Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Victoria L. Cowger Date

The role of supplementary environmental surveillance to complement acute flaccid paralysis surveillance for wild poliovirus in Pakistan and Afghanistan – January 2011 through September 2013

By

Victoria L. Cowger

MPH

Department of Epidemiology

Walter A. Orenstein, MD
Committee Chair – Faculty Advisor

Cara Burns, Ph.D
Committee Chair – Field Advisor

The role of supplementary environmental surveillance to complement acute flaccid paralysis surveillance for wild poliovirus in Pakistan and Afghanistan – January 2011 through September 2013

By

Victoria L. Cowger

Bachelors of Science
University of Minnesota – Twin Cities
2012

Thesis Committee Chairs: Walter A. Orenstein, MD and Cara Burns, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2014

ABSTRACT

The role of supplementary environmental surveillance to complement acute flaccid paralysis surveillance for wild poliovirus in Pakistan and Afghanistan – January 2011 through September 2013

By Victoria L. Cowger

Introduction.
Since 1988, poliomyelitis incidence has decreased more than 99% worldwide. The current gold-standard for poliovirus surveillance is clinical Acute Flaccid Paralysis (AFP) surveillance; however, there is evidence of decreased sensitivity of AFP surveillance with decreasing infection prevalence. Environmental surveillance (ENV) can detect circulating polioviruses from sewage excreted in stool without relying on clinical presentation of disease, making it a potentially powerful complement to AFP surveillance. Because of the extensive ENV and continued endemicity, Pakistan presents a unique opportunity to quantify the role of ENV as a supplement to AFP surveillance alone.

Methods.
Teams at the U.S. Centers for Disease Control and Prevention (CDC) provided genetic, geographic and temporal data for Pakistan and Afghanistan poliovirus isolates from January 2011 through September 2013. Descriptive analysis was conducted to quantify WPV positive isolates by surveillance type, year and country. AFP and ENV isolates were mapped by year, location, and genetic cluster to compare genetic distribution of isolates detected by respective surveillance types. Pakistan AFP isolates were classified into sub-lineages based genetic match to other Pakistan AFP isolates and similar genetic circulation preceding and following sub-lineage isolates were plotted and circulation time for both ENV and AFP surveillance was quantified.

Results.
A total of 803 WPV isolates were detected by AFP and ENV surveillance in Pakistan and Afghanistan from January 2011 through September 2013. In six of the first seven months of 2011, ENV detected poliovirus in samples in three provinces that did not detect any polio cases, suggesting silent transmission in these areas. Overall, ENV detected circulation first in more sub-lineages than did AFP (51.7% vs. 26.3%), and ENV detected circulation approximately 2 months sooner on average for each sub-lineage than did AFP surveillance.

Discussion.
This study presents evidence that suggests ENV in Pakistan is providing earlier and more sensitive detection than AFP surveillance alone. Overall, targeted ENV through strategic selection of sites has proven useful in Pakistan, and has important applications in the eradication Endgame strategy, including detecting cVDPVs, monitoring the switch from live to inactivated polio vaccine, and certification of a polio-free world.

The role of supplementary environmental surveillance to complement acute flaccid paralysis surveillance for wild poliovirus in Pakistan and Afghanistan – January 2011 through September 2013

By

Victoria L. Cowger

Bachelors of Science
University of Minnesota – Twin Cities
2012

Thesis Committee Chairs:
Walter A. Orenstein, MD (Faculty Advisor)
Cara Burns, Ph.D. (Field Advisor)

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2014

Acknowledgements

I would like to thank my faculty adviser, Dr. Walt Orenstein, for providing invaluable direction for the project, feedback on methods, analyses and drafts, and sharing with me his expertise and passion for polio eradication. I would also like to thank Dr. Cara Burns for her consistent support and guidance throughout the project and for sharing her immense knowledge of polio genetics and surveillance. Thanks to Dr. Howard Gary for not only sharing constructive and useful feedback on drafts, but for providing statistical and methodological advice.

I would like to thank Jane Iber and Beth Henderson for their ability to translate my abstract descriptions of my planned analyses into complicated MatLab algorithms to generate workable datasets. Special thanks to the National Institute of Health Laboratory in Pakistan, including Salmaan Sharif, who collected, processed and sequenced the samples used in this study. And thanks to Dr. Mark Pallansch and Sara Lowther for their participation and feedback in the planning and analysis of this project.

Finally, I would like to thank my fiancé, family and friends, especially my late grandparents Sally and Robert Cowger, for always encouraging me to exercise my love for science and humanity, and for reminding me that "to whom much is given, much is expected."

TABLE OF CONTENTS

Distribution Agreement

ABSTRACT

BACKGROUND/LITERATURE REVIEW **1**

Overview of Polio and poliovirus. **1**

History: *1*

Virology: *2*

Clinical features: *3*

Polio vaccines: *4*

Polio epidemiology and eradication **6**

Surveillance for Polio **7**

Acute Flaccid Paralysis surveillance *8*

Environmental Surveillance *9*

Epidemiology and Environmental surveillance for poliovirus in Pakistan **11**

METHODS **14**

Data: *14*

Genetic classification: *14*

Descriptive analysis: *15*

Orphan lineage analysis: *16*

Sub-lineage analysis: *16*

RESULTS **19**

Descriptive analysis *19*

Orphan lineage analysis *24*

Sub-lineage analysis *24*

DISCUSSION **27**

Detection above AFP surveillance alone *27*

Strengths and weaknesses of analyses: *30*

Limitations and considerations for implementation of ENV *32*

Role of environmental surveillance in endgame strategy *34*

Conclusions *36*

REFERENCES **38**

TABLES **41**

FIGURES AND FIGURE LEGENDS **50**

APPENDICES **66**

BACKGROUND/LITERATURE REVIEW

Overview of Polio and poliovirus.

History:

Descriptions of paralytic diseases compatible with that of poliomyelitis pre-date written history, with exhumation of 19th dynasty Egyptian priest, Ruma, showing shortening and withering of his right leg, buried with a crutch (1580-1350 BC) (1). Poliomyelitis was first described by 18th century physician, Michael Underwood as "debility of the lower extremities in children"(2). The first outbreaks of polio in Europe were recorded during the early 19th century, with the first outbreaks in the United States recorded in 1843 (3). In the late 19th century, outbreaks of infantile paralysis appeared abruptly in the U.S. and several Scandinavian countries, which were attributed to increasing age at infection (4). Prior to the advent of public sanitation and improvements in personal hygiene, ubiquitous exposure to enteric infections such as poliovirus were thought to have occurred during early infancy, while passively acquired maternal antibodies were still present, precluding viremia and preventing paralysis, but not impeding acquisition of active immunity (5). It is hypothesized that improved sanitation and hygiene measures delayed age of first exposure to enteric disease to late infancy or early childhood after passive immunity had waned (5). Observational evidence corroborates this hypothesis, marked by increased incidence from 1890-1950; outbreaks occurred earlier in more developed countries, while outbreaks were experienced later in developing countries following public health improvements (6).

Outbreaks occurred seasonally in the United States beginning in the early 1900s until declining substantially beginning in 1955 following the licensure of inactivated polio vaccine

(IPV), with further declines observed following introduction of oral polio vaccine (OPV) in 1961(3).

Virology:

Polioviruses are enteroviruses in the Picornavirus family of RNA viruses, sharing many biochemical properties with non-polio enteroviruses; they are resistant to inactivation by common detergents including some soaps, and are stable at an acidic pH for 1-3 hours. Polioviruses have an RNA genome about 7,500 nucleotides long, and can be inactivated by formaldehyde, chlorine, or UV light (7, 8). Polioviruses can remain infective for days at 30ºC, allowing them to persist in environments with adequate humidity (7).

Wild polioviruses (WPV) are classified into three antigenic types, WPV1, WPV2, and WPV3, with very limited cross-immunity introduced against infection by other serotypes. Complete genomic sequences for all three serotypes and three Sabin OPV strains have been determined, although only the sequences that encode its capsid proteins are genetically unique, with the remainder of the genetic material exchanged through frequent recombination with non-polio enteroviruses. Consequently, sequencing of the four capsid proteins (VP1 through VP4) is used for virus identification. Polioviruses have relatively simple replication cycles and genetic organization, and mutate at a rate of 10^{-4} substitutions per nucleotide copied, resulting in approximately 1% nucleotide substitutions per year; this exceeds the corresponding rate for DNA viruses (7, 9). Complete genome sequencing of polioviruses has added to the understanding of the epidemiology of poliomyelitis, facilitating determination of linkages between cases and ascertainment of importations from remaining polio reservoirs, and helping to document extinction of different genotypes.

