Comparison of the genotypes of the West Nile virus identified in Iran and neighboring countries

Hamid Staji[1], Seyedeh Fatemeh Angoshtan[2*]

1. Department of Pathobiology, Faculty of Veterinary Medicine, Semnan University, Semnan, Iran.

2. Department of Basic Sciences, Faculty of Veterinary Medicine, Semnan University, Semnan, Iran.

***Corresponding author.** Seyedeh Fatemeh Angoshtan

Semnan University, Semnan, Iran, Email: angoshtan.sf@gmail.com, +98-9214202835

Abstract

West Nile virus (WNV) is a Flavivirus transmitted by arthropods, and it is a public health concern globally. WNV was originally isolated in Uganda, and within a short period, it spread widely and caused outbreaks globally. We conducted this study to research strains from Iran, Azerbaijan, the United Arab Emirates, Pakistan, Iraq, and Turkey to understand WNV spread, genetic diversity, and evolution, and improve monitoring and control programs by evaluating the sources of introduction and modes of transmission.

A total of 93 gene sequences from Iran and neighboring countries were downloaded from the NCBI virus database. After filtering, 89 sequences were used for analysis and the construction of phylogenetic trees for several genes, alongside the genome sequences of reference lineages 1 and 2 for comparison. Sequence alignment was performed in MEGA 10 via ClustalW. Phylogenetic trees were constructed using the neighbor-joining method. Phylogenetic grouping stability was determined via the bootstrap analysis (1,000 replications).

Our analyses of the capsid, envelope protein, and NS5 genes demonstrated the importance of Iran in the circulation of WNV in the region. Although some sequences from Iran cluster with Turkish sequences, others present evidence of independent evolution. The clustering of Azerbaijani sequences with Turkish sequences and the divergence of Iraqi and Pakistani strains revealed varying degrees of connectivity and isolation. Furthermore, the hosts identified in the region include different species of mosquitoes, birds, and mammals, which were described in detail.

Our phylogenetic analysis revealed the geographical distribution of WNV transmission. The genetic diversity and connectivity of the strains provide crucial information for epidemic

prediction, virus transmission, and areas that require enhanced surveillance. In the future, these evolutionary patterns will also assist in monitoring WNV transmission in Iran and its neighboring countries.

Keywords: West Nile virus; phylogeny; whole-genome sequencing; viruses; phylogenetic analysis; zoonoses; arboviruses

1. Context

West Nile virus (WNV) is an enveloped, single-stranded positive-sense RNA virus belonging to the Japanese encephalitis serocomplex that includes a group of closely related viruses in the genus Flavivirus, family Flaviviridae (1). This virus is the main cause of a human disease known as West Nile fever, which is an important public health pathogen in humans and animals. It is considered one of the most widely distributed arthropod-borne viruses (2).

The genome of WNV consists of one strand of RNA approximately 11 kb in length. The viral RNA encodes ten proteins, which include three structural proteins, C (capsid), prM (premembrane), and E (envelope), as well as the seven nonstructural proteins NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5. The RNA is translated into a single large polyprotein that is subsequently cleaved by both viral and host proteases into individual proteins. The structural proteins comprise the virion, which encapsulates the viral RNA, and the nonstructural proteins comprise the replication complex (3).

This virus was first isolated from a febrile patient who lived in the West Nile District of Uganda in 1937. The first cases of WNV infection in horses were recorded in the 1960s in France and Egypt. Since then, the virus has been causing recurrent outbreaks in Africa, the Middle East, and parts of Europe. Reported cases of WNV in Turkey occurred during the 1970s and 1990s, as did severe epidemics from 2010--2011. The country of Iran has also reported cases dating back to the 1970s, 2008, and 2009, with patients showing fever and some neurological signs. A study conducted on horses in Iran from 2008--2009 revealed wide WNV spread in the western and southern provinces (4).

WNV has been identified in 65 distinct species of mosquitoes and more than 326 species of birds. The primary vectors for transmission to mammals, including humans, are *Culex spp.* (5). *Cx. pipiens, Ae. caspius, and Cx. theileri* are reported as potential vectors for WNV in Iran (6). It exists in its natural environment through the enzootic cycle, with *Culex* mosquitoes as vectors and particular species of birds as primary amplifying hosts. While mammals, including humans and horses, can develop WNV infections, they typically do not generate enough viremia to reinfect *Culex*; therefore, they are considered dead-end hosts (7, 8). Viremia most commonly lasts for a few days, but depending on the host species, viremia intensity varies greatly, being greater in birds than in mammals (5).

Assessing the actual incidence and prevalence of infections is difficult owing to the differing testing methods available. Nonetheless, whole-blood PCR is more sensitive than IgM ELISA is and needs more investigation as a possible means of diagnosis. In North America, the rapid spread of the West Nile virus ignited veterinary vaccine development by 2005; however, the development of this human vaccine remains in preclinical studies (9).

WNV has at least nine different lineages. Nevertheless, lineages 1a and 2 are the primary causes of most of the described human cases (10). The original strain lineage 2 was dominant in the African, Middle Asian, Arab, Russian, and European regions for more than 60 years, giving rise primarily to milder infections. Lineage 1, however, emerged in the mid-1990s, causing neurological infection in Europe, Russia, and Israel (9). Lineage 1 of the West Nile virus can be further divided into three sublineages: 1a, 1b, and 1c. Lineage 1b, also known as Kunjin virus, primarily occurs in Australia. Lineage 1c, which is also referred to as lineage 5, includes isolates from India (11). Lineage 3, also known as the Rabensburg virus, and several other less common lineages, including lineage 4 in Russia, lineage 5 in India, and lineage 6 in Spain, have all likely made their way into the Northern Hemisphere through separate introductions. Further lineages are being discovered in Africa (12).

Phylogenetic analysis employs phylogenetic trees to investigate evolutionary relationships among species or genes. The trees provide researchers with the ability to trace resulting lineages and relationships with a common ancestor to shed light on the evolution of organisms and gene sequences. These relationships are important for the prediction of gene function, tracking disease emergence, and guiding decisions in other areas, such as conservation biology and drug development (13). Recent studies on vector-borne pathogens have emphasized the importance of molecular detection, phylogenetic analysis, and serological surveillance in assessing disease transmission dynamics (14-18).

Recent advancements have enhanced our understanding of WNV, mainly investigating its genetic, transmission, and environmental factors linked to outbreak initiation. Studies about the genetics of the West Nile virus have revealed some interesting patterns in how it spreads and changes in Europe (10, 19). Recently, a study found West Nile Virus in an area that was considered low-risk before, which shows how it's spreading unexpectedly. Many people who got infected didn't even notice mosquito bites or realize they should take precautions (20). There are also some advancements in host selection of this virus, revealing ecological and environmental determinants influencing their role in transmission of West Nile virus (7). New evidence further shows that, besides ecological determinants conducive to the presence of birds and mosquitoes, land use within agricultural regions can have a pivotal role in the emergence of WNV (21).

