Vol. 4 No. 4 (2025)

Pakistan Journal of Medical & Cardiological Review

https://pakjmcr.com/index.php/1/about

Online ISSN 3007-2387

Print ISSN 3007-2379

# GENOMIC SURVEILLANCE: ITS BACKGROUND AND IMPLEMENTATION IN HIV SURVEILLANCE IN PAKISTAN

**Authors:**

1. Shafiq Azam
Email: shafiqazamrumi123@gmail.com
Affiliation: Assistant Scientific Officer, Jhpiego - Johns Hopkins University Affiliate, Pakistan
2. Muhammad Kamran Asif
Email: Kamranasif444@gmail.com
Affiliation: Assistant Scientific Officer, Jhpiego - Johns Hopkins University Affiliate, Pakistan
3. Dr. Sobhan Qadir
Email: Sobhan.Qadir@jhpiego.org
Affiliation: Team Lead EGHS, CDC, Jhpiego - Johns Hopkins University Affiliate, Pakistan
4. Fahad Humayun
Email: fahadhumayun@outlook.com
Affiliation: Bioinformatics Specialist, Jhpiego - Johns Hopkins University Affiliate, Pakistan
5. Kanwal gul
Email: Kanwalgul2005@yahoo.com
Affiliation: Public health Division, National institute of Health
6. Dr. Muhammad Yasir
Email: mohmdyasir@hotmail.com
Affiliation: NGS, Technical Specialist, Jhpiego - Johns Hopkins University Affiliate, Pakistan

**Author Details**

**Keywords:**.

**Received on 10 Aug 2025**
**Accepted on 05 Sep 2025**
**Published on 17 Sep 2025**

**Corresponding E-mail & Author*:**

**Dr. Muhammad Yasir***
Jhpiego (Johns Hopkins University Affiliate) Pakistan.
Email:mohmdyasir@hotmail.com

**Abstract**

Genomic Surveillance is a powerful technique to understand disease outbreaks. Pathogen sequencing has become more accessible and cost-effective than previously, offering a valuable complement to epidemiological approaches. In this review, we describe the advanced techniques employed in sequencing and phylogenetic analysis to trace transmission, and evolutionary patterns of the pathogen. Furthermore, this study also provides an overview of various genomic surveillance studies conducted in Pakistan to explore the genetic diversity, distribution of subtypes, drug resistance patterns, and molecular epidemiology of HIV-1 strains. These studies provide insights into the complexity of HIV subtypes circulating in the country and their distinct genetic characteristics. Key findings include diverse HIV subtypes in Punjab, an outbreak investigation in Larkana revealing CRF02_AG and subtype A1, multiple subtypes in Lahore with drug resistance mutations, genetic links between Pakistan and China, distinct clusters in a Karachi HIV-1 subtype A1 outbreak, a prevalent A1 subtype with resistance mutations in Pakistan, and the discovery of unclassified recombinant forms in children. They emphasized the importance of continuous surveillance, understanding risk factors, and developing targeted prevention and treatment strategies to combat HIV and its potential epidemics in various populations.

## Introduction

### BACKGROUND AND IMPORTANCE OF GENOMIC SURVEILLANCE

Genomic surveillance, a powerful tool in the fight against infectious diseases, involves the constant monitoring of pathogens and analysis of their genetic similarities and variations. This tool uses advancements in genetic sequencing technologies to examine the genetic makeup of pathogens, empowering health authorities and researchers to gain essential knowledge of disease transmission, evolution, and impact (1). By closely evaluating the genetic information of infectious agents such as viruses or bacteria, genomic surveillance facilitates the identification of newly emerging variants, monitoring drug resistance, tracking their transmission patterns, and enabling informed decision-making in public health interventions (2). Moreover, genomic surveillance contributes to the identification of genetic factors that lead to disease, contributing to more precise diagnoses, treatment approaches, and preventive measures. This tool involves collecting positive specimens and sequencing the entire genome to trace the viral lineages over time (3). This retrospective review analysis focuses on the global genomic surveillance of HIV in Pakistan, examining the probable limitations and challenges encountered during the surveillance process.

### HISTORY OF GENOMIC SURVEILLANCE

The history of genomic surveillance can be traced back to the early 1990s when polymerase chain reaction (PCR) methods were employed to sequence particular genes of infectious agents. These limited approaches provided initial insights into pathogen diversity and helped identify variants responsible for outbreaks (4). The introduction of next-generation sequencing (NGS) in the early 2000s enabled rapid and cost-effective sequencing of entire pathogen genomes, enabling researchers to explore genetic diversity on a larger scale and monitor the emergence of novel strains. With the accumulation of genomic data, phylogenetic analysis has become crucial for surveillance of infectious diseases. Phylogenetic trees visually represented the evolutionary relationships between pathogen strains, enabling researchers to track transmission routes and identify high-risk clusters (5).

### HIV PANDEMIC

The ongoing HIV pandemic has been a driving force behind the advancements in the genomic surveillance of virus, and the establishment of essential infrastructure and capabilities. Given the severe consequences of untreated HIV infection, the absence of a therapy or vaccine, and its significant socio-political impact, it is declared as a major global public health issue. Moreover, the HIV genome is comparatively complex among RNA viruses because of the presence of circulating recombinant forms (CRFs) and their capacity to acquire resistance to antiretroviral medications. Therefore, genomic surveillance offers several advantages. For instance, in Portugal, routine clinical care involves sequencing HIV genomes, and the resulting data are sent and archived in the HIV drug-resistance database (6). Besides its essential role in HIV programs, the advancement in virus sequencing has been instrumental in facilitating genomic surveillance for various other public health crises.

