymous Variant	NP_536605.1	4.20E-52	B17R	MPXVgp178	B17R
175668	A	T	snp	CDS	A-T		Synonymous Variant	NP_536605.1	4.20E-52	B17R	MPXVgp178	B17R
176526	C	T	snp	CDS	C-T		Synonymous Variant	NP_536605.1	4.20E-52	B17R	MPXVgp178	B17R
176466	G	A	snp	CDS	G-A		Synonymous Variant	NP_536605.1	2.60E-54	B17R	MPXVgp178	B17R
176130	G	A	snp	CDS	G-A		Silent Mutation	NP_536605.1	4.20E-52	B17R	MPXVgp178	B17R
136139	G	A	snp	CDS	G-A		Silent Mutation	NP_536564.1	3.20E-82	A27L	MPXVgp137	A27L
78-79	CC	TT	snp	CDS			Nonsense Mutation		2.20E-38			
47-49	TTT	AAA	snp	CDS			Nonsense Mutation		3.20E-34			
35-37	TTA	GAG	snp	CDS			Nonsense Mutation		7.60E-33			
22-23	TT	GA	snp	CDS			Nonsense Mutation		1.80E-37			
16-17	AC	TT	snp	CDS			Nonsense Mutation		7.60E-33			
3—4	TA	AG	snp	CDS			Nonsense Mutation		2.70E-33			
19049	A	T	snp	CDS			Nonsense Mutation		4.50E-74			
115	C	A	snp	CDS			Nonsense Mutation		8.80E-40			
55	A	T	snp	CDS			Nonsense Mutation		1.70E-36			
45	T	A	snp	CDS			Nonsense Mutation		2.00E-32			
40	T	A	snp	CDS			Nonsense Mutation		5.90E-37			

26	A	C	snp	CDS	Nonsense Mutation	7.60E-33
8	A	T	snp	CDS	Nonsense Mutation	5.40E-38
178501	G	A	snp	CDS	Nonsense Mutation	4.20E-52
173997	C	T	snp	CDS	Nonsense Mutation	4.20E-52
102	G	A	snp	CDS	Nonsense Mutation	8.80E-40
44	C	T	snp	CDS	Nonsense Mutation	2.00E-32
20	A	G	snp	CDS	Nonsense Mutation	7.60E-33
1	G	A	snp	CDS	Nonsense Mutation	5.60E-43
2798	A		snp	CDS	Nonsense Mutation	5.50E-15

Figure 2 shows the table above which is the result of the variants called which include the Position, Reference nucleotides, Variant nucleotides(Outbreak2022), Type, Ftype, Nucleotide change, Amino acid change, Effect, Protein ID, Variant P-value, Gene(J2L,B17R,F8R,A27L,N4R),Locus tag and the products.

© GSJ

3.4 PHYLOGENETIC ANALYSIS

Phylogenetic analysis of all studied sample sequences inclusive of the reference sample was constructed using the National Center for Biotechnology Information (NCBI) [18] and edited using the Interactive Tree of Life (ITOL) [19] to discover the common predecessor of each strain of all 2022 monkeypox virus strains. Comparative analysis of strains was performed on NCBI based on statistical analysis to determine positions of important differentiation between the strains. For each monkeypox virus sequence, we indicated the country, accession number and the date. The result shows that the sequence data generated exhibits various genetic profiles and that random mutations occurred during the evolution process. By applying the clade function, the samples analyzed shows variations going round the tree. The main clades were highlighted with different colours. A clade is a group of organisms that comprises of a similar ancestor and all its lineal descendants. The clades are embedded in each other, as each branch cleave into smaller branches. Of the 372 sequences studied, phylogenesis differentiated the sequences into 12 different clades separated by different colours. The first clade is denoted by the Colour red and has 29 sequences which includes,10 strains of Portugal,4 strains of Germany, a strain of Italy,3 strains of spain,5 strains of Canada,4 strains of UK, a strain of Switzerland, and a strain of Netherland all separated by their collection dates. The Second clade is represented by the colour purple and has 6 strains of Germany separated by their collection dates. The third clade is denoted by the colour brown and has a total of 27 sequences in the clad of which 21 strains are from Germany,2 strains from Portugal, a strain from Spain, a strain from Switzerland, a strain from USA, and a strain from Canada separated by their collection dates. The fourth clade is represented by a light green colour and has 8 strains of Germany separated by their collection dates. The fifth clade is denoted by a light blue colour and has 10 sequences in total, of which 4 strains are from Germany,4 strains from Canada, a strain from USA and a strain from Czech Republic all separated by their collection dates. The sixth clade is represented by the colour yellow and has 3 strains from Belgium separated by their collection date. The seventh clade is denoted by the dark green colour and has 3 sequences of which 2 strains are from Germany and a strain from Italy separated by their collection date. The eighth clade is represented by a light purple colour and has 3 sequences of which a strain is from Argentina and 2 strains from Argentina separated by their date of collection. The ninth clade has 2 sequences of which 1 is from Israel and another from USA.it is represented by the colour light blue separated by their collection date. The tenth clade is denoted by the dark green colour and has a total of 4

sequences of which 2 strains are from the UK, a strain from Argentina, and a strain from Australia separated by their date of collection. The eleventh clade is represented by a deep blue colour which has a strain of Slovenia as well as the date collection. The twelfth clade has a total of 264 sequences all separated by their collection dates.6 strains are from Slovenia,111 strains from Germany,33 strains from Portugal,61 strains from Canada,5 strains from Italy, a strain from Argentina, a strain from Australia,7 strains from UK,4 strains from spain,3 strains from Switzerland,4 strains from Belgium,2 strains from Brazil,4 strains from France, a strain from Taiwan, a strain from Mexico, a strain from Sweden,14 strains from USA, a strain from Ireland and 4 strains from Finland all denoted by the light blue colour.[19]