While WNV has been documented in Iran since the 1970s, reports of genetic data of Iranian strains in comparison to neighboring countries are still lacking. The concern here lies in the uncertainty over whether Iran serves as a major hub of transmission of the virus or merely acts as a pass-through corridor in the neighboring countries, especially through migratory birds or common ecosystems. This study is the first to report a multi-gene phylogenetic study of WNV strains from Iran and neighboring countries. We have demonstrated regional connections but also distinct evolutionary aspects of WNV strains, which suggests that Iran has been both a recipient of and

could represent a potential source of WNV spread in this region. WNV has been frequently reported in Iran, as well as in many bordering countries (22-25). In this study, we constructed phylogenetic trees using 93 available genome datasets from WNV samples collected in Iran and neighboring countries. This allowed us to trace the evolutionary relationships and diversity of the viruses in the region. By analyzing these trees, we can determine the patterns of transmission of this virus and the possible sources of introduction. This information is valuable for revealing the pathways through which the West Nile virus might be spread and for developing surveillance and control strategies. Our research will also shed light on the genetic diversity of the virus, providing essential insights for future development and research into vaccines and therapeutic antiviral agents.

2. Data Acquisition

For this study, gene sequences from Iran and bordering countries, including Azerbaijan, Iraq, Pakistan, the United Arab Emirates, and Turkey, were available and initially downloaded from the NCBI virus database. An initial analysis involving all WNV sequences from these countries was carried out. All WNV sequences from these countries, as submitted to the NCBI virus database, were selected for evaluation. Each sequence was then analyzed for the presence of genes of capsid protein, envelope protein, and NS5 protein. All sequences including these genes were selected. Furthermore, all whole genome sequences have also been included in the study. A total of 93 sequences were retrieved. The sequences were distributed as follows: 7 from Azerbaijan, 11 from Iran, 2 from Iraq, 4 from Pakistan, 68 from Turkey, and 1 from the United Arab Emirates. Among the 93 sequences analyzed, 89 were used due to the previously stated criteria to construct phylogenetic trees for the complete genome, capsid protein, envelope protein, and NS5 protein. One sequence from Turkey, two from Pakistan, and two from Iran were excluded because of their lack of relevance to these genes. Figure 1 shows the locations of the sequences analyzed in this study. Furthermore, genome sequences from lineages 1 and 2 were included to examine the relationships between these countries and the reference sequences.

Gene sequences in FASTA format were imported into MEGA10, and sequence alignments were performed via ClustalW. Neighborhood joining was used to generate phylogenetic trees. The reliability of the phylogenetic groupings was evaluated via a bootstrap test (1,000 replications). The study used neighborhood joining and the bootstrap test based on the research of Shah-Hosseini et al.. The authors presented a strong framework for neighborhood joining and its application to similar research questions, which supports the use of neighborhood joining (26). The neighbor-joining (NJ) method, relying purely on pairwise distances (rather than complex probabilistic models), offers a fast and efficient means of constructing phylogenetic trees. For purposes of phylogeny with datasets that are larger and complex, it is preferred to use maximum likelihood because of its simplicity and computational needs, although maximum likelihood provides superior accuracy for deep evolutionary analyses. The Neighborhood method's flexibility to variable evolutionary rates helps to make it a practical approach for analysis.

In addition to the detailed phylogenetic analyses, a thorough review of the literature was performed to provide evidence of an understanding of West Nile virus (WNV) transmission in Iran and neighbouring countries. We searched the scientific literature using PubMed, Scopus, and Web of

Science databases, drawing data from the last decade to capture as much relevant development as possible. Keywords such as "West Nile virus," "whole-genome sequencing," "phylogenetic analysis," "arboviruses," "zoonotic transmission," and "migratory bird routes" were used. The articles were selected according to their relevance to the subject areas of molecular epidemiology, vector-host ecology, and phylogeographic patterns. Priority was given to studies that used genomic or phylogenetic approaches. Studies that explored hosts, climate, as well as local transmission pathways were prioritized. This background of the literature provided insight into the sequence data, as well as the necessary data on the transmission potential of WNV strains, with specific emphasis on Iran and the surrounding countries, where regional patterns of transmission, typical ecosystems, and migratory bird routes offer critical background to understand virus circulation.

Although the approach taken in this study provides a good representation of the phylogenetic relationship of WNV sequences, it does have limitations. First, the inclusion of available sequences from the NCBI virus database can have some biases because there is a possibility of undercoverage of certain geographic areas due to limited sequencing. Also, the removal of certain sequences due to missing genes or incomplete genome data may affect the completeness of the study. These limitations will be addressed to improve the description of the phylogenetic analysis and provide a comprehensive assessment of WNV evolution in the region in future studies. Finally, Neighborhood joining measures exclusively rely on distance matrices, which offer less accuracy than a model-based approach.

Figure 1. Location of WNV strains isolated. Sequences shown in the figure were retrieved from the NCBI virus and were utilized in this study. Sequence 1, as shown in Figure 1, was isolated from Isfahan, Iran (KJ486150.1). Sequence 2 was isolated from the Sepid River, Iran

(MF462262.1). sequences 3 and 4 were isolated from Lorestan, Iran (MN238669.1, MN238670.1). sequences 5-9 were isolated from Hormozgan, Iran (MW455339.1, MW455338.1, MW455337.1, MW455336.1, and MW455335.1). Among the sequences analyzed, 67 originated from Turkey, 7 from Azerbaijan, 3 from Pakistan, 2 from Iraq, and 1 from the United Arab Emirates.

3. Results

3.1 Complete genome sequences

The full-genome nucleotide sequences were established for eight WNV isolates. WNV whole genome sequences were available in Azerbaijan, the United Arab Emirates, Turkey, and Iran. The analysis of WNV sequences revealed 99% site coverage for Iranian sequences aligned with lineage 2. Furthermore, the sequences from Turkey and the United Arab Emirates displayed significant similarities to one another and to lineage 1. Additionally, the WNV sequences from Azerbaijan presented high site coverage with these lineages. Figure 2 displays the phylogenetic analysis of complete genomes of WNV from Azerbaijan, the UAE, Turkey, and Iran.