### SARS COV-1 EPIDEMIC

Subsequent to HIV pandemic, the first outbreak of the 21$^{st}$ century emerged in rural China in 2002. About four months later, scientists identified the causative agent as a novel coronavirus, which was then named SARS-CoV-1. In April 2003, the initial complete genome sequences of the virus were submitted to GenBank. However, at that time, the technology for rapid sequencing, like the NGS facility was not yet available, leading to a delay in generating genomic data. Only 14 SARS-CoV-1 genomes were initially analyzed several months after the onset of the pandemic. Following the pandemic, scientists conducted a study using 61 virus sequences to examine mutational patterns during various phases of the outbreak and explore possible indications of its transmission to humans. Although the SARS pandemic occurred during the "genomic age," yet genomic investigations were limited and retrospective, primarily due to the lack of advanced technologies (5).

### INFLUENZA A VIRUS PANDEMIC

Rapid sequencing of the Influenza A virus was possible during the pandemic in 2009 because of modern sequencing technologies. Genomic investigations identified the animal origin of the virus

while the pandemic was still circulating. Genomic surveillance has been systematically adopted for influenza A virus, with World Health Organization (WHO) utilizing this data to anticipate antigenic variations and enhance vaccine strain selection. Additionally, consistent sequencing of human viruses allows for the identification and monitoring of markers that could influence transmissibility and vaccine efficacy (7).

## POLIO OUTBREAK 2014

Genomic surveillance plays a particularly critical role in the eradication or elimination programs. For instance, during the polio outbreak in 2014, a state of emergency was declared due to the circulation of wild-type poliovirus in multiple countries (8). Genomic surveillance was instrumental in distinguishing these wild-type viruses from vaccine-derived ones, providing crucial insights for optimizing control strategies. Moreover, routine sequencing aids in distinguishing between epidemics caused by new introductions and those resulting from cryptic local transmission. This information is important in identifying high-risk populations that need to be given priority in vaccination efforts. Consequently, real-time genomic surveillance act as a central component of the global poliovirus eradication initiative (9).

## EBOLA VIRUS (EBOV) EPIDEMIC

While over twenty outbreaks in sub-Saharan Africa have been caused by different species of the Ebola virus (a highly infectious pathogen), the majority remained confined to limited geographical regions, infecting a maximum of 425 cases before 2013. However, the Ebola virus epidemic that occurred in West Africa between 2013 and 2016 was more contagious, infecting a total 28,616 individuals with 11,310 reported deaths until June 10, 2016 (10). The EBOV epidemic is marked a crucial milestone for real-time viral genomic surveillance. During this epidemic, a substantial dataset of 1,610 cases (5% of the total) was subjected to sequencing, making it the most extensive dataset of its kind before the SARS CoV-2 pandemic. The breakthroughs in sequencing facilities, particularly the introduction of the Oxford Nanopore Technology (ONT) in 2014, played a pivotal role in facilitating this advancement (11).

## GENOMIC SEQUENCING EFFORTS FOLLOWING SARS COV-2 PANDEMIC

Since its beginning in Wuhan China in December 2019, a massive genome sequencing effort has been underway globally to document the genetic diversity of SARS-CoV-2. Over 14 million cases have been sequenced and several countries are producing their own genomic data. It took just two months for 25 other countries to submit at least one sequence to the Global Initiative on Sharing All Influenza Data (GISAID) when CDC released the first SARS-CoV-2 sequence in January 2020. Several genomic studies have reported significant gene mutation in numerous genes (ORF, E, N, S, and M) (12).

## TECHNIQUES USED IN GENOMIC SURVEILLANCE

Genomic surveillance utilizes an extensive range of molecular and computational tools to sequence, analyze, and evaluate the genetic material of infectious organisms (13). Here is the detail of some of the key techniques used:

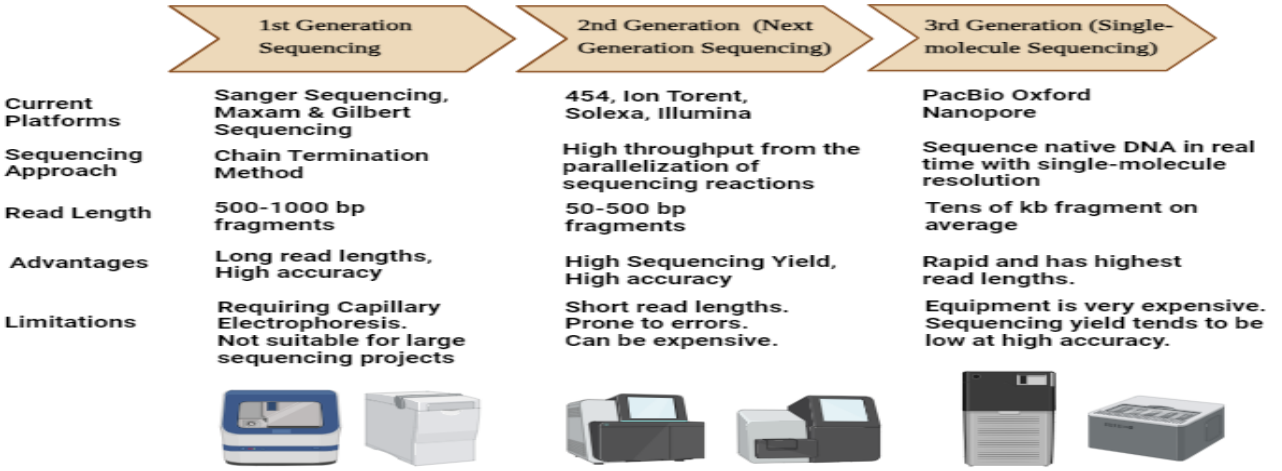
## SEQUENCING TECHNOLOGIES

Sequencing method offers an extensive amount of data and a precise way for subtyping of pathogens. Utilizing sequencing technologies in global surveillance enables researchers to gain comprehensive insights into the early emergence and transmission of pathogens. It provides a high-resolution view of the viral genetic variability and can identify mutations associated with drug resistance, and immune evasion (14). Sanger sequencing is the first commercially available method for DNA sequencing. Following the remarkable achievements of the Human Genome Project, notable progress has been made in Second-Generation Sequencing approaches, commonly known as Next-Generation Sequencing (NGS). Third-generation sequencing represents an emerging innovation tailored for the sequencing of individual molecules. Among these third-generation methods, nanopore sequencing is one of the most extensively applied (15).

**Figure-01: Schematic representation of the advancement in sequencing technologies (15) using dynamic BioRender.com.**

## PHYLOGENETIC ANALYSIS

After sequencing, the data undergoes phylogenetic analysis to determine the genetic relationships between different viral strains. Phylogenetic involves examining the evolutionary relationships between groups of viruses. Molecular phylogenetic, using sequence data, is employed to identify these relationships, encompassing both organisms and their genes. By conducting this analysis, researchers can gain insights into the virus's spread and evolution over time and across various




geographical. Tools such as BLAST (Basic Local Alignment Search Tool) are used for comparing genetic sequences, while other tools like Molecular Evolutionary Genetics Analysis (MEGA) or Bayesian Evolutionary Analysis by Sampling Trees (BEAST) are used for phylogenetic analysis (16).

## GENOME-WIDE ASSOCIATION STUDIES (GWAS)

This technique analyzes numerous genetic variations of multiple genomes to identify statistically significant associations with particular traits or diseases. These are observational studies that examine a comprehensive set of genetic variants across different individuals to determine if any variant is linked to a specific trait, in this instance, the response to viral infections (17).

## HIV GENOMIC SURVEILLANCE

HIV is an infectious agent accountable for inducing acquired immunodeficiency syndrome (AIDS) in human beings. This condition gradually weakens the immune system, creating an environment where life-threatening opportunistic infections can flourish. HIV-1 originated from immunodeficiency viruses found in Central African chimpanzees (SIVcpz), whereas HIV-2 emerged from immunodeficiency viruses in West African sooty mangabeys (SIVsm) (18). Up to now this virus has infected a huge number of people (84.1 million) and caused 40.1 million deaths all over the world (19).

## TYPES OF HIV

Based on genomic variation and other genetic characteristics it has been categorized into, HIV-1 and HIV-2. Type-1 HIV is more prevalent globally as compared to type-2 HIV, which is predominantly found in Western Africa and contributes only limited number of all HIV cases (20). HIV-2 consists of at least nine distinct groups, originally known as subtypes (A to I), with groups A and D currently being prevalent. However, information regarding HIV-2 subtypes is still relatively scarce, and only a small number of recombinants have been identified. HIV-1 is the most common type. When referring to the term "HIV," it is likely to be HIV-1 (21).

## HIV GENOMICS

HIV is a retrovirus characterized by a single-stranded RNA genome having a size of 9.75 kb, featuring a 5'-cap and a 3' poly-A tail. The virus has glycoproteins (gp120 and gp41) that form its outer envelope, facilitating its attachment to the host. Promoters and enhancers responsible for the gene expression of the viral DNA are located within the long terminal repeats (22). It encompasses

of GAG, POL and ENV genes. ENV segments of the genome encodes for gp160, which serves as the precursor to the envelope proteins gp120 and gp41 found in mature virions. These proteins are responsible for the formation of spike proteins that enable the virus to attach and fuse with specific target cells. GAG genes encode several proteins including p17 also known as matrix protein, p24 (capsid protein or CA, called p26 in HIV-2), and p7 (nucleocapsid protein), while POL reading frame encodes for viral enzymes, specifically reverse transcriptase (facilitates the conversion of viral RNA into double-stranded DNA) and integrase (which manages the integration of viral DNA into the host genome) and protease (converting the protein derived from gag or pol to functional proteins through cleaving). Additionally, there are individual genes named rev, tat, nef, vpr, vif, and vpu, each encoding a single protein with the same corresponding name (23).