Clinical features:

Typically afflicting children and infants, transmission of poliovirus occurs from person-to-person after the virus enters through the mouth and replicates in the pharynx and gastrointestinal tract, sometimes infecting local lymphatic tissue and causing viremia (3, 7). The virus can subsequently infect central nervous system cells, where replication in motor neurons of the anterior horn and brainstem result in cell destruction, causing polio's hallmark, acute flaccid paralysis without permanent sensory loss (3, 4). The vast majority of poliovirus infections are inapparent, with up to 72% of polio infections being asymptomatic, and less than 1% resulting in paralysis(8). Minor, non-specific illnesses occur in 24% of cases, and may present as three syndromes indistinguishable from other viral illnesses – respiratory tract infection, including sore throat and fever; gastrointestinal illness, such as nausea, vomiting and abdominal pain; and influenza-like illness. Non-paralytic aseptic meningitis, followed by complete recovery occurs in 4% of infections(3, 8). Estimates of the ratio of poliovirus infection to paralytic illness vary widely from 500:1 to 50:1 and depend on age and virus serotype, with highest paralytogenicity reported in WPV1 infections(5). On average for all serotypes, estimates suggest that only 1 in approximately 150 infections cause paralytic poliomyelitis (4, 5). Average incubation period from infection to onset of paralysis is 10 days, but can range from 5 to 25 days (3, 5, 10). Paralytic symptoms occur 1-10 days after prodromal symptoms and progress for 2-3 days, associated with fever. With defervescence, further paralysis stops. Although many recover some degree of muscle function, weakness or paralysis still present after 12 months is usually permanent (3, 8). Polio typically involves only the motor system and sensory deficits are usually not part of the clinical syndrome.

Polio vaccines:

Poliovirus is highly communicable and could infect nearly every individual in a completely unimmunized population(7). As there is no cure for polio, vaccination has been at the cornerstone of eradication efforts, with two available vaccines – inactivated polio vaccine (IPV), licensed in 1955 and oral polio vaccine (OPV), licensed in 1963 (3). OPV, used for routine immunization and in some mass campaigns, is a trivalent vaccine containing live-attenuated Sabin strains of poliovirus types 1, 2 and 3, and it has been the main tool in polio eradication because it confers high levels of both humoral and intestinal immunity, is inexpensive and easy to administer, and can provide protection to close contacts of vaccine recipients through secondary spread of vaccine virus(3). OPV is especially effective at producing intestinal immunity to poliovirus, with a single dose conferring immunity in approximately 50% of recipients, and in developed countries, seroconversion was observed in more than 95% of infants after 3 doses(3, 11). However, studies in tropical developing countries have suggested that OPV is less effective in these settings, conferring shorter duration of intestinal immunity and lower seroconversion rates to all virus types after 3 doses (73%, 90% and 70% to types 1, 2 and 3, respectively) (11, 12). Therefore, in tropical developing areas, >3 doses are required to achieve herd immunity levels to terminate transmission. There are a variety of risk factors for vaccine failure. Type 2 virus interferes with the immunogenicity to types 1 and 3, and bivalent vaccine eliminating type 2 has become the standard for mass immunization campaigns, known as "Supplemental Immunization Activities" (SIAs) since WPV2 appears to have been eradicated with the last naturally occurring case detected in 1999. To improve effectiveness of OPV, monovalent type 1 vaccine (mOPV1) and type 3 (mOPV3) were introduced in 2005 in remaining reservoirs in Asia and Africa, and shown to be more immunogenic than trivalent vaccine

against respective serotypes(11). After use of mOPV1 in campaigns from 2005-2008 promoted resurgence of WPV3, bivalent OPV (bOPV) containing types 1 and 3 was licensed after showing non-inferiority in immunogenicity as compared with respective monovalent vaccines(11).

Since the advent of The Global Polio Eradication Initiative (GPEI), more than 10 billion doses of OPV have been administered to more than 2.5 billion children throughout the world, preventing an estimated 10 million cases of paralytic polio and 1 million childhood deaths (13). Because OPV is a live-attenuated vaccine, it has the potential to regain neurovirulence as in wild poliovirus (WPV), leading to vaccine-associated paralytic poliomyelitis (VAPP). VAPP occurs in 2-4 cases per million births per year, and with the current use of OPV, 250-500 cases of VAPP are estimated to occur annually(14). As OPV contains live Sabin-strain viruses, vaccines can be shed in stool for up to 6 weeks after immunization, with maximum shedding occurring 1-2 weeks after the first dose(3). Before 2000, it was assumed that vaccine strains shed in stool, despite having adequate neurovirulence, would not have adequate transmissibility to cause outbreaks(14). Since then, it has been shown that accumulations of mutations in the poliovirus genome are adequate to regain sufficient transmissibility to cause outbreaks of circulating vaccine-derived poliovirus (cVDPV) (15). These cVDPVs have regained both the neurovirulence and transmissibility characteristics of WPVs. In 2012, 20 cVDPV outbreaks have occurred, resulting in 580 cases of paralytic polio; more countries suffered an outbreak of cVDPV than outbreaks from wild poliovirus(13).

In contrast, IPV recipients do not shed live-poliovirus in stool and more than 90% of recipients had serum antibodies to all three poliovirus types following 2 doses and at least 99% had antibodies to all three poliovirus types after 3 doses(3). Although effective in

producing humoral immunity and protecting against paralytic polio, IPV produces much less intestinal immunity than does OPV, and therefore IPV recipients are more likely to be infected than OPV recipients and continue to shed virus in stool(3).

Polio epidemiology and eradication

Following the introduction of polio vaccines in the United States, annual incidence of polio fell sharply, with the last cases of poliomyelitis from indigenous virus occurring in 1979(5). Disappearance of polio in the U.S. and successful control programs in Cuba and Brazil demonstrated the feasibility of eradication of wild virus in other regions (5).

In 1988, the forty-first World Health Assembly resolved to eradicate poliomyelitis globally by the year 2000, citing immense progress in routine immunization coverage in developing countries(16). The WHA also encouraged member states to strengthen routine immunization systems and intensify surveillance systems, promoting global collaborations as a means to strengthen health systems. This initiated the launch of the Global Polio Eradication Initiative (GPEI), which developed a polio eradication strategy with four main pillars: 1) Routine immunization; 2) Supplementary immunization, 3) Surveillance; and 4) Targeted mop-up campaigns(13). Since 1988, drastic reductions in polio cases have been observed, decreasing from an estimated 350,000 cases in 1988 to 719 reported cases in 2000 (13).

The Americas was the first WHO region to be certified free from poliovirus circulation in 1994, followed by the Western Pacific in 2000 and the European region in 2002 (17). Most recently, the Southeast Asian region was certified polio free in March of 2014; this region notably includes India, regarded as one of the most challenging locations for polio control, given its population density, geographic location and poor sanitation infrastructure

(17, 18). The last case of paralytic polio from naturally circulating WPV2 occurred in 1999 in Uttar Pradesh, India, with the number of countries with endemic wild poliovirus dropping from 20 in 2000 to 4 in 2009 (5, 19). As of 2014, only three countries are considered endemic (have never interrupted transmission of wild polioviruses), Pakistan, Nigeria and Afghanistan.

Additionally, the genetic diversity of the remaining WPV isolates has decreased substantially, suggesting limited circulation of WPV(19). Completion of eradication however, has proven difficult in endemic areas because of failure to vaccinate, failure of the vaccine, and viral epidemiology (5). High prevalence of diarrheal diseases in resource-limited countries, and interference of virus types within the bivalent and trivalent vaccine can limit OPV's efficacy (5, 11). These issues have combined with epidemiologic variables including geographic location, high population densities, and political instability to complicate eradication efforts in the three remaining reservoir countries (5, 20).

Surveillance for Polio

By the end of 2012, global incidence of polio had decreased by 99.9% from 1988 levels, and the World Health Assembly (WHA) declared completion of polio eradication a programmatic emergency, establishing the Global Polio Emergency Action Plan and development of a comprehensive polio eradication and endgame strategy (13, 21). The Plan identifies four main objectives: 1) poliovirus detection and interruption; 2) immunization systems strengthening and OPV withdrawal; 3) containment and certification; and 4) legacy planning (13). One of the major activities under objective 1, poliovirus detection and interruption, is strengthened global surveillance to detect virus circulation. The main priority

for improved surveillance involves closing remaining gaps in acute flaccid paralysis (AFP) surveillance.