© GSJ

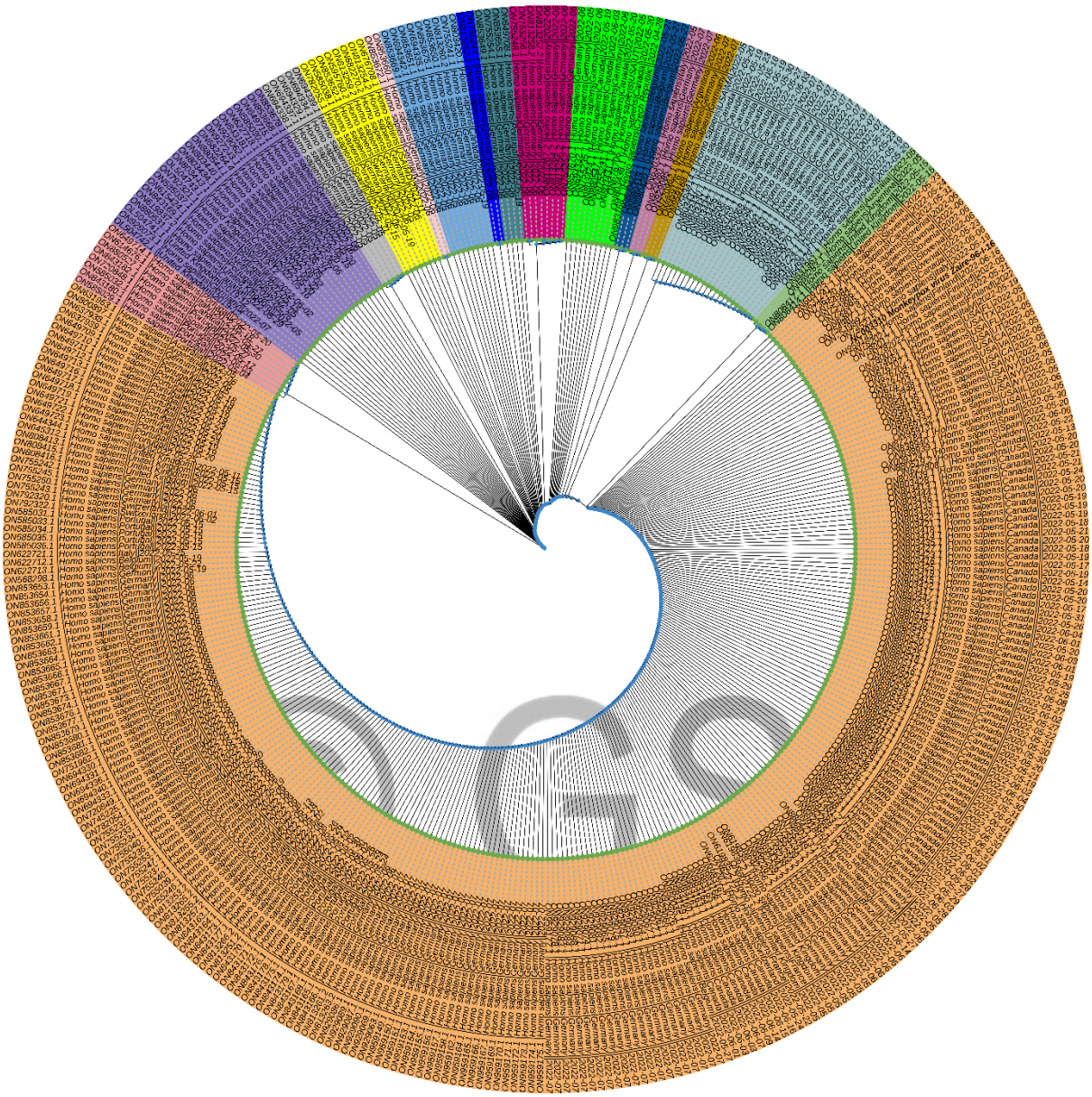


Figure 6: Phylogenetic tree.

CONCLUSION

Because of the dramatic increase in reported cases, researchers now have access to additional genomes that could provide light on population structure, most notably the potential for multiple introductions of monkeypox virus into the human population. The quantification of variants for monkeypox virus isolates reveals an active strategy for change, but long-term surveillance is required. The increase in the frequency of variants being transferred and introduced from humans to humans may be due to a number of reasons, including inoculation status and the invasion of Human land. Non-human primates and rodents are

prohibited entry due to various countries' strict import regulations. Using data from 376 samples of monkeypox virus in 2022, this study reveals any differences that may have occurred between the sample and reference sequence. For all samples with the identical settings, the variant calling produced around 33 SNPs. We analyzed the effects of polymorphism on protein structure, the types of changes that occurred in amino acids and their implications for the protein, the changes that each codon encodes, the frequency of variations, and the locus tags for each sequence. We went into great detail on the genes affected by variants, the products of those genes, the lengths of the sequences, and the CDS places where changes were found. The collected sequencing data displays a wide range of genetic profiles, suggesting that mutations happened at random during the evolutionary process. Phylogenetic tree construction shows 12 major clades distinguished by colors, each containing creatures with a shared ancestor and their direct genetic descendants. Each clade's branch has split off into subbranches, creating a tree structure. Our research reveals that the evolution of newly emerging strains of monkeypox virus is being influenced by a combination of genomic instability and genetic polymorphism.

Acknowledgement

We appreciate GENOMAC HUB, for the impact via research capacity building with funding for the project we carried out. Our appreciation also goes to the leading reseaercher in the person of Mr. Oluwaseyi Olawale for his kind support at all times.

© GSJ

REFERENCES

1. Moss B. Poxviridae: the viruses and their replication, vol. 2. Knipe, DMl. 2001.
2. Prevention CfDCa. Poxvirus 2020 [Available from: https://www.cdc.gov/poxvirus/index.html#:~:text=Poxviruses%20are%20brick%20or%20oval,skin%20nodules%2C%20or%20disseminated%20rash.
3. Fenner F, Wittek R, Dumbell K. Academic Press; San Diego, CA: 1989. The orthopoxviruses[Google Scholar].
4. Shchelkunov SN, Marennikova SS, Moyer RW. Orthopoxviruses pathogenic for humans: Springer Science & Business Media; 2006.
5. Ladnyj I, Ziegler P, Kima E. A human infection caused by monkeypox virus in Basankusu Territory, Democratic Republic of the Congo. Bulletin of the World Health Organization. 1972;46(5):593.

6. Organization WH. Monkeypox 19 May 2022 [Available from: https://www.who.int/news-room/fact-sheets/detail/monkeypox.

7. Shafaati M, Zandi M. Monkeypox virus neurological manifestations in comparison to other orthopoxviruses. Travel Medicine and Infectious Disease. 2022;49:102414.