| Type | Group | Origin | Epidemiology | Comments |
|---|---|---|---|---|
| **HIV-1** | M | SIVepz | Spread to almost all continents except Antarctica. | Responsible for 90% of HIV cases.<br>**Strains:** (**A-D, F-H, J, K**) along with numerous **CRFs** and **URFs**.<br>Subtypes A and F can be further classified into sub-subtypes A1, A2, A3, A4, F1, and F2. |
| | O | SIVgor or SIVepz | West-Central Africa | Very rare. No strain identification yet.<br>Show resistance to NNRTI. |
| | N | Recombinant group M / SIVepz | West & Central Africa | Found in very few individuals.<br>Show resistance to anti HIV-1 drugs. |
| | P | SIVgor | Undetermined | Limited cases |
| **HIV-2** | | SIVsm | Predominantly found in Western Africa & some cases found in Western Europe, USA, India, Japan and Brazil | Relatively scarce.<br>Slower progression<br>Strains: **A,B,C,D,E,F,G,H,I** along with very few recombinants identified |

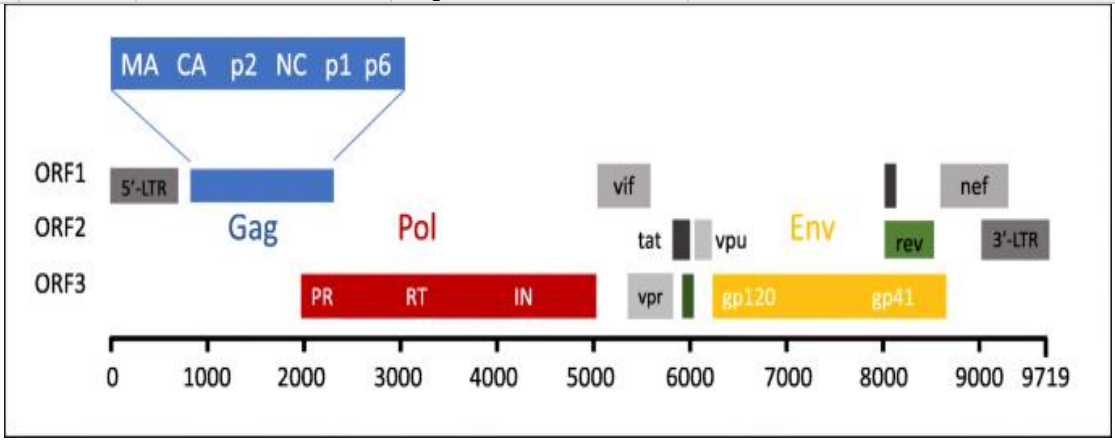



Figure-01: Representation of 9.8 kb genetic makeup of the HIV-1 virus.

## SUBTYPES OF HIV-1

Type-1 HIV demonstrates a distinguishing trait of having a rapid mutation and recombination rate, both within individual hosts and among different hosts. This phenomenon gives rise to distinct groups within HIV-1, including M (Major), O (Outlier), N (None-M, None-O), and the more recent addition, group P (24). Group M further categorized into nine distinct strains (A-D, F-H, J, K) along with numerous circulating recombinant forms (CRFs) and unique recombinant forms (URFs). Subtypes A and F can be further classified into sub-subtypes A1, A2, A3, A4, F1, and F2. As compared to the M group, Group O exhibits almost as many variations. But because it is so rare, scientists have not yet identified its different strains. They show natural resistance to NNRTI drugs. Group N is also very rare and has only been detected in a small number of people from West & Central Africa. Similar to Group O, it possesses genetic differences from Group M and are resistant to anti-HIV drugs. Group P is the most recent HIV-1 group and has been described in 2009 for the 1st time. It has been given its own name because of significant dissimilarities from the M, N, and O strains (20).

**Table-02: A Schematic Representation of HIV Classification Types, Groups and Subtypes**

## EMERGENCE OF GENETIC DIVERSITY OF HIV-1

The diversity of genetic makeup of HIV emerges due to the errors in reverse transcriptase enzyme during replication, causing frequent mutations and recombination events. Recombination requires an individual to be simultaneously infected with multiple different strains of HIV. Recombination of different HIV subgroups are categorized as either circulating recombinant forms (CRFs) or unique recombinant forms (URFs). Until now, over 120 unique CRFs have been identified, and the prevalence of recombinants has been on the rise over the years, both on a global scale and in most geographical areas. Currently, recombinant forms make up nearly 25% of all HIV-1 infections (25). This genetic diversity has significant implications for the control and transmission of the virus. The mutations in viral genome can make it difficult for epitopes to be processed and recognized, leading to immune evasion. These mutations can also cause resistance to anti-viral drugs. Moreover, certain mutations associated with immune evasion or drug resistance can facilitate rapid transmission of the virus (26). Therefore, genomic surveillance is an essential approach in characterizing the transmission dynamics and the prevalence of an HIV variant within a population.

## HIV-1 SUBTYPE DISTRIBUTION GLOBALLY

The rapid evolutionary process of HIV has led to the emergence of various subtypes that are unevenly distributed worldwide (27). Subtype A is primarily found in various parts of East Africa, Russia, and regions formerly belonging to the Soviet Union. Subtype B is prevalent in Europe and America. Subtype D dominates Eastern and Central Africa. Subtype F has been identified in Central Africa, South America, and Eastern Europe. Subtype G has been recorded in both Western and Eastern Africa, as well as in Central Europe. Subtypes G, H, J, and K are widespread in Africa, from where these subtypes have spread to Southern Europe and Asia. Notably, Numerous CRFs and URFs are also distributed simultaneously around the globe (28).