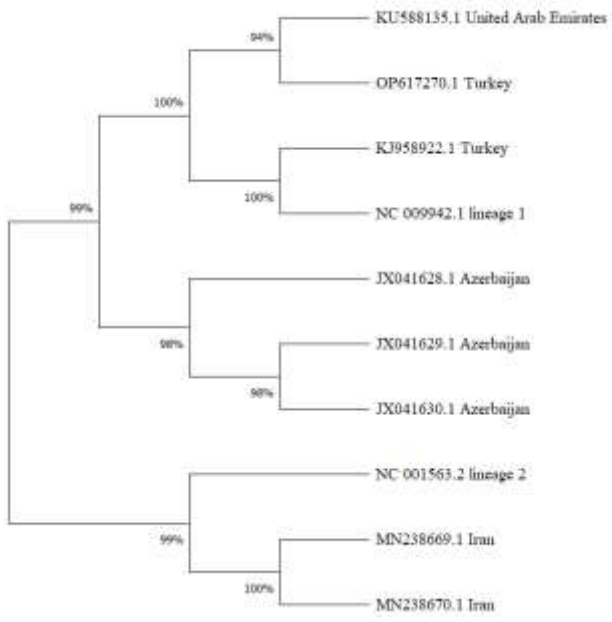

Figure 2. Phylogenetic analysis based on the complete genome sequence of WNV from Azerbaijan, the United Arab Emirates, Turkey, and Iran. The phylogenetic relationships of the full-length sequences were inferred via MEGA 10 via the neighbor–joining method and tested via 1000 bootstrap replicates. The sequences are indicated by the GenBank accession number and country of isolation. The percentage shown in the figure indicates the extent of branch site coverage.

3.2 Capsid gene sequences

Twenty-two capsid protein sequences, representing strains from Iran, Turkey, Pakistan, and Iraq, were obtained from the NCBI database. To have a more representative context in the phylogenetic analysis, reference sequences of Lineages 1 and 2 were also selected and aligned. Phylogenetic analysis was performed using sequences of the capsid gene. Figure 3 shows the phylogenetic analysis of West Nile Virus (WNV) capsid gene sequences from Iran, Turkey, Pakistan, and Iraq.

High site coverage of some branches formed by the clustering of Turkey's sequences and those from Iraq signifies the strong reliability of these evolutionary connections. Additionally, the Iranian sequences showed significant clustering. Lineages 1 (NC 009942.1) and 2 (NC 001563.2) served as outgroups, further confirming the phylogenetic structure. The phylogenetic analysis revealed relevant clustering of the sequences from Iran, implying a common ancestry with some genetic resemblance among the strains of WNV circulating in this area. Interestingly, the Iranian sequences also reveal a genetic relationship with strains from neighboring countries such as Iraq and Turkey. This regional clustering suggests that there may have been a historical or ongoing genetic exchange between these countries. This exchange could be driven by migratory birds or by other ecological and environmental factors. The presence of genetic similarities may suggest shared transmission pathways or common environmental conditions that have shaped the evolution of the virus across these areas. Conversely, the Pakistani strains appear to have diverged from the Iranian cluster, forming a separate group. This could signify either a unique evolutionary trajectory in that region or limited genetic exchange between the Pakistani strains and those in Iran.

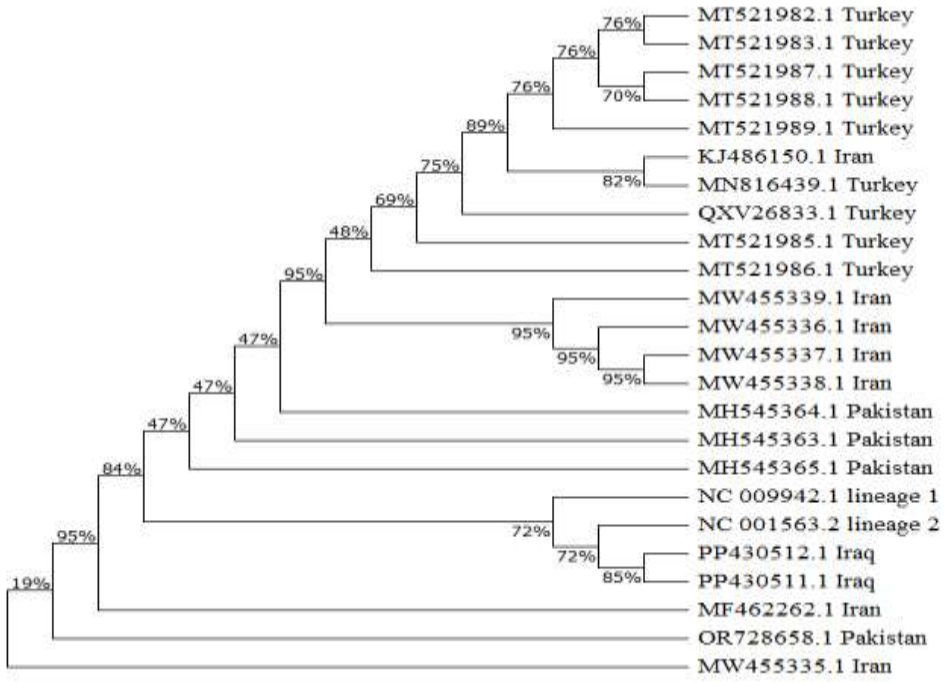

Figure 3. Phylogenetic analysis of the capsid gene sequences of WNV from Iran, Turkey, Pakistan, and Iraq. The phylogenetic relationships of the capsid gene sequences were inferred via MEGA 10 via the neighbor–joining method and tested via 1000 bootstrap replicates. The sequences are indicated by the GenBank accession number and country of isolation. The percentage shown in the figure indicates the extent of branch site coverage.

3.3 The envelope protein (E) gene sequences

A total of 54 sequences were included in the analysis, 52 of which originated from Iran, Turkey, Azerbaijan, and Iraq. Additionally, reference sequences for Lineage 1 and Lineage 2 were included. The majority of the tree sequences were from Turkey and formed several separate clusters. This may suggest that Turkey was one of the major sites of virus sampling. The sequences from Azerbaijan (AF237563.1, AF237564.1) are closely related to those from Turkey, suggesting a geographic link with Turkey. This may indicate that viruses spread or share evolutionary pressures in this region. Iran has a sequence (MF462262.1) placed further apart from these clusters; however, the lack of data for this gene in Iran can influence the results. The Iraq and Pakistani sequences (PH430512.1, PP430511.1, OR728658.1) also cluster far apart, possibly reflecting geographic or lineage-specific separation. These findings may be useful for comparison with other sequences to understand the relationships between Turkey and its neighbors with lineages around the globe. In this dataset, Turkey appears in the center and indeed could act as the hub, given its geographic placement connecting Europe, Asia, and the Middle East. Figure 4 shows the phylogenetic analysis of West Nile Virus (WNV) envelope protein (E) gene sequences from Iran, Turkey, Azerbaijan, and Iraq.