Acute Flaccid Paralysis surveillance

AFP surveillance is a clinical, case-driven system that detects recent paralysis from any cause, followed by subsequent confirmation of poliovirus through laboratory isolation from stool samples (19, 22, 23). AFP surveillance is currently considered the gold standard for polio surveillance, and GPEI has identified several quality indicators to assess aspects of quality of surveillance including completeness of reporting, sensitivity of surveillance, completeness of case investigation, completeness of follow-up and laboratory performance (Table 1). Effective AFP surveillance hinges on clinical recognition of paralysis, notification and investigation of paralysis cases, collection of quality stool specimens (i.e., 2 specimens at least 24 hours apart within 14 days of onset of paralysis), proper transport to laboratory and laboratory processing to confirm polio, with multiple factors impacting the sensitivity of AFP surveillance (22). Overall sensitivity of the AFP surveillance system depends on the following factors: 1) collection of stool specimens at a time shedding is likely (i.e., within 14 days of onset of paralysis; 2) the laboratory's probability of detecting virus in a specimen that contains live virus; 3) probability that stool samples from an infected individual will be collected and processed so that detectable virus is present upon arrival at the lab, impacted both by an individual's viral shedding and proper collection and transport; 4) probability that an infected person will have virus detected from at least one stool sample, which increases as the number of stool samples increases; and probability that at least one infected person will be detected from a population with circulating wild poliovirus (24). Mathematical modeling of the above factors suggests decreased sensitivity of AFP surveillance with decreasing prevalence of infection in the population (24). In a quantitative analysis of the AFP

surveillance system for poliovirus detection in Australia, mathematical models suggested a median sensitivity of 8.2% at a theoretical population prevalence of infection of 1 case per 100,000 population, but sensitivity decreased to nearly 0.9% when infection prevalence is decreased to 1 case per 1,000,000 population (25). Declines in surveillance sensitivity with decreased infection prevalence poses a problem as eradication progresses, and implies that other surveillance strategies may be needed to address the gaps present in AFP surveillance.

Environmental Surveillance

One such supplemental strategy identified in the Emergency Action Plan is environmental surveillance (ENV) to help identify residual transmission in endemic areas (18). The rationale for ENV is based on the natural course of poliovirus infection, as more than 99% of poliovirus infections are without paralysis and therefore, not detected by AFP surveillance. Irrespective of clinical symptoms, virus can be excreted in feces for several weeks to months after infection, and shed into the environment in variable amounts, reaching maximal amounts of 10^7 infectious doses/day per person (26). Since polioviruses are relatively stable in aqueous environments, environmental sampling can detect polioviruses circulating in the community without relying on clinical presentation of disease, making it a potentially powerful tool to complement AFP surveillance alone.

Sensitivity of ENV depends on several factors, including population parameters specific to poliovirus circulation such as number of infected individuals in the population, and duration of excretion, as well as environmental factors such as sewage infrastructure (i.e., whether there is a "central sewage" system), weather conditions, sample size, and frequency measured, that might impact the laboratory's ability to isolate the virus (27). Using simulation models, previous studies have suggested that when assuming the case to infection

ratio is low (<1:200), ENV surveillance may be more efficient than AFP surveillance in detecting circulating poliovirus, especially in areas with high vaccination coverage with IPV(27). However, this study did not recommend ENV as a replacement for AFP, as ENV cell culture isolation and detection is complicated by dilution of minority WPV with Sabin strain virus in populations vaccinated with OPV(27). Additionally, it may be unfeasible and expensive to cover the entire population through ENV alone.

ENV has contributed to documenting the elimination of WPV from Egypt and India, and continues to impact current surveillance strategies in remaining endemic areas(28). Environmental surveillance is currently being conducted in 24 countries – 2 endemic and 22 non-endemic. Endemic countries, Nigeria and Pakistan, are currently conducting ENV in 11 sites in 3 states and 23 sites in four states, respectively (29). Non-endemic states conducting environmental surveillance include India (15 sites in 4 states), Egypt (34 sites in 11 cities), and multiple countries in the WHO European region (29). Historically, ENV has been used as a supplement to AFP surveillance, with evidence to suggest increase in sensitivity and early detection, especially when paralytic polio cases are few. In Finland in the 1980's, after two decades without a case of poliomyelitis, nine cases of paralytic poliomyelitis were detected, with concurrent detection of WPV3 in sewage (30). After the final cases presented, sewage screening continued to detect circulation in areas without clinical cases, and depicted larger geographic dispersion and reservoirs for circulating virus (31). Evidence was used to implement vaccination campaigns in infected areas, and to document the cessation of circulation of WPV3, only possible by ENV in the absence of clinical presentation or selective stool sampling in the population (30, 31).

ENV has also been effective in capturing virus circulation in the absence of clinical cases in Israel and the Palestinian Authority, where the last poliomyelitis cases occurred in

1988. In the 8 years following those cases, between February 1989 and February 1998, four separate episodes of WPV circulation were detected (32). ENV allowed health officials to determine homology between ENV isolates as well as relatedness to previous AFP cases and to document the end of circulation. In this setting, active AFP surveillance would have been insufficient in detecting circulation, and provides evidence of the utility of ENV, especially as eradication efforts progress (33).

ENV in Egypt was established in 2001 and is ongoing, with sites selected in areas of known WPV circulation (34). ENV sites consisted of areas with both converging sewer networks as well as areas where sewage was channeled via pipes into a shared pit, demonstrating the feasibility of conducing ENV in areas with limited sanitation infrastructure(34). Throughout its tenure in Egypt, ENV has been able to identify reservoirs that sustain transmission during the low season, detect importation of WPV1, and isolate VDPV (35, 36).

Implemented at the tail-end of their epidemic, India initiated ENV in Mumbai in 2001. Expansion of ENV to northern India led to detection of poliovirus in wastewater in Delhi linked to 2009 WPV isolates from AFP cases (18, 37). Subsequently, ENV was able to document the elimination of polio in India, with no WPV isolated from any environmental samples since August 2010, suggesting that ENV may become an increasingly useful tool in documenting elimination and eradication in high-risk areas, potentially proving useful in the polio-free certification process(18).

Epidemiology and Environmental surveillance for poliovirus in Pakistan

Along with Nigeria, Pakistan and Afghanistan comprise the only countries with endemic wild poliovirus transmission. Afghanistan and Pakistan experienced increases in reported

WPV cases in 2011, with more than three times as many WPV cases experienced in Afghanistan, and a 37% increase in WPV cases in Pakistan over 2010 case counts. Immunization efforts in Pakistan have been confronted with various challenges; supplemental immunization activities (SIAs) were temporarily suspended in areas of Pakistan in 2012 and the beginning of 2013 because of attacks against health workers and polio immunization was banned by local Pakistani Taliban commanders in some districts in the Federally Administered Tribal Areas (FATA) where the highest incidence of polio is observed (20, 38).

Estimated routine immunization coverage with 3 doses of oral polio vaccine (POL3) in Pakistan was 75% in 2011 and 66% in Afghanistan, with substantial variability within each country and especially low coverage in areas of political instability (20). Overall immunization coverage increased in 2012, with 89% reported OPV3 coverage, and coverage among children with non-polio AFP in Pakistan was 65% nationally (39). Provincial OPV3 coverage in 2011 in Pakistan is as follows: 26% in conflict-affected Federally Administered Tribal Areas (FATA); 63% in Khyber Pakhtunkwa (KP), 52% in Sindh province, 18% in Balochistan province, with highest reported OPV3 coverage in Punjab province (77%) (40). In Pakistan, ENV started in 2009 with two cities in two provinces – Karachi, Sindh Province and Lahore, Punjab Province. ENV was then expanded to other major cities in the country, with 23 sampling sites at the end of 2012 (Figure 1) (28). ENV site locations were selected based on perceived risk for polio circulation, which was based on a combination of the following factors: 1) Sustained transmission for polio previously noted in district; 2) available and appropriate sewer systems; 3) poor sanitation and crowding; 4) insufficient vaccination coverage; or 4) suspicion of sub-optimal or limited sensitivity of AFP surveillance (28, 41). ENV in Pakistan has consistently yielded positive isolates in the absence of AFP cases and

informed direction of actions to enhance polio control measures in affected areas (28, 40). Because of the extensive environmental sampling being conducted and the endemicity of polio in the area, Pakistan presents a unique opportunity to quantitatively investigate the role of environmental surveillance as a supplement to AFP surveillance alone. This paper aims to characterize the benefit of environmental surveillance over AFP surveillance alone in Pakistan, and by extension, Afghanistan, because the poliomyelitis epidemics in both countries are neither genetically, geographically, nor epidemiologically independent.

METHODS

Data:

Data for this study were provided by teams in the Polio and Picornavirus Laboratory Branch (PPLB) in the Division of Viral Diseases, and in the Global Immunization Division (GID) at the U.S. Centers for Disease Control and Prevention (CDC) in Atlanta. Environmental sampling was conducted in accordance with WHO's Guidelines for Environmental Surveillance for Poliovirus by the National Institutes of Health laboratory in Pakistan (42). Virus isolates were prepared from stool specimens and environmental grab samples using the recommended methods of the World Health Organization Global Poliovirus Laboratory Network(43).

The dataset was derived from the completely de-identified and IRB exempt surveillance database maintained by PPLB, and consisted of genetic and geographic information for both clinical AFP samples and environmental samples. The variables that comprised the final dataset and their descriptions can be found in table 2.

All AFP WPV-positive isolates from Pakistan and Afghanistan and all WPV environmental isolates from environmental sampling in Pakistan collected from January 2011 through September 2013 were considered for analysis. Duplicate isolates from the same environmental sampling batch were removed.