## CURRENT STATUS OF HIV GENOMIC SURVEILLANCE IN PAKISTAN

### OVERVIEW OF HIV OUTBREAKS IN PAKISTAN

Pakistan has experienced a concerning number of HIV outbreaks in the past two decades experiencing its first full-fledged HIV outbreak in 2004 (29). A significant increase (57%) in the HIV-1 outbreak has been observed in Pakistan during the last ten years. It has infected 0.4 million people in Pakistan and causing the tragic loss of 9600 people. National Aids Control Program (NACP) has registered 57,505 patients living with HIV till March 2023 (30).

### GENOMIC SURVEILLANCE STUDIES IN PAKISTAN

In 2019 a study conducted in Punjab explored the genetic characteristics and drug resistance substitutions of prevalent HIV strains across four districts. Lahore exhibited the highest genetic diversity, with subtypes A and G identified, while Faisalabad, Gujranwala, and Sargodha only had CRF02_AG. Phylogenetic analysis revealed three distinct clades among the 18 pol sequences from the study. Fourteen sequences clustered with sub-subtype 02_AG, originating from Pakistan and

Ghana, two sequences clustered with subtype A from Uganda, South Africa, and Pakistan, and two sequences clustered with subtype G from Cameroon, Ghana, and Kenya (31).

In a recent investigation, researchers studied the subtype distribution of HIV-1, conducted phylogenetic analysis, and examined drug resistance mutations in HIV-1 sequences from an outbreak in Larkana, Pakistan, in 2019. This study involved 401 samples collected from individuals infected with HIV, and the results revealed that the predominant subtypes among the outbreak sequences were HIV-1 CRF 02_AG and subtype A1. In the group of participants who had not received treatment, the most frequently observed mutations were RT: E138A (8%) and RT: K219Q (8%) (32).

A study conducted in Lahore in 2020 aimed to assess the prevalence and genetic variability of HIV strains. 49 HIV-positive samples were analyzed, and the results showed the presence of subtype A, CRF02_AG, subtype C, and subtype G. Mutations associated with drug resistance were also observed, including in the protease region (M46L, N88D, L89V) and the reverse transcriptase region (D67T, K70R/Q, M184V, T215F, V108T, E138A, V179I, Y181C). The findings suggest the circulation of multiple HIV-1 subtypes among injection drug users (IDUs), emphasizing the need for continuous disease surveillance and understanding disease risk factors to effectively prevent and treat HIV and reduce epidemics among IDUs (33).

The phylogenetic analysis of samples from Pakistan, China, and India revealed variations in the sequences of the HIV pro gene across these South Asian countries. Interestingly, Pakistani isolates showed more genetic traits with Chinese isolates as compared to Indian isolates. This finding shows that factors like frequent travel, trade, and academic exchanges between Pakistan and China may be responsible for the closer genetic link between Chinese and Pakistani isolates (34).

In Karachi, a study conducted in 2022 aimed to characterize the HIV-1 subtype A1 outbreak in Pakistan using molecular epidemiology and phylodynamic. Analysis of 143 HIV-1 subtype A1 gag sequences, including 61 newly generated sequences, revealed three primary clusters. Two clusters showed close resemblance to Kenyan sequences, while the third cluster comprised 123 Pakistan-specific sequences with limited similarity to other subtype A1 sequences in the database. Specific amino acid positions exhibited signature variations that affected the structure of the Gag protein and Gag-specific T-cell epitopes, as demonstrated by structural protein modeling. The findings indicate that the majority of HIV-1 subtype A1 strains in Pakistan are unique to the country, with a distinct mutation pattern in the Gag genes (35).

Researchers at Dow University of Health Sciences in Karachi conducted a study focused on analyzing gene sequences of HIV to investigate drug resistance patterns, molecular epidemiology, and variations in amino acid sequences. Phylogenetic studies identified the prevalent HIV subtype in Pakistan as A1(68.75%), while also detecting emerging subtypes and recombinant forms like B, CRF02_AG, CRF10_CD, CRF35_AD, and CRF11_cpx. Geographically, the sequences clustered with those from the Middle East, Africa, and parts of Europe. The study unveiled a high level of resistance to non-nucleoside reverse transcriptase inhibitor (NNRTI) drugs, with 62.5% of patients showing resistance and specific mutations such as E138A and K103N. Furthermore, 75% of the sequences exhibited resistance mutation at M184V against nucleoside reverse transcriptase inhibitor (NRTI) drugs. Notably, specific regions of the reverse transcriptase exhibited significant sequence variability at amino acid positions p.119, p.130, p.157, and p.164 (36).

In a study conducted by Abidi and colleagues in 2022, novel unassigned unique recombinant forms (URFs) linked to CRF36_cpx were identified among children in the midst of an HIV-1 epidemic in Pakistan. The study employed pol gene sequencing to determine the HIV subtype and recombined forms, employing a comprehensive approach involving subtype analysis, recombination analysis, and phylogenetic evaluation. The investigation of a collection of 344 samples revealed the emergence of two distinct URFs that had not been previously classified yet shared a relationship with the CRF36_cpx lineage. These unique recombinant forms were detected within a subgroup of 15 individuals (37).