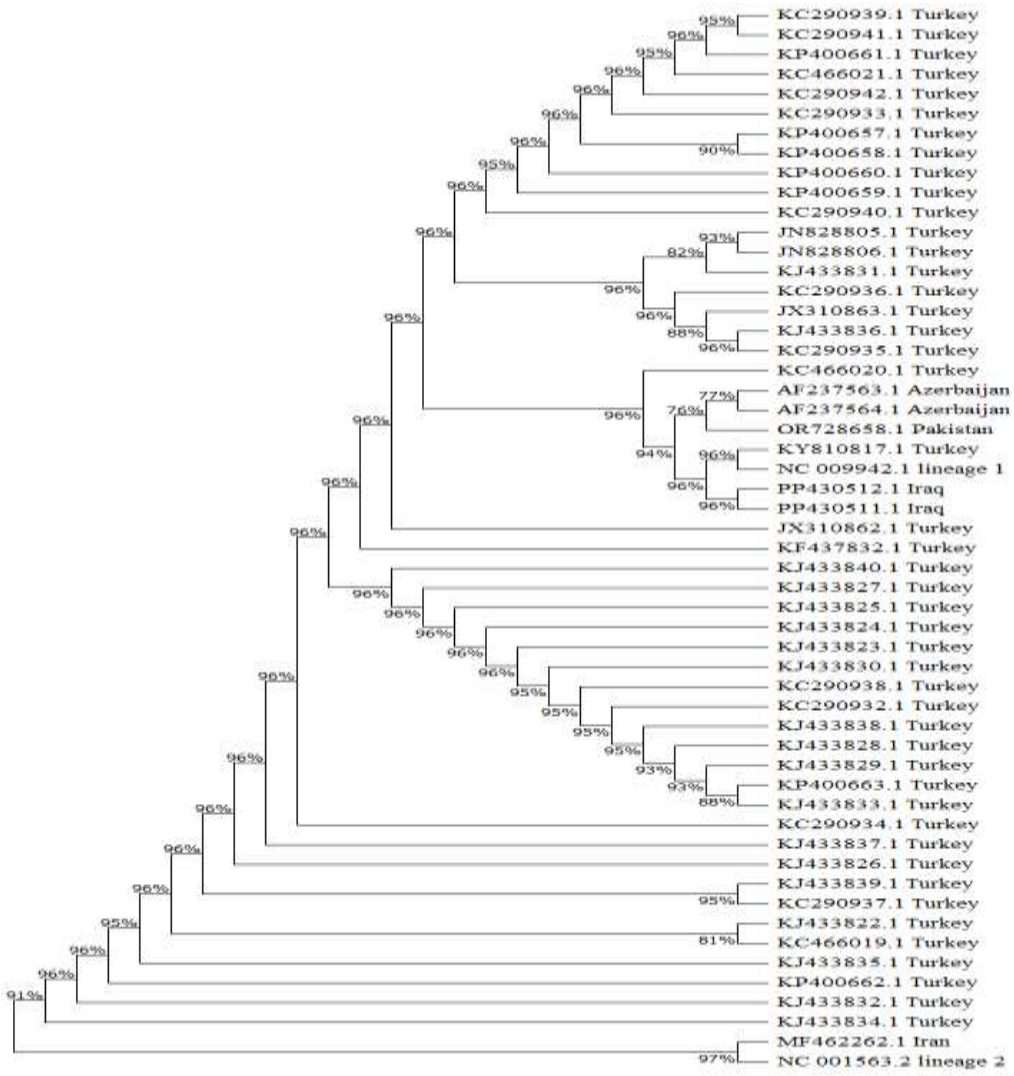

Figure 4. Phylogenetic analysis based on the envelope protein (E) gene sequence of WNV from Iran, Turkey, Azerbaijan, and Iraq. The phylogenetic relationships of the E gene sequences were inferred via MEGA 10 via the neighbor–joining method and tested via 1000 bootstrap replicates. The sequences are indicated by the GenBank accession number and country of isolation. The percentage shown in the figure indicates the extent of branch site coverage.

3.4 NS5 gene sequences

The phylogenetic tree was constructed based on 16 NS5 gene sequences, along with two reference sequences of the West Nile virus. The tree contains samples from Turkey, Iran, Azerbaijan, Pakistan, and Iraq. Some Turkish sequences cluster together, indicating their genetic similarities. Pakistan and Iraq had more distant branches. Considering Iran's central location, the distinct phylogenetic branch could point to virus introductions from eastern regions or neighboring countries such as Pakistan or Iraq, which show evolutionary paths separate from Turkey in NS5

phylogenetic analysis. However, the lack of data on the NS5 gene sequence from Iran influenced the findings. A different approach with greater sampling and tracing of biogeographic links could clarify the issues within the analyzed data. The phylogenetic analysis of West Nile Virus (WNV) NS5 gene sequences from Turkey, Iran, Azerbaijan, Pakistan, and Iraq is depicted in Figure 5.