Genetic classification:

Sequencing of the complete VP1 capsid protein was conducted using reverse transcription-PCR (34, 44). Both environmental and AFP isolates were classified into genetic clusters by

PPLB, and within genetic clusters, isolates are ≥95% similar in their VP1 capsid protein coding region.

Using the genetic, geographic and temporal data described above, this paper aims to answer the following research questions:

1. Does supplemental environmental surveillance for polio in Pakistan provide added benefit over AFP surveillance alone?
2. Is environmental surveillance able to show silent transmission of poliovirus, linking cases of polio AFP?

Descriptive analysis:

Attempts to characterize the added benefit of ENV surveillance over AFP alone began with basic descriptive analysis that included quantification of WPV positive isolates by surveillance type, year and country. Isolates were then further characterized to enumerate number of provinces and districts considered infected with WPV each month between January 2011 and September 2013. ArcGIS (Esri ® ArcMap™ 10.1) was used to map AFP and ENV isolates by year, location, and genetic cluster to compare genetic distribution of isolates detected by respective surveillance types. Overall, descriptive analysis aimed to point out the places and times when ENV added information or insight into polio circulation that would have otherwise been missed by AFP surveillance alone.

Orphan lineage analysis:

Orphan viruses are defined as isolates that are ≥1.5% different in their VP1 capsid nucleotide sequence from any previously detected isolate, and are used as a measure of completeness of surveillance. Isolation of orphan virus suggests silent circulation without detection for an extended period of time, indicating gaps in surveillance. Consequently, the number of orphan viruses detected by AFP and ENV would allow direct comparison of completeness of surveillance. Considering only AFP isolates with ENV isolates removed from the database, a MatLab software algorithm was used to generate a list of AFP orphan isolates. The same process was conducted using only ENV isolates to identify ENV orphan lineages, and finally to identify both AFP and ENV orphans when considering both AFP and ENV isolates. This list of orphans was imported into SAS® (Version 9.3), and analyzed to compare proportions of orphan viruses detected by each respective surveillance system to determine what proportion of AFP orphans could be linked through ENV.

Sub-lineage analysis:

To specifically address research question 1 above more quantitatively, this study aimed to determine the proportion of AFP cases which would have been detected sooner by ENV surveillance than with AFP surveillance alone. To do so, we aimed to quantify circulation of genetically similar virus, detected by both AFP and ENV surveillance before and after each AFP isolate.

All Pakistan AFP isolates were classified into sub-lineages based on pairwise percent match in VP1 capsid to other PAK AFP isolates. Gephi ® (Version 0.8.2) network visualization software was used to visualize genetic connectivity between Pakistan AFP isolates, restricting connections to pairwise genetic matches ≥99.5%(45). Therefore, isolates

within sub-lineages are ≥99.5% similar with respect to their VP1 capsid genetics to at least one other virus in that lineage. For the 326 Pakistan AFP isolates, network visualization allowed characterization into 209 sub-lineages, based on ≥99.5% VP1 genetic match. Of these 209 sub-lineages, 173 contained only a single AFP isolate, while the remaining 153 isolates were classified into sub-lineages comprised of more than one AFP isolate. Among sub-lineages with more than one isolate, number of isolates per sub-lineage ranges from 2-17 isolates per sub-lineage and a median of 3 isolates. The mean circulation time between first isolate in the sub-lineage to the last isolate was 68.9 days. (95% CI: 41.3, 96.5 days).
After sub-lineages were identified, MatLab was used to determine pairwise matches of genetically similar viruses, both ENV and AFP isolates, defined as ≥99.0% of the VP1 capsid for all isolates in a sub-lineage. For each sub-lineage, these pairwise matches were then plotted and quantified to determine the circulation time and first detection of circulation by either AFP surveillance or ENV. For isolates in sub-lineages with more than one AFP isolate, the median date of isolates was used to determine circulation time before and after. The unit of analysis for subsequent statistical analysis was sub-lineage, and all statistical analyses were completed in SAS.

Overall mean circulation time for each sub-lineage was calculated from the date of first matching isolate (≥99.0%), AFP or ENV, to the date of the last matching isolate for each respective sub-lineage. Mean circulation time before and after sub-lineage median date was calculated for ENV detected circulation alone, and for AFP detected circulation alone. Both AFP and ENV matching isolates (≥99.0%) for one sub-lineage could also be matches for other sub-lineages, therefore, sub-lineages could not be considered statistically independent of one another. As such, statistical tests comparing proportion of first detected isolate type by surveillance system and tests comparing mean circulation times for sub-

lineages violated pre-requisite assumptions of statistical independence and therefore were not considered. Instead, proportions and mean circulation times were compared qualitatively to assess utility of AFP and ENV surveillance systems.

RESULTS

Descriptive analysis

A total of 803 WPV isolates were obtained from samples collected through the AFP and ENV surveillance systems in Pakistan and Afghanistan from January 2011 through September 2013 (Table 3). With the exception of 5 WPV3 isolates from Pakistan (2 in 2011 and 3 in 2012), all detected wild poliovirus was type 1. More than twice the number of AFP isolates was reported in Pakistan than in Afghanistan during these years (326 vs. 129 isolates, respectively). During the same period 348 WPV1 isolates were derived from sewage samples from the 23 ENV sites in Pakistan.

Both ENV in Pakistan and AFP surveillance in Pakistan and Afghanistan showed a greater number of WPV positive isolates in 2011 than the combined total number of isolates for the remainder of the study period. Trends in number of positive isolates between AFP and ENV surveillance were similar, with peaks in number of WPV positive isolates observed from August 2011 to Jan 2012 in both Pakistan and Afghanistan (Table 4, Figure 2). In 2011, 4 provinces were conducing biweekly ENV at 17 sites in Pakistan. During this time, both polioviruses from AFP samples and ENV samples were consistently isolated from Balochistan province, KP province, Punjab province and Sindh province (Table 5). During this time AFP cases were also consistently reported in the Federally Administered Tribal Areas (FATA) and twice in Gilgit-Baltistan province (GB) where no environmental sampling sites are present.

In six of the first seven months of 2011, ENV detected poliovirus in samples in provinces (Balochistan, Punjab and Sindh) that did not detect any polio cases, suggesting silent transmission in these areas (Table 5). In the latter half of 2011, these provinces

experienced a large increase in number of reported polio AFP cases, suggesting that ENV might have detected virus circulation in these provinces before a symptomatic case presented. The remainder of the observational period showed a decrease in the number of provinces experiencing AFP cases. ENV continued to detect circulation in these areas, as evidenced by entire provinces without AFP cases, but detecting WPV in sewage isolates (Table 5).

Concurrent AFP and ENV surveillance in Pakistan from 2011 through September of 2013 both showed decreases in the number of infected districts from 64 in Pakistan and 36 in Afghanistan in 2011, to 23 in Pakistan and 15 in Afghanistan in 2013 (Table 6, Figure 3). In 2011, there were 8 districts conducting ENV in 4 provinces, increasing to 11 districts in 4 provinces in 2012 and 2013. In districts conducting both AFP and ENV, there was limited overlap observed in the number of these districts reporting both positive ENV and AFP isolates with only 3 of 8 (37.5%) in 2011, 3 of 11(27.2%) in 2012 and 1 of 11 (9.0%) in 2013. Figure 4 shows the number of districts positive for WPV by surveillance type, with the limited number of districts experiencing both AFP and ENV positive isolates concurrently, although inherently limited by the number of districts conducting ENV. Even in districts conducing both AFP and ENV, few reported both AFP cases and ENV positive isolates, implying that ENV is detecting circulation in districts not reporting polio cases.

Two ENV sites in Quetta district, Balochistan province detected WPV consistently from January 2011 through January 2012, with WPV positive AFP cases being first reported 6 months later in June 2011 (Figure 5A). Following consistent positives through January 2012, only two AFP cases were reported through September 2013. Results from ENV were consistent with AFP surveillance during 2012 through the first 9 months of 2013, with just 3

positive ENV isolates reported from 3 ENV sites in Balochistan during the same interval (Figure 5A).

Similar trends in both AFP and ENV were reported in Punjab Province, where relatively consistent detection of positive ENV isolates from six sites in three districts suggested circulation before one AFP case was reported in one of the three districts in December of 2011 (Figure 5B). From June 2012 through January 2013, ENV detected positive isolates from 6 ENV sites in 4 districts in the absence of AFP cases. Conversely, no ENV positive isolate was reported from Punjab from February 2013 through August 13, and during that time 5 polio AFP cases were reported from districts not conducting ENV surveillance (Figure 5B).

GB province observed only three AFP cases from January 2011 through September 2013 and no ENV was conducted there during this period (Figure 5C). KP province and conflicted FATA experienced the largest burden of AFP cases throughout the observational period, with ENV not conducted in FATA. However, consistent poliovirus isolation at two ENV sites in KP province, adjacent to FATA, and intermittent AFP cases in this area suggest ongoing transmission.