8. UK Health Security Agency. Monkeypox: background information. [Available from: https://www.gov.uk/guidance/monkeypox#transmission.

9. Organization WH. Multi-country monkeypox outbreak: situation update 27 June 2022 [Available from: https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON396.

10. Shchelkunov SN, Totmenin AV, Babkin IV, Safronov PF, Ryazankina OI, Petrov NA, et al. Human monkeypox and smallpox viruses: genomic comparison. FEBS Lett. 2001;509(1):66-70.

11. Alakunle E, Moens U, Nchinda G, Okeke MI. Monkeypox virus in Nigeria: infection biology, epidemiology, and evolution. Viruses. 2020;12(11):1257.

12. Organization WH. 2022 Monkeypox outbreak global map 2022 [Available from: https://www.cdc.gov/poxvirus/monkeypox/response/2022/world-map.html.

13. Thornhill JP, Barkati S, Walmsley S, Rockstroh J, Antinori A, Harrison LB, et al. Monkeypox virus infection in humans across 16 countries—April–June 2022. New England Journal of Medicine. 2022.

14. Likos AM, Sammons SA, Olson VA, Frace AM, Li Y, Olsen-Rasmussen M, et al. A tale of two clades: monkeypox viruses. Journal of General Virology. 2005;86(10):2661-72.

15. Wang L, Shang J, Weng S, Aliyari SR, Ji C, Cheng G, et al. Genomic annotation and molecular evolution of monkeypox virus outbreak in 2022. Journal of Medical Virology. 2022;n/a(n/a).

16. *"WHO Director-General declares the ongoing monkeypox outbreak a Public Health Emergency of International Concern". World Health Organization (WHO). 23 July 2022. Retrieved 4 August 2022*

17. Geneious(https://www.geneious.com/)

18. PubMed NCBI VIRUS [http://www.ncbi.nlm.nih.gov/virus]

19. Interactive Tree of Life (ITOL)

© GSJ

Supplement materials 1: Summary of monkeypox whole genome sequence isolate investigated. Included are; Accession Number and Demographics

© GSJ

Accession Number	Sample Name	Data Source	Location	Collection Date
ON800897	ARG003	NCBI	Argentina: Ciuda	2022-06
ON720962	MPV-ARG002	NCBI	Argentina: Ciuda	2022-06
ON669283	MPV-ARG001	NCBI	Argentina: Bueno	2022-05
ON631961	VIDRL01	NCBI	Australia	2022-05
ON631963	VIDRL01	NCBI	Australia	2022-05
ON950045	ITM_PT_31	NCBI	Belgium	2022-05
ON880419	U2_REGA_3	NCBI	Belgium	2022-05
ON880420	U2_REGA_4	NCBI	Belgium	2022-05
ON880421	U2_REGA_5	NCBI	Belgium	2022-05
ON880422	U2_REGA_6	NCBI	Belgium	2022-05
ON622712	U2-REGA-1	NCBI	Belgium	2022-05
ON622713	U2-REGA-2	NCBI	Belgium	2022-05
ON880413	BR0002	NCBI	Brazil	2022-06
ON751962	BR0001	NCBI	Brazil	2022-06
0N983159	NML-3293	NCBI	Canada	2022-06
ON983160	NML-3393	NCBI	Canada	2022-06
ON983161	NML-3530	NCBI	Canada	2022-06
ON983162	NML-3531	NCBI	Canada	2022-06
ON983163	NML-3532	NCBI	Canada	2022-06
ON983164	NML-3558	NCBI	Canada	2022-06
ON983165	NML-3686	NCBI	Canada	2022-06
ON983166	NML-3702	NCBI	Canada	2022-06
ON983167	NML-3725	NCBI	Canada	2022-06
ON880505	NML-2906	NCBI	Canada	2022-05
ON880506	NML-2959	NCBI	Canada	2022-05
ON880507	NML-3071	NCBI	Canada	2022-05
ON880508	NML-3127	NCBI	Canada	2022-05
ON880509	NML-3137	NCBI	Canada	2022-05

ON880510	NML-3193	NCBI	Canada	2022-05
ON88O511	NML-3214	NCBI	Canada	2022-05
ON880512	NML-3292	NCBI	Canada	2022-06
ON880513	NML-3407	NCBI	Canada	2022-06
ON880514	NML-3412	NCBI	Canada	2022-06
ON880515	NML-3467	NCBI	Canada	2022-06
ON880516	NML-3472	NCBI	Canada	2022-06
ON880517	NML-3476	NCBI	Canada	2022-05
ON880518	NML-3486	NCBI	Canada	2022-06
ON880519	NML-2833	NCBI	Canada	2022-05
ON880520	NML-2857	NCBI	Canada	2022-05
ON880521	NML-2858	NCBI	Canada	2022-05
ON880522	NML-2870	NCBI	Canada	2022-05
ON880523	NML-2898	NCBI	Canada	2022-05
ON880524	NML-3080	NCBI	Canada	2022-05
ON880525	NML-3086	NCBI	Canada	2022-05
ON880526	NML-3100	NCBI	Canada	2022-05
ON880527	NML-3141	NCBI	Canada	2022-05
ON880528	NML-3147	NCBI	Canada	2022-05
ON880529	NML-3153	NCBI	Canada	2022-05
ON880530	NML-3162	NCBI	Canada	2022-05
ON880531	NML-3165	NCBI	Canada	2022-05
ON880532	NML-3253	NCBI	Canada	2022-05
ON880533	NML-3257	NCBI	Canada	2022-05
ON880534	NML-3262	NCBI	Canada	2022-05
ON880535	NML-3269	NCBI	Canada	2022-05
ON880536	NML-3272	NCBI	Canada	2022-05
ON880537	NML-3273	NCBI	Canada	2022-05
ON880538	NML-3336	NCBI	Canada	2022-06
ON880539	NML-3339	NCBI	Canada	2022-06
ON880540	NML-3342	NCBI	Canada	2022-06