**Table-2: Systematic Review of Genomic Surveillance studies of HIV in Pakistan**

## CHALLENGES

There are many current significant issues in Pakistan that act as barriers to the implementation of public health genomics studies. One of the major challenges in the genomic surveillance of HIV in

Pakistan is the lack of resources and infrastructure. Shortage of professionally trained genomic personnel, well-equipped laboratories, and sequencing technologies are the barriers in the field of genomic surveillance studies. Limited funding for research and development from government authorities and heavy dependence on other countries and agencies further exacerbate this challenge (38).

| Sr. No | Name of Authors | Study Objective | Methodology | Results |
|---|---|---|---|---|
| **1** | Yaqub S, et al., (2019) | Identification of various subtypes (A, G, and CRFs 02_AG) of human immunodeficiency virus type 1 (HIV-1) in particular regions within the Punjab province. | Initially, a total of 130 samples underwent screening using an HIV-1 Ag/Ab Combo test, after which the pol gene was amplified for phylogenetic analysis to identify circulating recombinant forms (CRFs) and mutations. | HIV Positive: 45<br>Positive on Antigen: 27<br>Positive on both Ag & Ab: 18<br>Phylogenetic Study: 18 samples<br>CRF02_AG:14 (77.0%)<br>Subtypes A: 2 (11.5%)<br>Subtype G: 2 (11.1%) |
| **2** | Abidi SH, et al., (2021) | An in-depth investigation of the HIV type 1 sequences from a significant pediatric HIV-1 epidemic in District Larkana, Pakistan, involving both phylogenetic and drug-resistance analyses" | A comprehensive examination was conducted on 532 Pol sequences from the HIV-1 dataset in Pakistan. The investigation involved sequencing and phylogenetic analyses to identify the circulating recombinant forms and drug resistance patterns. | CRF02_AG: 338 (63.5%),<br>Subtype A1: 149 (28.0%),<br>Subtype C: 20 (3.8%),<br>Subtype G: 8 (1.5%),<br>CRF35_AD: 7 (1.3%),<br>Subtype B: 5 (0.9%),<br>Subtype D: 5 (0.9%).<br>Mutations: RT: E138A and RT: K219Q |
| **3** | Yaqub S, et al., (2020) | Prevalence, Molecular Tracking, and Drug Resistance Profile of Human Immunodeficiency Virus Type 1 among Drug Users in Lahore, Pakistan. | Total of 130 samples underwent screening using an HIV-1 Ag/Ab Combo test, after which the pol gene was amplified for phylogenetic analysis to identify circulating recombinant forms (CRFs) and mutations | HIV positive: 49<br>Phylogenetic Study: 26 samples<br>Subtype A:12 (46.15%),<br>CRF_AG: 6 (23.08%),<br>Subtype C: 5 (19.23%),<br>Subtype G 3 (11.54)<br>Mutations:<br>Protease Region - M46L, N88D, L89V mutations<br>Reverse transcriptase Region: "4 NRTI (K70R/Q, D67T, T215F and M184V) and 4 NNRTI (E138A, V108T, Y181C and V179) mutations. " |
| **4** | Shabir S, et al., (2022) | Examination of the pro gene within HIV samples from Pakistan demonstrate the investigation of Single Nucleotide Polymorphisms (SNPs), using phylogenetic studies. | The sequence of the HIV protease gene (pro-gene) was obtained from NCBI for three countries, with 174, 117, and 105 sequences retrieved from Pakistan, China, and India, respectively. These sequences were then subjected to analysis using bioinformatics tools, MEGAX, and CLUSTALW, to investigate their linkage and genetic mutations. | Pakistani isolates showed more genetic traits with Chinese isolates as compared to Indian isolates. This finding shows that factors like frequent travel, trade, and academic exchanges between Pakistan and China may be responsible for the closer genetic link between Chinese and Pakistani isolates. |

| | | | | |
|---|---|---|---|---|
| | | | | |
| **5** | Tariq U, et al, 2022 | Detailed examination of HIV-1 Subtype A1 outbreak in Pakistan using advance molecular epidemiology and Phylodynamics. | This involves studying 143 genetic sequences from the gag region of HIV-1 sub-subtype A1. The aim is to establish the link between the A1 strains identified in Pakistan and those prevalent in the global pandemic. Subtype A1 Clusters Identified: 03 . | Similarity to Kenyan Sequences: 2 clusters Unique to Pakistan: 01 cluster Notably, both the Pakistan-specific cluster and the reference strains from Kenya displayed distinct variations, particularly at specific amino acid positions: 312, 319, 331, 372, 373, 383, and 402 |
| **6** | Zahid M, et al., (2022) | Examination of sub-genomic sequences unveils the emergence of novel circulating recombinant forms of HIV-1 within Pakistan ” | This research examined the genetic sequences of HIV to investigate both drug resistance and molecular epidemiology. The variability in amino acid sequences was also studied. Moreover, the study involved conducting phylogenetic analysis to explore potential geographic connections among HIV strains in Pakistan. | **Subtype Distribution** Subtype A1: 68.75 % Subtype B, CRF02_AG, CRF10_CD, CRF35_AD, and CRF11_cpx: 6.25% each **Mutations** Drug Resistance: NNTRI: 62.5% NTRI: 75% RT region Amino Acid Seq variations: ”p.119, p.130, p.157, p.164. |
| **7** | Abidi SH, et al., (2022) | Identification of URFs associated with CRF36_cpx among children during an HIV-1 epidemic in Pakistan. | Through sequencing of the pol gene, the HIV subtype and recombinant forms were discerned using a combination of subtype analysis, recombination analysis, and phylogenetic assessment. | Within the repository of 344 samples, we noted the emergence of two distinct unique recombinant forms hitherto unclassified, which exhibited a connection to the CRF36_cpx lineage. These unique recombinant forms were identified in a subset of the 15 individuals |