ENV detected positive isolates from 5 sites in Karachi, Sindh Province from January 2011 to Feb 2012, preceding the 9 AFP cases first reported in this area from March 2011 to November 2012 (Figure 5D1). When Karachi is viewed in finer detail (Figure 5D2), ENV maintains specificity, detecting virus at all ENV sites in Karachi sub-districts – Baldia Town, Gaddap and Gulshan E-Iqbal –before AFP cases occurred in these areas. Two AFP cases occurred in Hyderabad, Sindh Province before ENV was initiated in this district. However when an ENV site was established in 2012, poliovirus was detected every month from July 2012 to August 2013 (Figure 5D1). Two ENV sites in Sukkur district, Sindh Province

detected WPV in sewage samples, with no reported AFP cases in these districts (Figure 5D1).

When further examined by genetic clusters detected by AFP as compared with detection by ENV, similar proportions of each genetic cluster were detected, suggesting that ENV was capturing virus with similar genetics to the ones causing disease in surrounding communities (Figure 6). In Pakistan, the most commonly isolated cluster for both AFP and ENV was R4B, accounting for 44.2% and 34.8% of isolated WPV, respectively (Table 7). In Afghanistan R4B was the second most commonly isolated cluster type, accounting for 24.8% of all virus isolated. Overall, percent of total isolates of each genetic cluster for AFP and ENV were similar, with several exceptions. Only seven (2.1%) of the 326 Pakistan AFP isolates were cluster R3A, while ENV detected 52 isolates, comprising 14.9% of all Pakistan ENV isolates. AFP surveillance in Afghanistan exhibited less genetic diversity overall, with seven different cluster types detected as compared with 15 cluster types in Pakistan. Both AFP and ENV documented the relative decline in genetic diversity and disease burden from 2011 through September 2013 (Figure 6).

Mapping genetic clusters also showed that genetic clusters detected by AFP and those detected by ENV in specific geographic areas were relatively similar. For instance, from January 2011 through June 2011, nearly all R4B detected by ENV was isolated in Karachi, with very few AFP cases due to R4B occurring outside Sindh province (Figure 7A). Genetic clusters detected at ENV sites closer to the Pakistan-Afghanistan border were similar to AFP viruses isolated in these areas, with no detection of these clusters elsewhere. For instance, ENV sites in Quetta district most commonly isolated clusters R2A and Q3, with the majority of AFP isolates of these clusters occurring in border districts (Figure 7A).

The latter half of 2011 experienced both the highest burden of AFP cases and highest number of ENV isolates during the observational period. As in the first six months of 2011, ENV sites detected similar genetic clusters of virus to AFP isolates in the surrounding districts. As seen previously, Karachi ENV sites most commonly isolated cluster R4B. However, AFP cases in this 6 month period (July 2011-Dec 2011) were more widely spread throughout Pakistan than in previous months (Figure 7B). Cluster R3A was the most common genetic cluster detected at Punjab ENV sites, with AFP cases of the same cluster also reported in Punjab province. In addition to detecting R4B, ENV sites in KP commonly isolated R4A cluster poliovirus, with the majority of R4A AFP cases reported from FATA. In the latter half of 2011, cluster R5 was only isolated from AFP cases in FATA, and all R5 ENV isolates were detected in KP (Figure 7B). In 2012, cluster R5 poliovirus was detected by ENV in Lahore, with only one AFP isolate of this cluster detected in KP during this time (Figure 7B). Fewer AFP cases were observed in 2012, but as in previous months, ENV detected similar virus to that detected in geographically close AFP cases. R4B virus was detected in Karachi area by three ENV sites; however, only one AFP case occurred in Sindh province during this period (Figure 7C). All three Lahore ENV sites isolated R2B virus, with no AFP cases of this cluster reported in this district. Decreased genetic diversity and geographic localization was observed from July 2012 to September 2013 (Figure 7D, 7E). Nearly all AFP cases reported during this period were cluster R4B, and observed in FATA and KP provinces. ENV sites in Multan, Punjab province and Karachi, Sindh province detected R4B in sewage samples, with few AFP cases reported in these areas. In 2013, the Hyderabad ENV site in Sindh province and the Faisalabad ENV site in Punjab province detected R3A virus exclusively, with no AFP cases of this cluster detected in 2013 (Figure 7E).

Orphan lineage analysis

When considering only isolates detected by AFP surveillance, 38 of 455 total isolates (8.4%) were orphans. When both AFP and ENV isolates were considered, 8 (21.1%) of all AFP orphans, regardless of location detected, were linked through ENV isolates and no longer considered orphans (Table 8). For the 8 orphans that were accounted for by ENV, the increases in percent match when considering ENV was 0.1% to 1.5% (Table 9A). Of the 38 orphan lineages detected by AFP surveillance alone, 6 (15.7%) had better percent matches by ENV, but remained orphans when ENV isolates were also considered, with increases in percent match ranging from 0.11% to 0.55% (Table 9B). Additionally, when considering both AFP and ENV isolates, 13 ENV orphan lineages were detected that were otherwise undetected by AFP surveillance (Table 10). Overall, when considered together, AFP and ENV surveillance exhibited the smallest proportion of orphans detected (5.4%). When considered independently, ENV detected fewer orphan lineages than did AFP surveillance alone (6.3% vs. 8.4%, respectively) (Table 8).

Sub-lineage analysis

Circulation of closely related genetic (≥99.0%) isolates before and after median date for each defined sub-lineage is depicted in Figure 8, where for a given sub-lineage, the median date is defined as the onset date for the median isolate from the first to the last isolate with a match within 99.5% . Of the 209 sub-lineages, 113 (54.1%) had closely matching (≥99.0%) genetic circulation detected before the median sub-lineage date by both AFP and ENV, while 46 (22.0%) had no detected circulation before by either AFP or ENV (Table 11). There were 24 (11.5%) sub-lineages with circulation detected by ENV only, with no AFP isolates, and 26 (12.4%) of isolates with circulation detected by AFP only, with no ENV isolates. The relative number of sub-lineages with both AFP and ENV circulation

detected before decreased by year, (57.8% in 2011 vs. 32.4% in 2013). Conversely, the number of sub-lineages with only ENV circulation before is the highest in 2013 when the number of AFP cases was fewest (Table 11).

Overall, ENV detected circulation first in more sub-lineages than did AFP (51.7% vs. 26.3%) (Table 12). In 2011 and 2012, ENV detected first circulation in the majority of sub-lineages. However, in 2013 differences in proportions of first circulation were less pronounced, with ENV detecting first circulating virus in 41.2% of sub-lineages and AFP detected first circulation in 29.4% of sub-lineages (Table 12).

For the entire observational period, ENV detected circulation approximately 64 days sooner on average for each sub-lineage than did AFP surveillance (Table 13). ENV detected closely related circulating virus 189.7 days before the median date for AFP isolates in each sub-lineage, while AFP detected circulation 126.2 days before median date for AFP isolates in each sub-lineage (Table 13). ENV detected circulating virus sooner than AFP in 2011 and 2013; ENV detected circulating WPV 71.8 days sooner in 2011 and 51.1 days sooner than AFP surveillance in 2013. In 2012 the difference between AFP and ENV detection was less pronounced, with ENV detected circulation 45.7 days sooner than AFP surveillance.

As 2013 sub-lineages are likely still circulating, current information about overall circulation time and overall circulation time for these isolates is likely biased for these sub-lineages. Therefore, only 2011 and 2012 sub-lineages were considered in calculations of circulation time after and overall circulation time (Table 13). Mean circulation time after sub-lineage median date did not appear to be different as detected by AFP surveillance vs. ENV (Table 13). Overall, ENV detected genetically similar circulation for 266.8 vs. 197.4 days as detected by AFP surveillance. ENV detected circulation for longer than AFP; however,

differences were presumably due to differences between ENV and AFP in circulation time before sub-lineage median date.

DISCUSSION

Detection above AFP surveillance alone

Both AFP and ENV documented the decline in WPV burden in Pakistan from January 2011 through September 2013. However, this study describes instances and geographic areas where poliovirus circulation was detected by ENV that would have otherwise been missed by AFP surveillance alone.

In six of the first seven months of 2011, ENV detected circulating WPV1 in three provinces of Pakistan without reports of AFP cases. In the following six months, these provinces reported the largest number of AFP cases observed during the study period, implying extensive circulation of WPV in these areas only detected by ENV. As fewer AFP cases were reported, the provinces reporting WPV infection by ENV only became more consistent, suggesting that ENV may become increasingly important as eradication efforts progress and fewer polio AFP cases occur. ENV sites in Pakistan were established in areas that local health authorities deemed at high risk for circulation, and although few polio cases were reported, ENV consistently detected WPV circulation in these districts (Table 6).