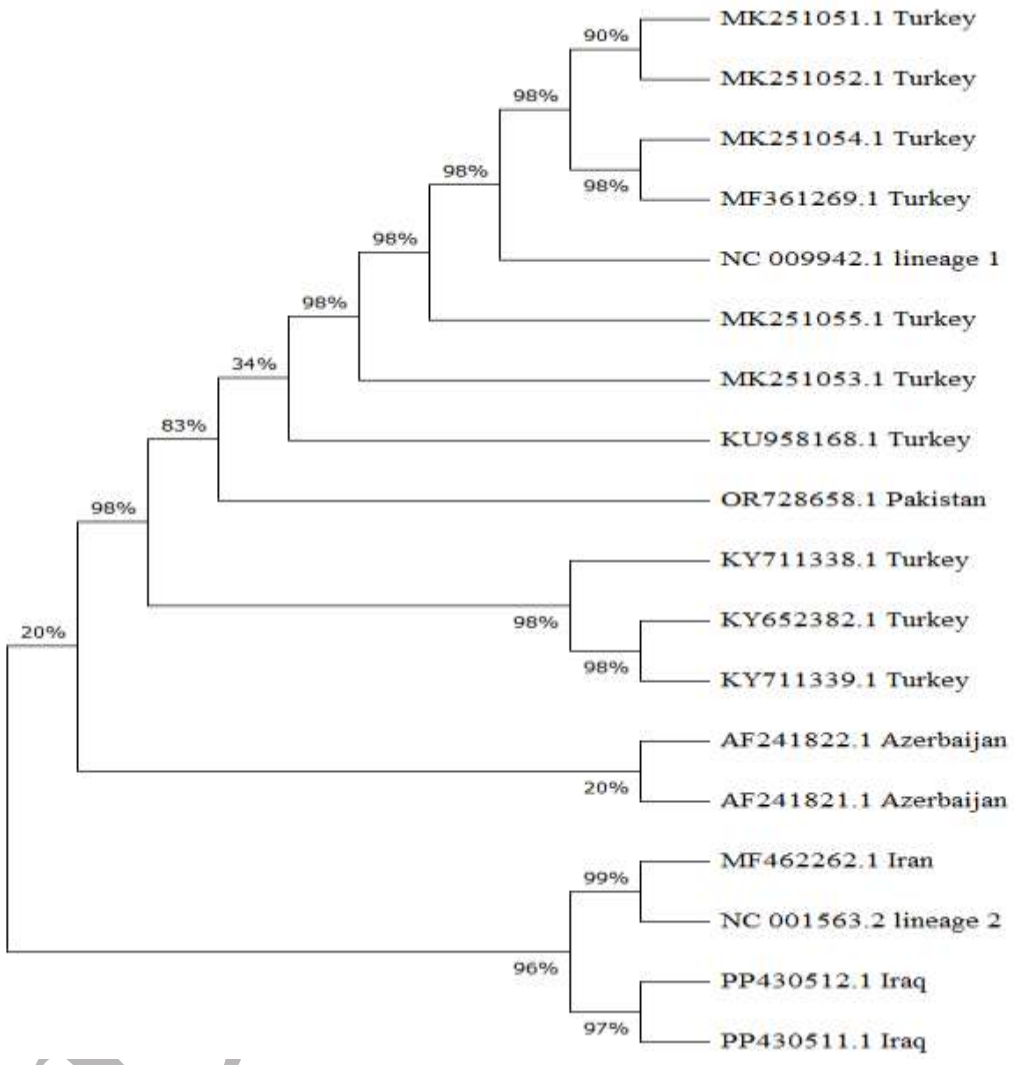

Figure 5. Phylogenetic analysis based on the NS5 gene sequence of WNV from Turkey, Iran, Azerbaijan, Pakistan, and Iraq. The phylogenetic relationships of the NS5 sequences were inferred via MEGA 10 via the neighbor–joining method and tested via 1000 bootstrap replicates. The sequences are indicated by the GenBank accession number and country of isolation. The percentage shown in the figure indicates the extent of branch site coverage.

3.5 Observations across all four trees

Turkish sequences consistently dominate the trees and form multiple distinct clusters. This highlights Turkey as a significant hub of genetic diversity for the West Nile virus. Its geographical

location, which links Europe, Asia, and the Middle East, can provide another perspective toward filling this role in the evolution of a virus. Iranian sequences often appear in relatively isolated positions or distinct branches. While some Iranian sequences cluster with Turkish sequences, others show signs of independent evolution. Sequences from Azerbaijan tend to group very closely together. Their close association with Turkish sequences indicates strong genetic connectivity, probably resulting from the migratory routes of birds or ecological interactions. Sequences from Iraq and Pakistan are much more divergent across all trees. The clustering of Azerbaijani sequences together with Turkish sequences and the divergence of Iraqi and Pakistani strains indicate uneven degrees of connectivity and isolation. The inclusion of reference sequences for Lineages 1 and 2 in all trees provides a global perspective. Most Turkish and some Azerbaijani and Iranian sequences align closer to these references, indicating their involvement in globally significant lineages. However, strains from Iraq and Pakistan tend to show greater divergence from these lineages.

3.6 Comparative Phylogenetic Implications and Interpretations

In this study, the analyzed sequences identified hosts of WNV in this region. A total of 32 sequences from Culicidae hosts collected from Iran and Turkey were considered. Twelve sequences were from *Culex*. Among the *Culex* hosts, eight sequences were from *Culex pipiens*, whereas the others were from *Culex perexiguus*, *Culex sitiens*, and *Culex theileri*. In addition, 27 sequences were associated with human hosts. In Turkey, there were also sequences from *Corvus cornix, Ochlerotatus caspius, Aedes albopictus, Clanga pomarina, Equus caballus*, and *Pica pica*. WNV sequence analysis in Iran identified Anatidae and *Anopheles stephensi* as hosts. Anatidae was also identified in Iraq. In the UAE, Camelus was confirmed as a host. Reports from Azerbaijan confirmed *Alectrobius capensis*, *Aves*, and *Turdus merula* to be hosts. These results provide evidence for the capacity of the virus to cycle through different host species and the ability to cycle through different environments, thereby emphasizing targeted surveillance. A thorough examination of these hosts could enhance our understanding of virus transmission pathways.

How strains of WNV cluster genetically provides useful information about the evolution of the virus, and helps with outbreak planning. Strains from Pakistan, Iraq, Iran, Turkey, and Azerbaijan exhibit strong genetic relatedness, with the genetic disparity of these strains suggesting that we will need to tailor surveillance and response efforts based on where we are targeting our efforts. Local conditions, including climate, mosquitoes, and the wildlife hosting WNV, provide critical context that also determines the spread of WNV related to the above factors. These conditions will affect how transmission occurs and how well surveillance and control measures can be implemented. Recognizing the genetic differences can help improve surveillance and prevention measures in the region. Surveillance programmes should consider the genetic groupings in addressing the surveillance and diagnosis of WNV infections. The clustering of Iranian and Turkish strains confirms the possibility for cross-border communication on the monitoring of the virus. Countries that harbour similar WNV strains may benefit from joint surveillance initiatives designed to monitor mutations of the virus and coordinate preventative measures. The degree of divergence of the strains in Iraq and Pakistan suggests local ecological factors must be taken into account to formulate independent control strategies linking vector control interventions,

management, and warning systems. The identification of genetically differentiated strains emphasizes the importance of regionally optimized vector control programs. Indeed, genetic diversity could influence virus-host interactions, and in turn, reduce transmission efficiency and the efficacy of existing mosquito control efforts. Health authorities need to revisit existing strategies to ensure they adequately consider the diversity of circulating WNV strains being encountered in different areas.

West Nile virus has been gaining attention as a global health concern. The virus was once restricted to certain areas of the globe but has recently become widespread. Factors contributing to its spread include the migration of infected birds, climate change, and changes in mosquito populations that have allowed transmission within many different ecosystems (8, 20, 27). The analysis of phylogenetic data is critical for understanding the evolution and spread of WNV. Phylogenetic data reveal genetic relationship patterns of transmission and variation in the different distributions of WNV in different regions and can be used to track outbreaks, understand the origin of WNV, and support public health efforts. This work is especially important in locations where WNV has increased, such as Iran and surrounding countries (21). This phylogenetic analysis investigated the relationships between WNV transmission patterns in various parts of Iran and other neighboring countries. By comparing sequences of the complete genome, capsid, envelope protein, and NS5, we identified several important patterns relevant to epidemiological surveillance. The results provide a significant understanding of the linkages between bird and mosquito hosts across countries and WNV evolution.

For several years, WNV has been endemic to the Middle East, with documented outbreaks being observed in countries such as Turkey and Iran (28, 29). In Turkey, the virus has been reported since the 1970s (28). WNV cases were reported in the 1970s in Iran (29). There have been many reports of this virus in Iran since then (25). WNV has also been reported in many other countries surrounding Iran, such as Iraq(30), the United Arab Emirates (23), and Pakistan (24).