In Pakistan, the lack of a centralized national framework for genomic surveillance of HIV poses obstacles to data sharing and collaboration among various regions and institutions. This absence of coordination significantly restricts the creation of comprehensive and representative datasets, hindering the effective analysis and interpretation of HIV genomic data. Stigma and discrimination related to HIV/AIDS continue to persist in Pakistan. The fear of social exclusion and persecution creates a climate of underreporting and reluctance among individuals living with HIV to engage in research studies, including those focused on genomic surveillance initiatives.

The implementation of genomic surveillance gives rise to ethical considerations regarding the privacy and confidentiality of individuals' genetic information. In Pakistan, the absence of strong legal frameworks and ethical guidelines for genomic research intensifies these concerns. The lack of clear regulations may hinder individuals' willingness to take part in genomic surveillance studies, thereby restricting the availability of comprehensive genomic data (39).

## CONCLUSION

Pakistan lacks a comprehensive molecular phylogenetic characterization of HIV strains. To address this, the government needs to initiate HIV genomic surveillance involving experts in public health, molecular epidemiology, and bioinformatics. By integrating patient interviews, demographic

profiling, and phylogenetic analysis of HIV sequences from blood samples, they can establish evidence on the origin and spread of outbreaks. Through robust data collection, sharing, and analysis systems, Pakistan can enhance its ability to detect and respond to changes in the HIV landscape. This genomic surveillance has the potential to transform the country's response to the epidemic, leading to improved prevention strategies, optimized treatments, and reduced HIV burden.

**REFERENCES**


Lambrou AS, Shirk P, Steele MK, Paul P, Paden CR, Cadwell B, et al. Genomic surveillance for SARS-CoV-2 variants: predominance of the Delta (B. 1.617. 2) and Omicron (B. 1.1. 529) variants—United States, June 2021–January 2022. Morbidity and Mortality Weekly Report. 2022;71(6):206.

Organization WH. Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health 2021 [cited 2023 August 01]. Available from: https://www.who.int/publications/i/item/9789240018440.

Stockdale JE, Liu P, Colijn C. The potential of genomics for infectious disease forecasting. Nature Microbiology. 2022;7(11):1736-43.

Shi W, Li J, Zhou H, Gao GF. Pathogen genomic surveillance elucidates the origins, transmission and evolution of emerging viral agents in China. Science China Life Sciences. 2017;60:1317-30.

Hill V, Githinji G, Vogels CB, Bento AI, Chaguza C, Carrington CV, et al. Toward a global virus genomic surveillance network. Cell host & microbe. 2023.

Vasylyeva TI, Du Plessis L, Pineda-Peña AC, Kühnert D, Lemey P, Vandamme A-M, et al. Tracing the impact of public health interventions on HIV-1 transmission in Portugal using molecular epidemiology. The Journal of Infectious Diseases. 2019;220(2):233-43.

Morris DH, Gostic KM, Pompei S, Bedford T, Łuksza M, Neher RA, et al. Predictive modeling of influenza shows the promise of applied evolutionary biology. Trends in microbiology. 2018;26(2):102-18.

Moturi EK, Porter KA, Wassilak SG, Tangermann RH, Diop OM, Burns CC, et al. Progress toward polio eradication—worldwide, 2013–2014. Morbidity and Mortality Weekly Report. 2014;63(21):468.

Macklin G, O'Reilly K, Grassly N, Edmunds W, Mach O, Santhana Gopala Krishnan R, et al. Evolving epidemiology of poliovirus serotype 2 following withdrawal of the serotype 2 oral poliovirus vaccine. Science. 2020;368(6489):401-5.

Organization WH. Ebola virus disease 2019 [Available from: https://www.who.int/health-topics/ebola/#tab=tab_1.

Dudas G, Carvalho LM, Bedford T, Tatem AJ, Baele G, Faria NR, et al. Virus genomes reveal factors that spread and sustained the Ebola epidemic. Nature. 2017;544(7650):309-15.

Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579(7798):265-9.

Dulanto Chiang A, Dekker JP. From the pipeline to the bedside: advances and challenges in clinical metagenomics. The Journal of infectious diseases. 2020;221(Supplement_3):S331-S40.

Rifaie S, Patil V, Jangid K. Advances in sequencing technology, databases, and analyses tools for the assessment of microbial diversity. Microbial Diversity in Hotspots: Elsevier; 2022. p. 317-47.

Partha P. Paul P, MBA-Managing. The Future of DNA Sequencing & How it impacts the Job Market! : blog.connexissearch.com; 2021 [cited 2023 Aug 20]. Available from: https://blog.connexissearch.com/employersblog/dna-sequencing-the-future-and-how-it-impacts-the-job-market.