ENV in Pakistan, although inherently limited in geographic reach by site selection, detected relatively comprehensive WPV virus genetics, with only 6.3% of total isolates detected considered orphan isolates as compared with 8.4% orphan lineages detected by AFP surveillance alone. When considered together, AFP and ENV provided the most comprehensive surveillance; the addition of ENV to AFP surveillance explained 8 AFP orphan lineages. Above this, even for AFP cases that were still considered orphans by the

combination of AFP and ENV surveillance, several of these isolates had better matches by ENV than by AFP alone, suggesting that ENV is filling in gaps present in AFP surveillance.

With sites in 11 districts and 4 provinces by the end of 2013, circulation detected by ENV corresponded to the AFP cases that presented in those communities. Proportions of total isolates of each genetic cluster detected by ENV were similar to those detected by AFP, although the above evidence suggests that ENV surveillance is detecting viruses of certain clusters before AFP in some geographic areas. In several instances throughout the observational period, it appeared that WPV circulation was limited to areas experiencing AFP cases; however, ENV suggested that silent circulation was actually occurring in areas not reporting AFP cases. For example, in the second half of 2012, R4B cluster AFP cases were only reported in Pakistan's Northern provinces, FATA, GB and KP. However, R4B cluster virus was detected by ENV not only in the KP sites, but also at two sites in Punjab, one site in Balochistan and four ENV sites in Karachi, Sindh province, the Southernmost province. In the following months, Karachi districts experienced 3 AFP cases due to R4B virus, corroborating evidence of silent circulation in these districts, not detected by AFP. Although fixed at certain sites, strategic ENV sampling can indicate more widespread circulation than AFP alone. ENV during this period suggested that circulation was far more widespread than AFP data suggested, and if ENV was not being conducted, the incorrect conclusions about localized circulation may have been drawn.

Consistently throughout the study period, ENV detected circulation sooner than did AFP surveillance, as shown by the sub-lineage analyses. For all years in the study period 11.5% of sub-lineages had circulation detected by ENV only with no genetically similar preceding AFP isolates; for these isolates, detection preceding these isolates was limited to ENV only and was completely undocumented by AFP. Even for sub-lineages with preceding

detection by both ENV and AFP, ENV detected circulating wild virus sooner. ENV detected circulation of genetically similar virus sooner than did AFP surveillance in the majority of sub-lineages. On average, ENV uncovered virus circulation more than two months before detection by AFP surveillance, finding similar genetic virus circulation on average almost 190 days before the median date for the sub-lineage. This suggests that ENV surveillance could provide early warning of WPV circulation without having to wait for AFP cases to present, allowing for more time to conduct control measures and immunization activities, preventing further cases.

Overall, this study provides substantial evidence to suggest that ENV in Pakistan is providing information that AFP surveillance alone would not have provided; namely, ENV in Pakistan is detecting genetically similar virus sooner, providing evidence of more geographically widespread circulation than AFP surveillance, and showing evidence of infection in areas not reporting AFP cases.

Even with supplemental ENV surveillance in Pakistan, there was still evidence of remaining gaps in surveillance. Because the vast majority of polio infections are asymptomatic, most infections go unreported by AFP surveillance. These observed gaps are not entirely explained by supplemental ENV due to its inherently limited ability to capture the entire population by sampling site selection. Sub-lineage analysis suggested that 22% of sub-lineages had neither AFP nor ENV detection preceding clinical presentation of these isolates, and were missed by the combination of AFP and ENV. Similarly, when considering both AFP and ENV, 43 orphan lineages were detected (5.4% of all isolates considered), corroborating evidence of persistent gaps in surveillance provided by sub-lineage analyses above.

Strengths and weaknesses of analyses:

This analysis represents one of the first attempts to quantify the added benefit of environmental surveillance over AFP surveillance alone. Pakistan has one of the most extensive ENV systems currently operating, with concurrent endemic circulating WPV virus, providing a unique opportunity to examine the role of ENV while polio AFP cases are still present. This study provides extensive evidence from a diverse set of analyses that enumerate the ways that ENV is supplementing AFP surveillance in Pakistan. These analyses give insight into ways that ENV is functioning in areas where polio remains endemic, and suggesting that ENV may be useful in remaining endemic reservoirs or those areas at high risk for reinfection; namely, ENV detected circulation around Pakistan's borders suggesting that ENV might be able to provide early warning of reintroduction into neighboring districts in Afghanistan and India.

Because the data were pulled from a surveillance system database of positive isolates, limited denominator data on the number of samples tested and their results were available. Although Pakistan's concurrent ENV and continued endemicity of indigenous polio circulation allows for quantification of added benefit of ENV over AFP surveillance alone, the benefits incurred in Pakistan may not be of the same magnitude in areas without endemic circulation of poliovirus. The advantages of ENV rely entirely on the configuration of sampling sites and sampling procedures as well as the local epidemiology of polio. As Pakistan is one of only three areas with remaining wild virus circulation, it is unlikely that the observed results would be the same in non-endemic areas, and the benefits of ENV in these areas need to be explored separately.

Analyses presented in this study, although providing a detailed overview of ENV in Pakistan, are largely qualitative with limited opportunity for statistical testing. Quantification

of circulation for sub-lineage analysis provided an estimate for the amount of time ENV detected circulation preceding detection by AFP surveillance; however, because of the way that sub-lineages were defined, statistical tests for significance of these estimates were not feasible. Some isolates considered closely related to one sub-lineage (≥99.0%) were also considered to be matches for other sub-lineage if they fell within ≥99.0% for both sub-lineages. Because of this, the sub-lineages could not be considered statistically independent, invalidating the basis for the statistical tests. Even so, the above analyses provide useful quantitative information about the detection of closely related ENV isolates before AFP detection that could have practical implications when assessing the role of ENV in end game eradication strategies.

To obtain more statistically sound estimates of ENV's detection preceding AFP detection, discrete sub-lineages could be defined to circumvent issues experienced herein with statistical independence. Further description of ENV in Pakistan could also be useful if denominator data from the laboratories involved about number of samples tested and number of sites reporting at the time were employed to help discern sensitivity and surveillance quality for specific ENV sites. Comprehensive genetic information collected by surveillance systems may also be conducive to further geospatial and network analyses that may help discern transmission networks through bridging of epidemiologic and laboratory data. Future analyses might also include expansion of the evidence provided above to determine if the benefits of ENV described by this study are exhibited with geographic specificity; that is, is ENV detecting WPV circulation sooner than AFP in the same geographic areas in which AFP cases are occurring? These analyses might have more consequential programmatic implications, as detection specific to geographic locations

would allow for implementation of control measure and prevention in the correct localities, limiting spread.

Limitations and considerations for implementation of ENV

ENV provides non-invasive surveillance at regular intervals, with especially useful applications in high-risk areas where lapses in AFP surveillance are expected. Experience in settings including India, Egypt, Israel and other countries, and evidence provided herein suggests that ENV may provide more timely and sensitive detection of WPV, independent of clinical presentation and health infrastructure. Although ENV surveillance has demonstrated benefits, several factors may limit its wider application and use in all settings. Systematic ENV surveillance is most efficient in settings with converging sewer networks, allowing catchment of downstream samples that represent a large number of people. ENV sites in India and Egypt have demonstrated that poliovirus can be isolated from samples not collected at a waste processing facility, through strategic and regular sampling from open canal settings; however, these samples, although abundantly available, do not represent a large number of people, and identifying sampling sites in these settings that capture adequate representation of the population may prove difficult (34, 35, 37).

Converging sewer networks allow catchment of larger numbers of individuals in the underlying population through collection of downstream samples, however, large catchment areas may decrease sample sensitivity, making it difficult to detect poliovirus in populations with relatively few excretors (19). Poliovirus has been successfully isolated by ENV sites with catchment areas of ten thousand to a few million individuals, with Pakistan catchment sites capturing 100,000 individuals (19, 41). In large catchment areas, collection of larger samples can compensate for decreased sensitivity, although larger samples require more

laboratory resources to process. Sensitivity of ENV is also influenced by sample collection schedule and distance from excretors to collection site(19). Probability of detection increases with closer proximity of the source to the sampling site, however if the excretor is very close to the sampling site, hourly pooled aliquots may be necessary to capture peak virus concentrations in sewage (19). Collecting these hourly pooled aliquots may not be feasible without a timed sampling device, which are generally less available in low- and middle-income countries, and run the risk of being stolen if used in an open-channel ENV system. Variability in laboratory detection has been noted for individual Egypt samples tested in parallel at different laboratories, where the larger the number of flasks processed and tested, the larger proportion of samples found to contain WPV, suggesting that improved sensitivity can be achieved through multiple tests of the same sample, although this would require more resources (35).

Implementation of ENV can be resource demanding, even for laboratories with existing poliovirus isolation capacity. An estimated $50,000 in start-up costs are required for a laboratory already isolating poliovirus in clinical specimen, with an additional $33,000 required for supplies to analyze 100 specimens(19). Although overall sample collection is less difficult than for clinical specimens, sample processing is laborious. Ten years of ENV in Egypt has suggested that processing and analyses of 100 ENV specimens requires trained staff at the same levels as for twice the number of AFP cases with two specimens per case (19). ENV sample processing would increase the workload and cost for laboratories, often with already limited trained personnel, time and resources and should not occur at the expense of AFP surveillance.