The phylogenetic analysis of WNV strains in Iran and neighboring regions shows many similarities with patterns seen in Europe. As described by Simonin et al., birds play a critical role in the dissemination of WNV in both locations, with their migration having been found to affect patterns of transmission across regions. As in Europe, WNV circulation in Iran is largely ecological, with wetlands serving as the focal point of virus amplification, as these wetland features contain and support populations of mosquitoes, plus the proper ecological space for the virus to amplify. Although WNV strains from Europe demonstrate genetic variation generated from repeated introductions of the virus, along with endemic evolution, strains of WNV in Iran appear to be less genetically diverse than those seen in Europe (8). The results from Iran and neighboring countries are consistent with a recent investigation in Europe showing how temperature, precipitation, and habitat suitability moderate viral circulation. Both studies sufficiently demonstrate that bird migration and mosquito vector populations are important determinants for the continuation of WNV transmission cycles. The study by Erazo et al. demonstrated that WNV expansion is also mostly related to climate change (19). However, limited sequences and analysis in Iran and some neighboring countries limit this comparison.

In Iran, WNV has many different vectors and hosts that are considered important for the virus cycle of transmission. Among mosquitoes, many species, such as *Culex pipiens, Culex theileri,* and *Culex perexiguus. Culex pipiens, Aedes caspius* are indeed universal transmission vectors. Birds, particularly *Anas sparsa* (*Cape teal*), are also critical since they can act as reservoirs of viruses within nature and cause human transmission. Other mammals, such as humans, pigs, cattle, sheep, and dogs, have been identified as infected hosts (6). Various hosts in Iran were identified in the sequences analyzed in this study. These included species from the family Culicidae, such as *Culex pipiens, Culex sitiens, Culex theileri,* and *Culiseta longiareolata*; other hosts included Anatidae, *Anopheles stephensi*, and *Homo sapiens*. *Culex pipiens*, an important vector of the West Nile virus, is the most common mosquito species in some areas of Iran. The presence of *Cx. pipiens* in these areas would greatly increase the probability of transmission of West Nile virus to humans (6). In Europe, similar to this study, bird species such as Anatidae and some vector species, such as *Anopheles* species, have been identified as hosts (31). Although some avian species, such as *Feral Pigeons* (*Columba livia domestica*), do not show much evidence of WNV detection, reports from a few studies have revealed that many birds in Iran show positive evidence of WNV circulation (25, 32, 33).

The main mosquito vectors of WNV in Turkey are *Culex pipiens, Culex quinquefasciatus, Culex perexiguus, and Ochlerotatus caspius*. other reservoirs of the virus are ducks (Anas platyrhynchos), while incidental hosts with high exposure include but are not limited to sheep, horses, or humans (22). Iran and Turkey share common hosts and vectors for WNV, as shown by the phylogenetic analysis. *Culex pipiens* is an important mosquito vector in both regions.

Culex pipiens and *Culex quinquefasciatus* have been defined as the predominant WNV mosquito vectors in the United Arab Emirates (34). In southern Iraq, both species have also been shown to be WNV vectors (30). Joseph et al. reported the presence of WNV in dromedary camels in the United Arab Emirates (23). An analysis of the sequences from this study revealed that *Camelus* was the host in the United Arab Emirates, whereas Anatidae was the host in Iraq. In Pakistan, several *Culex* mosquito species act as significant vectors for WNV (24). This finding highlights the importance of *Culex* mosquitoes in the WNV cycle.

In Iran, WNV has been associated with a number of lineages, of which lineage 2 has been the most commonly reported. Phylogenetic analysis revealed that the Russian, Romanian, and Italian WNV strains presented some closeness in terms of genetic similarity (6). Lineage 2 was strongly associated with WNV in our phylogenetic trees. The predominant lineages varied among the countries that we studied. In Turkey and Pakistan, lineage 1 is considered a major lineage (22, 24, 35). Lineage 2 has also been seen in Turkey (36). Although a study demonstrated the existence of WNV lineage 1 in the UAE, few studies regarding WNV lineages have been conducted there (23). Iraq and Azerbaijan also lacked data for particular lineages. Our NS5 and capsid phylogenetic trees indicated that sequences from Iraq clustered with lineage 2. The envelope phylogenetic tree, however, showed some clustering of Iraq strains with lineage 1 and Turkey strains. The Azerbaijan strains analyzed in our study also presented some associations with lineage 1. This study and previous work strongly emphasize the need for more investigations to define the geographical impacts of WNV lineages.

WNV likely entered Iran via two migratory bird pathways. First, the virus potentially arrived in central and southern Iran via East Africa--West Asia Flyway when it was introduced from Central or East Africa. Second, WNV was likely to have been introduced into northwestern and western Iran from Eastern Europe, likely Romania or Russia, via the Black Sea-Mediterranean Flyway. Genetic analyses suggest that they were introduced from separate origins. Although there is evidence of similarities in WNV strains in Iran and neighboring countries, geographical proximity alone is unlikely to play a substantive role in the distribution of the virus. Instead, migratory bird pathways are responsible for WNV epidemiology (6). Environmental factors, such as proximity to wetlands and climate, impact WNV transmission. The climate and wetlands of Iran attract migratory birds and lead to conditions conducive to WNV transmission. Although studies have focused on humans, little work has been conducted to determine the distribution of WNV in avian species in Iran. Given the recent increase in outbreaks in neighboring countries and the intercontinental migratory patterns of birds into Iran, it is important to implement effective control measures (25).

Additionally, Iran's health system should prioritize mosquito control planning and put programs into action, particularly for WNV vectors. The WNV surveillance in Iran is missing mosquito control as well as diagnostics. WNV vector targeted management should focus on *Culex pipiens*, *Culex theileri*, and *Culex perexiguus*, and be implemented through larviciding, insecticide resistance survey, and environmental management of stagnant waters. Improved mosquito control will help limit the potential for WNV outbreaks and improve the health of at-risk people. Increasing clinician awareness and implementing WNV-specific diagnostic tests are important in hospitals to increase the detection and care of patients diagnosed with WNV (6). WNV surveillance of migrating birds, an important transporting factor, also needs attention (25). In addition, our study emphasizes the need to increase sampling in Iran and neighboring countries to better understand host-vector dynamics. Further genome sequencing of mosquitoes and birds may shed further light on the genetic variation of WNV and its spread to new regions. Some additional knowledge that may support the spread of the virus can be gained by identifying the biogeographic links between hosts in Iran and those in neighboring regions. Owing to the important role of hosts in the spread of WNV, we suggest considering other ecological parameters related to interactions between bird species, mosquito vectors, and environmental factors in Iran. Collaboration with neighboring countries will shed more light on the ecological and genetic determinants of WNV transmission. Researchers have suggested that WNV surveillance can be improved in Africa through expanding genomic monitoring in strategic locations, connecting human, animal, and environmental data to better predict when outbreaks will occur, and enhancing global collaboration through international databases. These recommendations aim to mitigate existing surveillance gaps and improve overall preparedness (37). Another recent study in Europe indicated that by combining indicator-based surveillance (IBS) and event-based surveillance (EBS), researchers can improve risk mapping for WNV. While IBS is based on structured reports of activity through official sources, EBS focuses on outbreak signals through unofficial sources, such as news articles and expert networks. When both sources are integrated with environmental data, the research findings identified additional areas of concern in Spain, Italy, France, and Greece that were not identified by the traditional approaches (38). An integrated approach will provide the highest

likelihood of control for WNV disease, especially when mixed, synergistic techniques, such as physical barriers, genetically engineered mosquitoes, and immunological targets, are utilized (39).