ON880541	NML-3344	NCBI	Canada	2022-06
ON880542	NML-3348	NCBI	Canada	2022-05
ON880543	NML-3357	NCBI	Canada	2022-06
ON880544	NML-3363	NCBI	Canada	2022-06
ON880545	NML-3366	NCBI	Canada	2022-06
ON880546	NML-3368	NCBI	Canada	2022-06
ON880547	NML-3441	NCBI	Canada	2022-06
ON880548	NML-3448	NCBI	Canada	2022-06
ON880549	NML-3450	NCBI	Canada	2022-06
ON803413	NML-2928	NCBI	Canada	2022-05
ON803414	NML-2930	NCBI	Canada	2022-05
ON803415	NML-2936	NCBI	Canada	2022-05
ON803416	NML-2835	NCBI	Canada	2022-05
ON803417	NML-2836	NCBI	Canada	2022-05
ON803418	NML-2837	NCBI	Canada	2022-05
ON803419	NML-2844	NCBI	Canada	2022-05
ON803420	NML-2846	NCBI	Canada	2022-05
ON803421	NML-2849	NCBI	Canada	2022-05
ON803422	NML-2850	NCBI	Canada	2022-05
ON803423	NML-2851	NCBI	Canada	2022-05
ON803424	NML-2852	NCBI	Canada	2022-05
ON803425	NML-2853	NCBI	Canada	2022-05
ON803426	NML-2856	NCBI	Canada	2022-05
ON803427	NML-2861	NCBI	Canada	2022-05
ON803428	NML-2869	NCBI	Canada	2022-05
ON803429	NML-2873	NCBI	Canada	2022-05
ON803430	NML-2874	NCBI	Canada	2022-05
ON803431	NML-2883	NCBI	Canada	2022-05
ON803432	NML-2884	NCBI	Canada	2022-05
ON803433	NML-2885	NCBI	Canada	2022-05
ON803434	NML-2886	NCBI	Canada	2022-05

ON803435	NML-2893	NCBI	Canada	2022-05
ON803436	NML-2894	NCBI	Canada	2022-05
ON803437	NML-2896	NCBI	Canada	2022-05
ON803438	NML-2897	NCBI	Canada	2022-05
ON803439	NML-2899	NCBI	Canada	2022-05
ON803440	NML-2954	NCBI	Canada	2022-05
ON803441	NML-2955	NCBI	Canada	2022-05
ON803442	NML-2967	NCBI	Canada	2022-05
ON803443	NML-2975	NCBI	Canada	2022-05
ON803444	NML-2969	NCBI	Canada	2022-05
ON736420	AB1	NCBI	Canada	2022-05
ON983168	Prague_001	NCBI	Czech Republic	2022-04
ON959143	MPX-96	NCBI	Finland	2022-06
ON782021	MPX_37	NCBI	Finland	2022-05
ON782022	MPX_42	NCBI	Finland	2022-05
ON755039	MPXVIRBA22-11	NCBI	France	2022-05
ON755040	MPXVIRBA22-14	NCBI	France	2022-05
ON959149	RKI140	NCBI	Germany	2022-07
ON959150	RKI145	NCBI	Germany	2022-07
ON959151	RKI160	NCBI	Germany	2022-07
ON959152	RKI139	NCBI	Germany	2022-07
ON959153	RKI165	NCBI	Germany	2022-07
ON959154	RKI137	NCBI	Germany	2022-07
ON959155	RKI138	NCBI	Germany	2022-07
ON959156	RKI141	NCBI	Germany	2022-07
ON959157	RKI142	NCBI	Germany	2022-07
ON959158	RKI143	NCBI	Germany	2022-07
ON959159	RKI144	NCBI	Germany	2022-07
ON959160	RKI146	NCBI	Germany	2022-07
ON959161	RKI147	NCBI	Germany	2022-07

ON959162	RKI148	NCBI	Germany	2022-07
ON959163	RKI149	NCBI	Germany	2022-07
ON959164	RKI150	NCBI	Germany	2022-07
ON959165	RKI151	NCBI	Germany	2022-07
ON959166	RKI152	NCBI	Germany	2022-07
ON959167	RKI153	NCBI	Germany	2022-07
ON959168	RKI154	NCBI	Germany	2022-07
ON959169	RKI155	NCBI	Germany	2022-07
ON959170	RKI156	NCBI	Germany	2022-07
ON959171	RKI157	NCBI	Germany	2022-07
ON959172	RKI158	NCBI	Germany	2022-07
ON959173	RKI159	NCBI	Germany	2022-07
ON959174	RKI161	NCBI	Germany	2022-07
ON959175	RKI162	NCBI	Germany	2022-07
ON959176	RKI163	NCBI	Germany	2022-07
ON959177	RKI164	NCBI	Germany	2022-07
ON929057	RKI070	NCBI	Germany	2022-07
ON929058	RKI071	NCBI	Germany	2022-07
ON929059	RKI072	NCBI	Germany	2022-07
ON929060	RKI102	NCBI	Germany	2022-07
ON929061	RKI103	NCBI	Germany	2022-07
ON929062	RKI104	NCBI	Germany	2022-07
ON929063	RKI105	NCBI	Germany	2022-07
ON929064	RKI106	NCBI	Germany	2022-07
ON929065	RKI107	NCBI	Germany	2022-07
ON929066	RKI108	NCBI	Germany	2022-07
ON929067	RKI109	NCBI	Germany	2022-07
ON929068	RKI110	NCBI	Germany	2022-07
ON929069	RKI111	NCBI	Germany	2022-07
ON929070	RKI112	NCBI	Germany	2022-07
ON929071	RKI113	NCBI	Germany	2022-07