Another limitation of ENV is that detection is geographically constrained to areas around which the surveillance is conducted, with a large number of sampling sites required

to capture large portions of the population. Although ENV may provide genetic links to previously sequenced polioviruses, ENV isolates are not inherently linked to infected individuals, in contrast to AFP isolates. Thus, elucidation of epidemiological linkages and traceback are required to determine source of infection in a population to implement control measures. This traceback and determination of source can be incredibly complicated for ENV sites with large catchment areas, or for catchment areas representing highly transient populations. For instance, for several ambiguous VDPVs (aVDPVs) detected in Slovakia from 2003-2005, extensive traceback through sewer networks was conducted to determine source. Although researchers were able to trace the excretion to approximately 500 individuals living in flats within a few blocks, detection of the aVDPVs ceased before a definitive source could be identified (46).

Results of ENV, both positive and negative can be difficult to interpret. Negative results do not exclude the possibility of the circulating polioviruses in that community and could be due to several factors including, but not limited to 1) dilution; 2) environmental conditions that affect virus viability; 3) inefficient sampling schedules or 4) inadequate laboratory processing. To validate surveillance, non-polio enteroviruses need to be consistently detected in a geographic area to confirm the absence of poliovirus in samples. Similarly, a single positive isolate does not necessarily indicate circulation within a population, and instead may suggest reintroduction or transient presences in the area (22, 28). To determine circulation, virus should be isolated at multiple time points.

Role of environmental surveillance in endgame strategy

Given its ability for earlier detection and determining extent of virus circulation, ENV could play an important role in the endgame strategy. ENV may be especially useful in settings

where case-driven, passive reporting by AFP surveillance may be insufficient in determining infection, especially for populations with heterogeneous immunization coverage. As previously documented, ENV may help explain gaps in existing surveillance, and help to track importations, reintroductions, and circulation, using genetic information to link these isolates to remaining polio reservoirs. This enables intensification of immunization campaigns preceding paralytic cases, and may help minimize more widespread infection and circulation.

Similarly, ENV may be useful in documenting the extinction of genetic lineages and document the end of circulation, as was done in India beginning in 1999. After the initiation of ENV, numerous separate importations, reintroductions and lineage eliminations were noted by ENV, providing India with invaluable information to implement aggressive immunization activities to protect highly vulnerable control programs (37).

Current certification for polio elimination requires absence of WPV for a minimum of three years in all countries in the region, with concurrent certification-standard surveillance in all countries during that three-year period (13). WPV detection for certification is currently based entirely on AFP surveillance; however, with demonstrated gaps in AFP's ability to detect underlying circulation, incorporating certification requirements based on detection by ENV could be especially useful in areas like Israel where ENV has consistently uncovered WPV positives in the absence of AFP cases(33).

ENV may also have applications in documenting the transition from OPV to IPV. Intestinal immunity induced by IPV is inferior to that induced by OPV. IPV does not reduce the proportion, who shed following an OPV challenge in contrast to OPV vaccinated individuals, although IPV may reduce the titer shed and duration of shedding. Thus, individuals immunized with IPV may still become infected with OPV-derived viruses

excreted by those vaccinated with OPV during previous periods or from viruses excreted by chronically infected, immune-compromised shedders (iVDPVs). While circulation of vaccine-derived viruses is usually short-lived, ENV could be used to detect and monitor these lineages. For instance, IPV is used exclusively in Finland's routine schedule, however importations of OPV strains have been detected by ENV in Finland from neighboring Estonia and Russia, where OPV is still used (19). ENV may be used to monitor the disappearance of Sabin strain viruses as OPV is phased out. ENV may be implemented to detect emergence of cVDPV strains for vulnerable areas with sub-optimal vaccination coverage, and monitor the transition from tOPV to bOPV in Pakistan, planned for 2016.

Conclusions

This study presents evidence that suggests ENV in Pakistan is providing earlier and more sensitive detection than AFP surveillance alone. This does not, however, guarantee that benefits of ENV observed in Pakistan would be completely generalizable to other areas implementing ENV; it is unlikely that the results presented above would be the same in non-endemic areas, and the benefits of ENV in these areas need to be explored separately.

ENV provides anonymous, routine surveillance for poliovirus circulation independent of healthcare infrastructure. However, ENV is not without limitations. The greatest restriction to larger implementation and success of ENV is absence of requisite converging sewer networks in many areas that would otherwise be of interest to monitor. ENV is most effective in areas with converging sewer networks that allows catchment of a larger population, and ENV sensitivity may be influenced by sampling procedures, laboratory practices, and environmental conditions. ENV may be a costly undertaking, even in laboratories already processing clinical polio isolates, and supply and personnel costs as well as risks for poliovirus circulation should be considered before implementation. Results

from ENV, both positive and negative, may be difficult to interpret, but provide important supplemental information to other surveillance systems for response and containment. After polio eradication, ENV systems and laboratories may potentially be adaptable to other pathogens isolated in feces (e.g. other enteroviruses, etc.). Overall, targeted ENV through strategic selection of sites has proven useful in Pakistan, and has useful and important applications in the eradication Endgame strategy, including detecting cVDPVs, monitoring the switch from OPV to IPV, and certification of a polio-free world.

REFERENCES

1. POLIO'S LEGACY: AN ORAL HISTORY. *ORAL HISTORY REVIEW* 1996.
2. Underwood M. A Treatise on the Diseases of Children with Directions for the Management of Infants from the Birth. 1799.
3. Atkinson W WS, Hamborsky J, eds. *Centers for Disease Control and Prevention. Epidemiology and Prevention of Vaccine-Preventable Diseases*. 12th ed., second printing ed.: Washington DC: Public Health Foundation, 2012.; 2012.
4. Nathanson N, Martin JR. The epidemiology of poliomyelitis: enigmas surrounding its appearance, epidemicity, and disappearance. *Am J Epidemiol* 1979;110(6):672-92.
5. Nathanson N, Kew OM. From emergence to eradication: the epidemiology of poliomyelitis deconstructed. *Am J Epidemiol* 2010;172(11):1213-29.
6. Pearce JM. Poliomyelitis (Heine-Medin disease). *Journal of neurology, neurosurgery, and psychiatry* 2005;76(1):128.
7. Plotkin S, Orenstein W, Offit P. *Vaccines*. 2012.
8. Sutter RW, Kew OM, Cochi SL, et al. Poliovirus vaccine—live. *Vaccines*, 2013:598-645.
9. Savolainen-Kopra C, Blomqvist S. Mechanisms of genetic variation in polioviruses. *Reviews in medical virology* 2010;20(6):358-71.
10. Sartwell PE. The Incubation Period of Poliomyelitis. *American Journal of Public Health and the Nations Health* 1952;42(11):1403-8.
11. Estivariz CF, Pallansch MA, Anand A, et al. Poliovirus vaccination options for achieving eradication and securing the endgame. *Current opinion in virology* 2013;3(3):309-15.
12. Grassly NC, Jafari H, Bahl S, et al. Waning intestinal immunity after vaccination with oral poliovirus vaccines in India. *The Journal of infectious diseases* 2012;205(10):1554-61.
13. Initiative GPE. Polio Eradication & Endgame Strategic Plan 2013-2018. 2013.
14. Heinsbroek E, Ruitenberg EJ. The global introduction of inactivated polio vaccine can circumvent the oral polio vaccine paradox. *Vaccine* 2010;28(22):3778-83.
15. Duintjer Tebbens RJ, Pallansch MA, Kim JH, et al. Oral poliovirus vaccine evolution and insights relevant to modeling the risks of circulating vaccine-derived polioviruses (cVDPVs). *Risk analysis : an official publication of the Society for Risk Analysis* 2013;33(4):680-702.
16. Global eradication of poliomyelitis by the year 2000. Forty First World Health Assembly, 1988.
17. Dutta A. Epidemiology of poliomyelitis—Options and update. *Vaccine* 2008;26(45):5767-73.
18. Chatterjee A, Vidyant S, Dhole TN. Polio eradication in India: progress, but environmental surveillance and vigilance still needed. *Vaccine* 2013;31(9):1268-75.
19. Hovi T, Shulman LM, van der Avoort H, et al. Role of environmental poliovirus surveillance in global polio eradication and beyond. *Epidemiology and infection* 2012;140(1):1-13.
20. Centers for Disease C, Prevention. Progress toward eradication of polio--worldwide, January 2011-March 2013. *MMWR Morbidity and mortality weekly report* 2013;62(17):335-8.
21. Poliomyelitis: intensification of the global eradication initiative. World Health Organization, 2012.