The study has some biases that may limit conclusions. the unequal sampling of WNV sequences may over-represent Turkey's genetic diversity and under-represent strains from other countries such as Iran. Also, the geographical biases introduced by exclusively using sequences within specific locations may miss genetic variation present in regions that were undersampled. Because fewer sequences were sampled from Iran than Turkey, the study may have also underestimated the diversity of strains within Iran and strains connecting Iran with its neighboring regions. overall, the sampling distribution was not equal for sequences collected across regions. All things considered, sampling across all the countries for WNV completely reasonably considers biases, and this expectation of broader sampling may improve phylogenetic studies of viruses to resolve the conclusions while improving regional surveillance measures moving forward.

4. Conclusion

The phylogenetic analysis revealed the importance of understanding regional dynamics in WNV transmission. The genetic diversity and regional connectivity documented in our study are important sources of information for epidemiological studies, helping in the prediction of WNV spread and the identification of areas that can be monitored and controlled. The analysis of the four phylogenetic trees revealed relationships among WNV strains in the region. Turkey has a central position in the region. Azerbaijan sequences are genetically very close to those of Turkey. The Iranian strains occasionally cluster with the Turkish strains but also occasionally manifest distinct clusters, suggesting differentiation in their genetic pathways. These results were somewhat constrained by the lack of genome sequences. An expanded study in these areas will shed more light on this topic. These findings indicate an urgent need for more research into the factors driving these patterns.

Acknowledgments

This study was made possible with the expertise and guidance of Professor Staji, whose advice was truly valuable. Appreciation also goes to Semnan University for providing the necessary support and environment for this research.

Author's contributions

S.F.A. was responsible for the acquisition of data, analysis and interpretation of data, drafting of the manuscript, statistical analysis, and administrative, technical, and material support.

H.S. was responsible for the study concept and design, critical revision of the manuscript for important intellectual content, and study supervision.

Study concept and design: H.S.

Acquisition of data: S.F.A.

Analysis and interpretation of data: S.F.A.

Drafting of the manuscript: S.F.A.

Critical revision of the manuscript for important intellectual content: H.S.

Statistical analysis: S.F.A.

Administrative, technical, and material support: S.F.A.

Study supervision: H.S.

Disclosure of interest

On behalf of all coauthors, I hereby confirm that I have reviewed and complied with the relevant Instructions to Authors, the Ethics in Publishing policy, and Conflicts of Interest disclosure.

Author Disclosures

The article I have submitted to the journal for review is original, has been written by the stated authors, and has not been published elsewhere.

The article is not currently being considered for publication by any other journal and will not be submitted for such review while under review by this journal.

The article contains no libelous or other unlawful statements and does not contain any materials that violate any personal or proprietary rights of any other person or entity.

I have obtained written permission from copyright owners for any excerpts from copyrighted works that are included and have credited the sources in my article.

If the article was prepared jointly with other authors, I have informed the coauthor(s) of the terms of this publishing agreement and that I am signing on their behalf as their agent, and I am authorized to do so.

Data availability

The sequences of WNV are available in the NCBI virus database

Financial Disclosure

No financial support or sponsorship was received for this study.

Funding/Support

No financial support or sponsorship was received for this study.

Ethics

This study did not involve human or animal subjects. The virus gene sequences analyzed were sourced from the National Center for Biotechnology Information (NCBI), using publicly available data under ethical research practices.

Conflict of interest

The authors declare no conflicts of interest related to this study.

References

1. Calisher CH, Karabatsos N, Dalrymple JM, Shope RE, Porterfield JS, Westaway EG, Brandt WE. Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. J Gen Virol. 1989;70 (Pt 1):37-43.
2. Shomaker TS, Green EM, Yandow SM. Perspective: One Health: A Compelling Convergence. Academic Medicine. 2013;88(1):49-55.
3. Suthar MS, Diamond MS, Gale Jr M. West Nile virus infection and immunity. Nature Reviews Microbiology. 2013;11(2):115-28.
4. Sayed-Ahmed M. Incidence History of West Nile Virus in Africa and Middle East, With an Emphasis on Egypt: A Review. Journal of Dairy, Veterinary & Animal Research. 2016;3::00080.
5. Fiacre L, Pagès N, Albina E, Richardson J, Lecollinet S, Gonzalez G. Molecular Determinants of West Nile Virus Virulence and Pathogenesis in Vertebrate and Invertebrate Hosts. International Journal of Molecular Sciences. 2020;21(23):9117.
6. Shahhosseini N, Moosa-Kazemi SH, Sedaghat MM, Wong G, Chinikar S, Hajivand Z, et al. Autochthonous Transmission of West Nile Virus by a New Vector in Iran, Vector-Host Interaction Modeling and Virulence Gene Determinants. Viruses. 2020;12(12):1449.
7. Riccetti N, Fasano A, Ferraccioli F, Gomez-Ramirez J, Stilianakis NI. Host selection and forage ratio in West Nile virus–transmitting Culex mosquitoes: Challenges and knowledge gaps. PLOS Neglected Tropical Diseases. 2022;16(10):e0010819.
8. Simonin Y. Circulation of West Nile Virus and Usutu Virus in Europe: Overview and Challenges. Viruses. 2024;16(4):599.