Subbotin SA. Phylogenetic analysis of DNA sequence data. Techniques for work with plant and soil nematodes: CABI Wallingford UK; 2021. p. 265-82.

Uffelmann E, Huang QQ, Munung NS, De Vries J, Okada Y, Martin AR, et al. Genome-wide association studies. Nature Reviews Methods Primers. 2021;1(1):59.

Waheed U, Noor FA, Saba N, Wazeer A, Qasim Z, Arshad M, et al. Genetic Diversity of Human Immunodeficiency Virus Type 1 in Asymptomatic Blood Donors in Islamabad, Pakistan. Journal of Laboratory Physicians. 2020;12(02):092-7.
Organization WH. HIV/AIDS 2023 [Available from: https://www.who.int/news-room/fact-sheets/detail/hiv-aids.
Johansson E. Virus-host interactions in HIV-1 and HIV-2 infections. 2023.
Giovanetti M, Ciccozzi M, Parolin C, Borsetti A. Molecular epidemiology of HIV-1 in African countries: a comprehensive overview. Pathogens. 2020;9(12):1072.
Program NT. Report on Carcinogens Monograph on Human Immunodeficiency Virus Type 1. 2022.
Boswell M. Impact of host restriction factors and capsid evolution on HIV disease progression: University of Oxford; 2019.
Santoro MM, Perno CF. HIV-1 genetic variability and clinical implications. International Scholarly Research Notices. 2013;2013.
25. Hemelaar J, Elangovan R, Yun J, Dickson-Tetteh L, Kirtley S, Gouws-Williams E, et al. Global and regional epidemiology of HIV-1 recombinants in 1990–2015: a systematic review and global survey. The Lancet HIV. 2020;7(11):e772-e81.
Johnson MM, Jones CE, Clark DN. The Effect of Treatment-Associated Mutations on HIV Replication and Transmission Cycles. Viruses. 2022;15(1):107.
Sharp PM, Hahn BH. The evolution of HIV-1 and the origin of AIDS. Philosophical Transactions of the Royal Society B: Biological Sciences. 2010;365(1552):2487-94.
Bbosa N, Kaleebu P, Ssemwanga D. HIV subtype diversity worldwide. Current Opinion in HIV and AIDS. 2019;14(3):153-60.
Ahmed A, Hashmi FK, Khan GM. HIV outbreaks in Pakistan. The Lancet HIV. 2019;6(7):e418.
30. Unit CM. NACP – National AIDS Control Programme 2023 [Available from: https://www.cmu.gov.pk/nacp-national-aids-control-programme.
Yaqub S, Yaqub T, Shabbir MZ, Nadeem A, Mukhtar N, Rehman FAU, et al. Circulation of multiple subtypes (A, G and CRFs 02_AG) of human immunodeficiency virus type 1 (HIV-1) in selected districts of Punjab province, Pakistan. Archives of Virology. 2019;164:3081-7.
Abidi SH, Nduva GM, Siddiqui D, Rafaqat W, Mahmood SF, Siddiqui AR, et al. Phylogenetic and drug-resistance analysis of HIV-1 sequences from an extensive paediatric HIV-1 outbreak in Larkana, Pakistan. Frontiers in microbiology. 2021;12:658186.
Yaqub S, Yaqub T, Shabbir MZ, Nadeem A, Shahid MF, Tahir Z, et al. Prevalence, Molecular Epidemiology and Drug Resistance Pattern of Human Immunodeficiency Virus Type 1 among Injecting Drug Users in Lahore, Pakistan. Pakistan Journal of Zoology. 2021;53(2):563.
Sana Shabir HT, Khaver Ali, Sana Riaz, Mehboob Ahmed, Tariq Javed. Single Nucleotide Polymorphisms (SNPs), Phylogeny and Spatiotemporal analysis of pro gene in HIV samples from Pakistan.
. Pakistan Journal of Medical & Health Sciences 2021;15(2021):4.
Tariq U, Nazziwa J, Sasinovich S, Shah SA, Naeem S, Abidi SH, et al. Phylogenetic Characterization of HIV-1 Sub-Subtype A1 in Karachi, Pakistan. Viruses. 2022;14(10):2307.
Zahid M, Khan S, Qureshi MA, Raza Y. Subgenomic sequence analysis reveals emergence of new circulating recombinant forms of HIV-1 in Pakistan. Pakistan Journal of Medical Sciences. 2022;38(7):1857.
Abidi SH, Siddiqui D, Mahmood SF, Siddiqui AR, Nathwani AA, Hotwani A, et al. Unassigned complex unique recombinant forms related to CRF36_cpx in children identified in an HIV-1 outbreak in Pakistan. AIDS research and human retroviruses. 2022;38(10):806-11.
Gandra S, Alvarez-Uria G, Turner P, Joshi J, Limmathurotsakul D, van Doorn HR. Antimicrobial resistance surveillance in low-and middle-income countries: progress and challenges in eight South Asian and Southeast Asian countries. Clinical Microbiology Reviews. 2020;33(3):10.1128/cmr. 00048-19.
Riaz M, Tiller J, Ajmal M, Azam M, Qamar R, Lacaze P. Implementation of public health genomics in Pakistan. European Journal of Human Genetics. 2019;27(10):1485-92.