ON929072	RKI114	NCBI	Germany	2022-07
ON929073	RKI115	NCBI	Germany	2022-07
ON929074	RKI119	NCBI	Germany	2022-07
ON929075	RKI120	NCBI	Germany	2022-07
ON929076	RKI121	NCBI	Germany	2022-07
ON929077	RKI122	NCBI	Germany	2022-07
ON929078	RKI123	NCBI	Germany	2022-07
ON929079	RKI124	NCBI	Germany	2022-07
ON929080	RKI125	NCBI	Germany	2022-07
ON929081	RKI126	NCBI	Germany	2022-07
ON929082	RKI127	NCBI	Germany	2022-07
ON929083	RKI128	NCBI	Germany	2022-07
ON929084	RKI129	NCBI	Germany	2022-07
ON929085	RKI130	NCBI	Germany	2022-07
ON929086	RKI131	NCBI	Germany	2022-07
ON929087	RKI132	NCBI	Germany	2022-07
ON929088	RKI133	NCBI	Germany	2022-07
ON929089	RKI134	NCBI	Germany	2022-07
ON929090	RKI135	NCBI	Germany	2022-07
ON929091	RKI136	NCBI	Germany	2022-07
ON853649	RKI068	NCBI	Germany	2022-06
ON853650	RKI069	NCBI	Germany	2022-06
ON853651	RKI073	NCBI	Germany	2022-06
ON853652	RKI074	NCBI	Germany	2022-06
ON853653	RKI076	NCBI	Germany	2022-06
ON853654	RKI087	NCBI	Germany	2022-06
ON853655	RKI088	NCBI	Germany	2022-06
ON853656	RKI070	NCBI	Germany	2022-06
ON853657	RKI071	NCBI	Germany	2022-06
ON853658	RKI072	NCBI	Germany	2022-06
ON853659	RKI075	NCBI	Germany	2022-06

ON853660	RKI078	NCBI	Germany	2022-06
ON853661	RKI079	NCBI	Germany	2022-06
ON853662	RKI080	NCBI	Germany	2022-06
ON853663	RKI081	NCBI	Germany	2022-06
ON853664	RKI082	NCBI	Germany	2022-06
ON853665	RKI083	NCBI	Germany	2022-06
ON853666	RKI084	NCBI	Germany	2022-06
ON853667	RKI085	NCBI	Germany	2022-06
ON853668	RKI086	NCBI	Germany	2022-06
ON853669	RKI089	NCBI	Germany	2022-06
ON853670	RKI090	NCBI	Germany	2022-06
ON853671	RKI091	NCBI	Germany	2022-06
ON853672	RKI1093	NCBI	Germany	2022-06
ON853673	RKI1094	NCBI	Germany	2022-06
ON853674	RKI095	NCBI	Germany	2022-06
ON853675	RKI096	NCBI	Germany	2022-06
ON853676	RKI098	NCBI	Germany	2022-06
ON853677	RKI099	NCBI	Germany	2022-06
ON853678	RKI100	NCBI	Germany	2022-06
ON853679	RKI1077	NCBI	Germany	2022-06
ON853680	RKI1097	NCBI	Germany	2022-06
ON853681	RKI1092	NCBI	Germany	2022-06
ON853682	RKI101	NCBI	Germany	2022-06
ON813251	RKI051	NCBI	Germany	2022-06
ON813252	RKI052	NCBI	Germany	2022-06
ON813253	RKI053	NCBI	Germany	2022-06
ON813254	RKI054	NCBI	Germany	2022-06
ON813255	RKI055	NCBI	Germany	2022-06
ON813256	RKI056	NCBI	Germany	2022-06
ON813257	RKI057	NCBI	Germany	2022-06
ON813258	RKI058	NCBI	Germany	2022-06

ON813259	RKI059	NCBI	Germany	2022-06
ON813260	RKI060	NCBI	Germany	2022-06
ON813261	RKI061	NCBI	Germany	2022-06
ON813262	RKI062	NCBI	Germany	2022-06
ON813263	RKI063	NCBI	Germany	2022-06
ON813264	RKI064	NCBI	Germany	2022-06
ON813265	RKI065	NCBI	Germany	2022-06
ON813266	RKI056	NCBI	Germany	2022-06
ON813267	RKI067	NCBI	Germany	2022-06
ON755231	RKI013	NCBI	Germany	2022-06
ON755232	RKI026	NCBI	Germany	2022-06
ON755233	RKI027	NCBI	Germany	2022-06
ON755234	RKI028	NCBI	Germany	2022-06
ON755235	RKI029	NCBI	Germany	2022-06
ON755236	RKI030	NCBI	Germany	2022-06
ON755237	RKI031	NCBI	Germany	2022-06
ON755238	RKI032	NCBI	Germany	2022-06
ON755239	RKI033	NCBI	Germany	2022-06
ON755240	RKI034	NCBI	Germany	2022-06
ON755241	RKI035	NCBI	Germany	2022-06
ON755242	RKI036	NCBI	Germany	2022-06
ON755243	RKI037	NCBI	Germany	2022-06
ON755244	RKI038	NCBI	Germany	2022-06
ON755245	RKI039	NCBI	Germany	2022-06
ON755246	RKI040	NCBI	Germany	2022-06
ON755247	RKI041	NCBI	Germany	2022-06
ON755248	RKI042	NCBI	Germany	2022-06
ON755249	RKI043	NCBI	Germany	2022-06
ON755250	RKI044	NCBI	Germany	2022-06
ON755251	RKI045	NCBI	Germany	2022-06
ON755252	RKI046	NCBI	Germany	2022-06

ON755253	RKI047	NCBI	Germany	2022-06
ON755254	RKI048	NCBI	Germany	2022-06
ON755255	RKI049	NCBI	Germany	2022-06
ON755256	RKI050	NCBI	Germany	2022-06
ON694329	RKI025	NCBI	Germany	2022-05
ON694330	RKI019	NCBI	Germany	2022-05
ON694331	RKI015	NCBI	Germany	2022-05
ON694332	RKI023	NCBI	Germany	2022-05
ON694333	RKI022	NCBI	Germany	2022-05
ON694334	RKI021	NCBI	Germany	2022-05
ON694335	RKI018	NCBI	Germany	2022-05
ON694336	RKI024	NCBI	Germany	2022-05
ON694337	RKI020	NCBI	Germany	2022-05
ON694338	RKI017	NCBI	Germany	2022-05
ON694339	RKI016	NCBI	Germany	2022-05
ON694340	RKI014	NCBI	Germany	2022-05
ON694341	RKI012	NCBI	Germany	2022-05
ON694342	RKI011	NCBI	Germany	2022-05
ON694363	RKI03	NCBI	Germany	2022-05
ON694364	RKI05	NCBI	Germany	2022-05
ON694365	RKI04	NCBI	Germany	2022-05
ON694366	RKI09	NCBI	Germany	2022-05
ON694367	RKI010	NCBI	Germany	2022-05
ON694368	RKI08	NCBI	Germany	2022-05
ON694369	RKI07	NCBI	Germany	2022-05
ON694370	RKI06	NCBI	Germany	2022-05
ON637938	RK01	NCBI	Germany	2022-05
ON637939	RK02	NCBI	Germany	2022-05
ON568298	MPXV-BV-IMB25241	NCBI	Germany	2022-05
ON872184	WA-2022	NCBI	Ireland	2022-05
ON649879	MPXV_ISR001_2022	NCBI	Israel	2022-05