9. Ronca SE, Ruff JC, Murray KO. A 20-year historical review of West Nile virus since its initial emergence in North America: Has West Nile virus become a neglected tropical disease? PLoS neglected tropical diseases. 2021;15(5):e0009190.
10. Koch RT, Erazo D, Folly AJ, Johnson N, Dellicour S, Grubaugh ND, Vogels CBF. Genomic epidemiology of West Nile virus in Europe. One Health. 2024;18:100664.
11. Bondre VP, Jadi R, Mishra A, Yergolkar P, Arankalle V. West Nile virus isolates from India: evidence for a distinct genetic lineage. Journal of General Virology. 2007;88(3):875-84.
12. Rizzoli A, Jimenez-Clavero MA, Barzon L, Cordioli P, Figuerola J, Koraka P, et al. The challenge of West Nile virus in Europe: knowledge gaps and research priorities. Euro Surveill. 2015;20(20).
13. Eisen JA. Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. Genome research. 1998;8(3):163-7.
14. Staji H, Badagliacca P, Salehi TZ, Lopes F, Iorio M, Tonelli A, Masson L. Pathotyping of diarrhoeagenic cattle Escherichia coli strains isolated in the Province of Tehran, Iran. Vet Ital. 2017;53(4):345-56.
15. Mehrabifar N, Staji H, Keywanlou M, Salimi Bejestani M, Gallehdar Kakhki E. Identification and Genotyping of Anaplasma phagocytophylum Strains with Zoonotic Potential in Dogs from Mashhad Shelters, Khorasan-Razavi Province, Iran. Iranian Journal of Medical Microbiology. 2022;16(3):244-50.
16. Staji H, Yousefi M, Hamedani MA, Tamai IA, Khaligh SG. Genetic characterization and phylogenetic of Anaplasma capra in Persian onagers (Equus hemionus onager). Veterinary Microbiology. 2021;261:109199.
17. Saffar Benis O, Salimi Bejestani M, Staji H, Rassouli M. Molecular detection of the species Ehrlichia ewingii among dogs referred to veterinary clinics in Semnan, Iran. Journal of Veterinary Laboratory Research. 2021;12(2):157-64.
18. Mohajer F, Sheikh Y, Staji H, Keyvanlou M, Hashemzadeh H. Evaluation of the Seroprevalence of AKABANE and Bluetongue viruses using competitive-ELISA in dairy cattle from industrial herds, Semnan suburb, Iran. Iranian Veterinary Journal. 2019;15(3):78-84.
19. Erazo D, Grant L, Ghisbain G, Marini G, Colón-González FJ, Wint W, et al. Contribution of climate change to the spatial expansion of West Nile virus in Europe. Nature Communications. 2024;15(1):1196.
20. Centrone F, Procacci R, Melilli R, Orlando VA, Colella V, Loconsole D, et al. First seroprevalence study of West Nile Virus (WNV) infection in blood donors after the upsurge of West Nile Neuroinvasive Disease (WNND) cases in southern Italy in 2023. BMC Infectious Diseases. 2025;25(1):200.
21. Lu L, Zhang F, Oude Munnink BB, Munger E, Sikkema RS, Pappa S, et al. West Nile virus spread in Europe: Phylogeographic pattern analysis and key drivers. PLoS Pathogens. 2024;20(1):e1011880.
22. Ergunay K, Gunay F, Erisoz Kasap O, Oter K, Gargari S, Karaoglu T, et al. Serological, molecular and entomological surveillance demonstrates widespread circulation of West Nile virus in Turkey. PLoS neglected tropical diseases. 2014;8(7):e3028.
23. Joseph S, Wernery U, Teng JL, Wernery R, Huang Y, Patteril NA, et al. First isolation of West Nile virus from a dromedary camel. Emerging Microbes & Infections. 2016;5(1):1-12.
24. Zohaib A, Niazi SK, Saqib M, Sajid MS, Khan I, Sial A-u-R, et al. Detection of West Nile virus lineage 1 sequences in blood donors, Punjab Province, Pakistan. International Journal of Infectious Diseases. 2019;81:137-9.
25. Sharti M, Amouakbari MJ, Pourjabari K, Hashemzadeh MS, Tat M, Omidifar A, Dorostkar R. Detection of West Nile virus by real-time PCR in crows in northern provinces of Iran. Journal of Vector Borne Diseases. 2020;57(1):37-9.
26. Shah-Hosseini N, Chinikar S, Ataei B, Fooks AR, Groschup MH. Phylogenetic analysis of West Nile virus genome, Iran. Emerg Infect Dis. 2014;20(8):1419-21.
27. Carrasco L, Utrilla MJ, Fuentes-Romero B, Fernandez-Novo A, Martin-Maldonado B. West Nile Virus: An Update Focusing on Southern Europe. Microorganisms. 2024;12(12):2623.

28. Meço O. [West Nile arbovirus antibodies with hemagglutination inhibition (HI) in residents of Southeast Anatolia]. Mikrobiyol Bul. 1977;11(1):3-17.
29. Naficy K, Saidi S. Serological survey on viral antibodies in Iran. Trop Geogr Med. 1970;22(2):183-8.
30. Barakat AM, Smura T, Kuivanen S, Huhtamo E, Kurkela S, Putkuri N, et al. The Presence and Seroprevalence of Arthropod-Borne Viruses in Nasiriyah Governorate, Southern Iraq: A Cross-Sectional Study. Am J Trop Med Hyg. 2016;94(4):794-9.
31. Nikolay B. A review of West Nile and Usutu virus co-circulation in Europe: how much do transmission cycles overlap? Transactions of The Royal Society of Tropical Medicine and Hygiene. 2015;109(10):609-18.
32. Staji H, Keyvanlou M, Geraili Z, Shahsavari H, Jafari E. The First Study of West Nile Virus in Feral Pigeons (Columba livia domestica) Using Conventional Reverse Transcriptase PCR in Semnan and Khorasane-Razavi Provinces, Northeast of Iran. J Arthropod Borne Dis. 2021;15(1):136-42.
33. Bakhshi H, Beck C, Lecollinet S, Monier M, Mousson L, Zakeri S, et al. Serological evidence of West Nile virus infection among birds and horses in some geographical locations of Iran. Vet Med Sci. 2021;7(1):204-9.
34. Camp JV, Karuvantevida N, Chouhna H, Safi E, Shah JN, Nowotny N. Mosquito biodiversity and mosquito-borne viruses in the United Arab Emirates. Parasites & Vectors. 2019;12(1):153.
35. Yıldırım Y, Yılmaz V, Yazıcı K, Öziç C, Ozkul A, Çağırgan AA. Phylogenetic analysis of West Nile virus: first report of lineage 1 in donkey in Turkey. Tropical Animal Health and Production. 2021;53(5):453.
36. Erdogan Bamac O, Cizmecigil UY, Mete A, Yilmaz A, Aydin O, Tali HE, et al. Emergence of West Nile virus lineage-2 in resident corvids in Istanbul, Turkey. Vector-Borne and Zoonotic Diseases. 2021;21(11):892-9.
37. Moir M, Sitharam N, Hofstra M, Dor G, Mwanyika G, Ramphal Y, et al. Identifying genomic surveillance gaps in Africa for the global public health response to West Nile Virus. bioRxiv. 2025:2024.12. 18.629123.
38. Serres K, Erazo D, Despréaux G, Vincenti-González MF, Van Bortel W, Arsevska E, Dellicour S. Integrating indicator-based and event-based surveillance data for risk mapping of West Nile virus, Europe, 2006 to 2021. Euro Surveill. 2024;29(44).
39. Cendejas PM, Goodman AG. Vaccination and control methods of West Nile virus infection in equids and humans. Vaccines. 2024;12(5):485.