ON780016	INMI-Pt3	NCBI	Italy	2022-05
ON780017	INMI-Pt4	NCBI	Italy	2022-05
ON745215	INMI_Pt2	NCBI	Italy	2022-05
ON630332	INMI_Pt2	NCBI	Italy	2022-05
ON911481	UAWL001	NCBI	Mexico: Nuevo Leon	2022-06
ON615424	MPXV-2022-NL001	NCBI	Netherlands	2022-05
ON843163	PT0048	NCBI	Portugal	2022-05
ON843164	PT0037	NCBI	Portugal	2022-05
ON843165	PT0033	NCBI	Portugal	2022-05
ON843166	PT0038	NCBI	Portugal	2022-05
ON843167	PT0036	NCBI	Portugal	2022-05
ON843168	PT0032	NCBI	Portugal	2022-05
ON843169	PT0034	NCBI	Portugal	2022-05
ON843170	PT0031	NCBI	Portugal	2022-05
ON843171	PT0030	NCBI	Portugal	2022-05
ON843172	PT0035	NCBI	Portugal	2022-05
ON843173	PT0029	NCBI	Portugal	2022-05
ON843174	PT0046	NCBI	Portugal	2022-05
ON843175	PT0043	NCBI	Portugal	2022-05
ON843176	PT0041	NCBI	Portugal	2022-06
ON843177	PT0047	NCBI	Portugal	2022-06
ON843178	PT0040	NCBI	Portugal	2022-06
ON843179	PT0039	NCBI	Portugal	2022-06
ON843180	PT0042	NCBI	Portugal	2022-06
ON843181	PT0045	NCBI	Portugal	2022-06
ON843182	PT0044	NCBI	Portugal	2022-05
ON649708	PT0023	NCBI	Portugal	2022-05
ON649709	PT0028	NCBI	Portugal	2022-05
ON649710	PT0027	NCBI	Portugal	2022-05
ON649711	PT0026	NCBI	Portugal	2022-05

ON649712	PT0025	NCBI	Portugal	2022-05
ON649713	PT0020	NCBI	Portugal	2022-05
ON649714	PT0015	NCBI	Portugal	2022-05
ON649715	PT0017	NCBI	Portugal	2022-05
ON649716	PT0016	NCBI	Portugal	2022-05
ON649717	PT0019	NCBI	Portugal	2022-05
ON649718	PT0018	NCBI	Portugal	2022-05
ON649719	PT0012	NCBI	Portugal	2022-05
ON649720	PT0024	NCBI	Portugal	2022-05
ON649721	PT0021	NCBI	Portugal	2022-05
ON649722	PT0022	NCBI	Portugal	2022-05
ON649723	PT0013	NCBI	Portugal	2022-05
ON649724	PT0014	NCBI	Portugal	2022-05
ON649725	PT0011	NCBI	Portugal	2022-05
ON585029	PT0001	NCBI	Portugal	2022-05
ON585030	PT0002	NCBI	Portugal	2022-05
ON585031	PT0003	NCBI	Portugal	2022-05
ON585032	PT0004	NCBI	Portugal	2022-05
ON585033	PT0006	NCBI	Portugal	2022-05
ON585034	PT0007	NCBI	Portugal	2022-05
ON585035	PT0009	NCBI	Portugal	2022-05
ON585036	PT0010	NCBI	Portugal	2022-05
ON585037	PT0005	NCBI	Portugal	2022-05
ON585038	PT0008	NCBI	Portugal	2022-05-
ON927243	HGUGM-6642566	NCBI	Spain	2022-06
ON838939	8887	NCBI	Spain	2022-06
ON838940	4061	NCBI	Spain	2022-05
ON745225	HGUGM-66649483	NCBI	Spain	2022-05
ON720848	HGUGM-6513479	NCBI	Spain	2022-05
ON720849	HGUGM-6532064	NCBI	Spain	2022-05
ON622718	ES0001	NCBI	Spain	2022-05

ON782054	HUNSC-ITER-0001A	NCBI	Spain	2022-05
ON782055	HUNSC-ITER-0001B	NCBI	Spain	2022-05
OX009124	Chromosome:1	NCBI	Sweden	2022-05
ON838178	SI2022_S7	NCBI	Slovenia	2022-06
ON754984	MPVX-SI2022-S3	NCBI	Slovenia	2022-06
ON754985	MPVX-SI2022-S4	NCBI	Slovenia	2022-06
ON754986	MPVX-SI2022-S5	NCBI	Slovenia	2022-06
ON754987	MPVX-SI-VEROE6	NCBI	Slovenia	2022-05
ON631241	2SLO	NCBI	Slovenia	2022-05
ON622725	SLO	NCBI	Slovenia	2022-05
ON792320	ZH-UZH-IMV-3ba6449f	NCBI	Switzerland: ZH	2022-06
ON792321	ZH-UZH-IMV-3ba6450e	NCBI	Switzerland: ZH	2022-06
ON792322	ZH-UZH-IMV-3ba64538	NCBI	Switzerland: ZH	2022-06
ON622720	MPXV-CH-38156923	NCBI	Switzerland	2022-05
ON595760	MPXV-CH-38134631	NCBI	Switzerland	2022-05
ON918656	110-231642	NCBI	Taiwan	2022-06
OPO19275	MUW1525179	NCBI	Austria	2022-07
OPO19276	MUW1527495	NCBI	Austria	2022-07
OPO19277	MUW1531254	NCBI	Austria	2022-07
ON808413	CVR_MPXV1a	NCBI	United Kingdom	2022-05
ON808414	CVR_MPXV1b	NCBI	United Kingdom	2022-05
ON808415	CVR_MPXV1c	NCBI	United Kingdom	2022-05
ON808416	CVR_MPXV1d	NCBI	United Kingdom	2022-05

ON808417	CVR_MPXV1e	NCBI	United Kingdom	2022-05
ON645312	MPXV_GSTT_Patient1	NCBI	United Kingdom	2022-05
ON619835	MPXV_UK_2022_1	NCBI	United Kingdom	2022-05
ON619836	MPXV_UK_2022_2	NCBI	United Kingdom	2022-05
ON619837	MPXV_UK_2022_3	NCBI	United Kingdom	2022-05
ON619838	MPXV_UK_2022_4	NCBI	United Kingdom	2022-05
ON959131	IL001	NCBI	USA:IL	2022-05
ON959132	IL003	NCBI	USA:IL	2022-06
ON959133	NY002	NCBI	USA: NY	2022-05
ON959134	NY003	NCBI	USA: NY	2022-05
0N959135	NY004	NCBI	USA: NY	2022-05
ON959136	NY005	NCBI	USA: NY	2022-06
ON954773	CA002	NCBI	USA:CA	2022-05
ON674051	MPXV-USA-2022-FL001	NCBI	USA: FL	2022-05
ON675438	MPXV-USA-2022-VA001	NCBI	USA:VA	2022-05
ON676703	MPXV-USA-2022-CA001	NCBI	USA:CA	2022-05
ON676704	MPXV-USA-2022-FL002	NCBI	USA: FL	2022-05
ON676705	MPXV-USA-2022-UT001	NCBI	USA: UT	2022-05
ON676706	MPXV-USA-2022-UT002	NCBI	USA: UT	2022-05

OPO18607	RKI185	NCBI	Germany	2022-07
OPO18606	RKI184	NCBI	Germany	2022-07
ON627808	UT-UPHL-82200022	NCBI	USA	2022-05
ON563414	MPXV-USA-2022-MA001	NCBI	USA:MA	2022-05

Supplement Materials 2;

SNP Analysis with no Effect on Protein

Based on the 33 results gotten from variant analysis between the 387 samples mapped to Zaire-96-I-16 as reference in this study,6 of the variants encodes changes in amino acid but have no effects on amino acid. This is a Silent mutation, which occurs when a change in the DNA sequence of a gene does not affect the subsequent amino acid sequence of the protein it codes for. This is shown in Fig 4. 20 of the variants have no effects on amino acid and therefore do not encode changes. This is a Nonsense mutation, which occurs in DNA when a sequence change gives rise to a stop codon rather than a codon specifying an amino acid. This is shown in Fig 5.

SNP136139, with variant G, as reference nucleotide and A as the variant nucleotide (Guanine-Adenine) encodes a codon change (AGC-AGT) resulting to no effect in the amino acid with protein id (NP_536564.1), a length of 1 and locus tag (MPXVgp137) at a variant frequency of 82.1% and a variant P-Value 3.2E-82. SNP175668, with variant A as reference nucleotide and T as variant nucleotide (Adenine-Thymine) encodes a codon change (ACA-ACT) resulting to no effect in the amino acid with protein id (NP_536605.1), a length of 1 and locus tag (MPXVgp178) at a variant frequency of 61.1% and a variant P-Value (approximate) 4.2E-52. SNP176130, with variant G, as reference nucleotide and A as variant nucleotide (Guanine-Adenine) encodes a codon change (GTG-GTA) resulting to no effect in the amino acid with protein id (NP_536605.1), a length of 1 and locus tag (MPXVgp178) at a variant frequency of 61.1% and a variant P-Value (approximate) 4.2E-52. SNP176466, with variant G, as reference nucleotide and A as variant nucleotide (Guanine-Adenine) encodes a codon change (GTG-GTA) resulting to no effect in the amino acid

with protein id (NP_536605.1), a length of 1 and locus tag (MPXVgp178) at a variant frequency of 61.1% and a variant P-Value (approximate) 2.6E-54. SNP176526, with variant C, as reference nucleotide and T as variant nucleotide (Cytosine-Thymine) encodes a codon change (CAC-CAT) resulting to no effect on the amino acid with protein id (NP_536605.1), a length of 1 and locus tag (MPXVgp178) at a variant frequency of 61.1% and a variant P-Value (approximate) 4.2E-52. SNP177696, with variant T, as reference nucleotide and A as variant nucleotide (Thymine-Adenine) encodes a codon change (GTT-GTA) resulting to no effect in the amino acid with protein id (NP_536605.1), a length of 1 and locus tag (MPXVgp178) at a variant frequency of 61.1% and a variant P-Value (approximate) 4.2E-52.SNP 1 with G, as reference nucleotide and A as variant nucleotide (Guanine-Adenine) encodes no codon change resulting in no effect to amino acid with no protein id, a length of 1 and no locus tag at a variant frequency of 80.0% and a variant P-Value 5.6E-43. SNP 3-4 with TA, as reference nucleotide and AG as variant nucleotide (Thymine Adenine- Adenine Guanine) encodes no codon change resulting in no effect to amino acid with no protein id, a length of 2 and no locus tag at a variant frequency of 66.7% and a variant P-Value 2.7E-33.SNP 8 with A as reference nucleotide and T as the variant nucleotide (Adenine-Thymine)) encodes no codon change resulting in no effect to amino acid with no protein id, a length of 1 and no locus tag at a variant frequency of 73.3% and a variant P-Value 5.4E-38. SNP 16-17 with AC as reference nucleotide and TT as variant nucleotide (Adenine Cytosine-Thymine Thymine) encodes no codon change resulting in no effect to amino acid, no protein id, length of 2 and no locus tag at a variant frequency of 64.5% and a variant P-Value 7.6E-33.SNP 20 with A as reference nucleotide and G, as variant nucleotide (Adenine-Guanine) nucleotide encodes no codon change resulting in no effect to amino acid, no protein id, a length of 1 and no locus tag at a variant frequency of 64.5% and a variant P-Value 7.6E-33. SNP 22-23 with TT as reference nucleotide and GA as variant nucleotide (Thymine Thymine-Guanine Adenine) encodes no codon change resulting in no effect to amino acid, no protein id, length of 2 and no locus tag at a variant frequency of 71.0% and a variant P-Value 1.8E-37. SNP 26 with A as reference nucleotide and C as variant nucleotide (Adenine-Cytosine) encodes no codon change resulting in no effect to amino acid, no protein id, length of 1 and no locus tag at a variant frequency of 64.5% and a variant P-Value 7.6E-33. SNP 35-37 with TTA as reference nucleotide and GAG as variant nucleotide (Thymine Thymine Adenine- Guanine Adenine Guanine) encodes no codon change resulting in no effect to amino acid, no protein id, length of 3 and no locus tag at a variant frequency of 64.5% and a variant

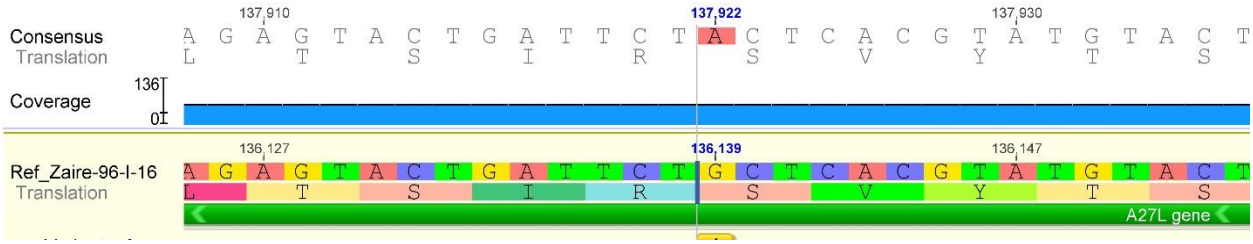
© GSJ

P-Value 7.6E-33. SNP 40 with T as reference nucleotide and A as variant nucleotide (Thymine-Adenine) encodes no codon change resulting in no effect to amino acid and therefore no protein id, a length of 1 and no locus tag at a variant frequency of 68.8% and a variant P-Value 5.9E-37. SNP 44 with C as reference nucleotide and T as variant nucleotide (Cytosine-Thymine) encodes no codon change resulting in no effect to amino acid, no protein id, with a length of 1 and no locus tag at a variant frequency of 62.5% and a variant P-Value 2.0E-32. SNP 45 with T as reference nucleotide and A as variant nucleotide (Thymine-Adenine) encodes no codon change resulting in no effect to amino acid, no protein id, a length of 1 and no locus tag at a variant frequency of 62.5% and a variant P-Value (approximate) 2.0E-32. SNP 47-49 with TTT as reference nucleotide and AAA as variant nucleotide (Thymine Thymine Thymine-Adenine Adenine Adenine) encodes no codon change resulting in no effect to amino acid, no protein id, length of 3 and no locus tag at a variant frequency of 60.6% and a variant P-Value 3.2E-34. SNP 55 with A as reference nucleotide and T as variant nucleotide (Adenine-Thymine) encodes no codon change resulting in no effect to amino acid, no protein id, a length of 1 and no locus tag at a variant frequency of 66.7% and a variant P-Value 1.7E-36. SNP 78-79 with CC as reference nucleotide and TT as variant nucleotide (Cytosine Cytosine-Thymine Thymine) encodes no codon change resulting in no effect to amino acid, no protein id, length of 2 and no locus tag at a variant frequency of 61.5% and a variant P-Value 2.2E-38. SNP 102 with G as reference nucleotide and A as variant nucleotide (Guanine-Adenine) encodes no codon change resulting in no effect to amino acid, no protein id, a length of 1 and no locus tag at a variant frequency of 61.0% and a variant P-Value 8.8E-40. SNP 115 with C as a reference nucleotide and A as variant nucleotide (Cytosine-Adenine) encodes no codon change resulting in no effect to amino acid, no protein id, a length of 1 and no locus tag at a variant frequency of 61.0% and a variant P-Value 8.8E-40. SNP 2798 has a deleted reference nucleotide and A as a variant nucleotide (-Adenine) encodes no codon change resulting in no effect to amino acid, no protein id, a length of 1 and no locus tag at a variant frequency of 64.7% and a variant P-Value 5.5E-15. SNP 19049 with A as a reference nucleotide and T as a variant nucleotide (Adenine-Thymine) encodes no codon change resulting in no effect to amino acid, no protein id, a length of 1 and no locus tag at a variant frequency of 72.1% and a variant P-Value 4.5E-74. SNP 173997 with C as a reference nucleotide and T as a variant nucleotide (Cytosine-Thymine) encodes no codon change resulting in no effect to amino acid, no protein id, a length of 1 and no locus tag at a variant frequency of 61.1% and a variant P-Value 4.2E-52. SNP 178501 with G, as a reference

© GSJ

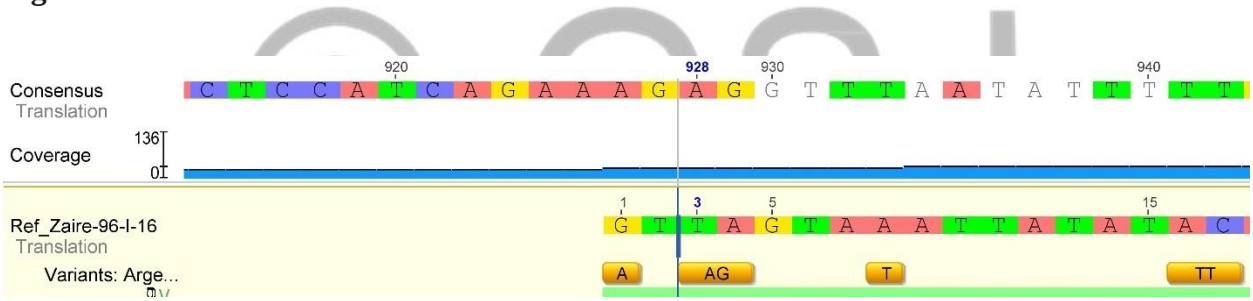
nucleotide and A as a variant nucleotide (Guanine-Adenine) encodes no codon change resulting in no effect to amino acid, no protein id, a length of 1 and no locus tag at a variant frequency of 61.1% and a variant P-Value 4.2E-52.

Figure 4. Silent Mutation



SNP 136139 shows the codon change on the coding region of the A27L gene.

Figure 5 Nonsense Mutation



SNP 3-5 shows no effect in amino acid and shows a stop codon.