22. Hovi T. Surveillance for polioviruses. *Biologicals : journal of the International Association of Biological Standardization* 2006;34(2):123-6.
23. Birmingham ME, Linkins RW, Hull BP, et al. Poliomyelitis surveillance: the compass for eradication. *The Journal of infectious diseases* 1997;175 Suppl 1:S146-50.
24. Howard E. Gary J, Ray Sanders, and Mark A. Pallansch. A Theoretical Framework for Evaluating the Sensitivity of Surveillance for Detecting Wild Poliovirus: I. Factors Affecting Detection Sensitivity in a Person with Acute Flaccid Paralysis. *Journal of Infectious Diseases* 1997.
25. Watkins RE, Martin PA, Kelly H, et al. An evaluation of the sensitivity of acute flaccid paralysis surveillance for poliovirus infection in Australia. *BMC infectious diseases* 2009;9:162.
26. Dowdle W, Van Der Avoort H, De Gourville E, et al. Containment of Polioviruses After Eradication and OPV Cessation: Characterizing Risks to Improve Management. *Risk Analysis* 2006;26(6):1449-69.
27. Ranta J, Hovi T, Arjas E. Poliovirus surveillance by examining sewage water specimens: Studies on detection probability using simulation models. *Risk Analysis* 2001;21(6):1087-96.
28. Lowther S, Asghar H, Weldegebriel G, et al. Environmental Surveillance for Polioviruses in the Global Polio Eradication Initiative. *The Journal of infectious diseases* 2014 (In Press).
29. Centers for Disease C, Prevention. Evaluating surveillance indicators supporting the Global Polio Eradication Initiative, 2011-2012. *MMWR Morbidity and mortality weekly report* 2013;62(14):270-4.
30. Pöyry T, Stenvik M, Hovi T. Viruses in sewage waters during and after a poliomyelitis outbreak and subsequent nationwide oral polio vaccination campaign in Finland. *Applied and environmental microbiology* 1988.
31. Hovi T, Cantell K, Huovilainen A, et al. Outbreak of paralytic poliomyelitis in Finland: widespread circulation of antigenically altered poliovirus type 3 in a vaccinated population. *Lancet* 1986;1(8495):1427-32.
32. Manor Y, Handsher R, Halmut T, et al. Detection of poliovirus circulation by environmental surveillance in the absence of clinical cases in Israel and the Palestinian authority. *Journal of clinical microbiology* 1999;37(6):1670-5.
33. Shulman LM, Gavrilin E, Jorba J, et al. Molecular epidemiology of silent introduction and sustained transmission of wild poliovirus type 1, Israel, 2013. *Euro Surveill* 2014;19(7):20709.
34. El Bassioni L. Prolonged Detection of Indigenous Wild Polioviruses in Sewage from Communities in Egypt. *American Journal of Epidemiology* 2003;158(8):807-15.
35. Hovi T, Blomqvist S, Nasr E, et al. Environmental surveillance of wild poliovirus circulation in Egypt--balancing between detection sensitivity and workload. *Journal of virological methods* 2005;126(1-2):127-34.
36. Blomqvist S, El Bassioni L, El Maamoon Nasr EM, et al. Detection of imported wild polioviruses and of vaccine-derived polioviruses by environmental surveillance in Egypt. *Applied and environmental microbiology* 2012;78(15):5406-9.
37. Deshpande JM, Shetty SJ, Siddiqui ZA. Environmental Surveillance System To Track Wild Poliovirus Transmission. *Applied and environmental microbiology* 2003;69(5):2919-27.
38. Walsh D. Taliban Block Vaccinations in Pakistan. The New York Times, 2012.

39. Centers for Disease C, Prevention. Progress toward poliomyelitis eradication - Pakistan, January 2012-September 2013. *MMWR Morbidity and mortality weekly report* 2013;62(46):934-8.
40. Centers for Disease C, Prevention. Progress toward poliomyelitis eradication - Afghanistan and Pakistan, January 2011-August 2012. *MMWR Morbidity and mortality weekly report* 2012;61(39):790-5.
41. Sharif S. Environmental Surveillance Site Selection in Pakistan. 2014.
42. Guidelines for environmental surveillance of poliovirus circulation. World Health Organization, 2003.
43. Polio Laboratory Manual. Geneva, Switzerland: World Health Organization, 2004.
44. Thompson KM, Pallansch MA, Duintjer Tebbens RJ, et al. Preeradication vaccine policy options for poliovirus infection and disease control. *Risk analysis : an official publication of the Society for Risk Analysis* 2013;33(4):516-43.
45. Mathieu Bastian SH, Mathieu Jacomy. Gephi: An Open Source Software for Exploring and Manipulating Networks. Presented at Proceedings of the Third International ICWSM Conference2009.
46. Cernakova B, Sobotova Z, Rovny I, et al. Isolation of vaccine-derived polioviruses in the Slovak Republic. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 2005;24(6):438-9.

TABLES

Table 1. Global Polio Eradication Initiative minimum level indicators for the certification of Acute Flaccid Paralysis Surveillance

Indicator	Description	Target
Completeness of reporting	Expected routine reports (weekly or monthly) received on time, including zero reports in the absence of AFP cases. Distribution of reporting sites should be representative of the geography and demographics of the country.	≥80%
Sensitivity of surveillance	Adequate number of non-polio AFP detected in population aged less than 15 years.	≥1:100,000 annually[ǂ]
Completeness of case investigation	AFP cases should have a full clinical and virologic investigation with 'adequate' stool specimen collected, defined as: 1) Two stool specimens of sufficient quality for laboratory analysis 2) Collected at least 24 hours apart 3) Within 14 days of paralysis onset 4) Arriving at an accredited the laboratory by reverse cold chain and with proper documentation	≥80%
Completeness of follow up	AFP cases should have a follow-up examination for residual paralysis at 60 days after onset of paralysis	≥80%
Laboratory performance	AFP specimens must be processed in a WHO-accredited laboratory within the Global Polio Laboratory Network (GPLN)	100%

Table adapted from GPEI (http://www.polioeradication.org/Dataandmonitoring/Surveillance.aspx)
[ǂ] In endemic regions, for improved sensitivity, rate of >2:100,000 is suggested.

Table 2. Complete list of variables and their descriptions included in the final dataset for analysis

Variable	Description
ID Number	Unique identifier for both AFP and ENV isolates. Completely de-identified
Isolate type	AFP if sample came from polio case or close contact; ENV if sample was isolated in the environment. Since all environmental sampling in our study conducted in Pakistan, there are only AFP samples from Afghanistan
Country	Country where isolate was obtained – either Pakistan or Afghanistan
State/Province	State or province in Pakistan or Afghanistan where the isolate was collected. For AFP isolates, this represents where the case presented. For ENV isolates, this represents where the environmental site is located
Locality	A finer measure of geographic location for where the isolate was obtained, usually district. This was the finest level of geographic detail available for each isolate.
Date	For AFP polio cases, the date variable represents date of onset of paralysis. When not available, for close contacts of paralysis cases, or for ENV samples, this date represents the date the specimen was collected.
Serotype	Samples included in our analysis represent wild poliovirus type 1 (WPV1), and type 3 WPV3. Only 5 samples from were WPV3, all from AFP isolates in Pakistan
Cluster	Genetic clusters are determined by PPLB and classified once a year. All isolates within a cluster are <5% different from each other with respect to nucleotides in their VP1 capsid gene
Percent match	For each Pakistan AFP isolate, their percent pairwise match in VP1 capsid was obtained for all other isolates in the dataset. A measure of genetic similarity between isolates.

Table 3. Number of wild poliovirus (WPV) positive isolates by country, year, serotype and surveillance system

	AFGHANISTAN	PAKISTAN				
	AFP	PAK AFP		PAK ENV	PAK TOTAL	PAK/AFG TOTAL
Year	*WPV1*	*WPV1*	*WPV3*	*WPV1*	*WPV1/WPV3*	*WPV1/WPV3*
2011	85	210	2	196	408	493
2012	36	60	3	110	173	209
2013	8	51	0	42	93	101
Total	129	321	5	348	674	803

Table 4. Number of WPV positive isolates by country, surveillance type (AFP or ENV), year and month

	Number of Positive WPV Isolates			
2011	**AFG**	**PAK**		**AFG/PAK**
	AFP	**AFP**	**ENV**	**TOTAL**
Jan	1	9	18	28
Feb	0	8	6	14
Mar	0	12	14	26
Apr	1	12	10	23
May	6	13	11	30
Jun	4	9	6	19
Jul	1	12	9	22
Aug	17	20	20	57
Sep	17	36	26	79
Oct	14	40	22	76
Nov	13	19	26	58
Dec	11	22	28	61
2011 Total	85	212	196	493
2012	**AFG**	**PAK**		**AFG/PAK**
	AFP	**AFP**	**ENV**	**TOTAL**
Jan	4	13	15	32
Feb	1	3	13	17
Mar	1	1	5	7
Apr	0	1	12	13
May	3	7	9	19
Jun	3	2	6	11
Jul	4	5	11	20
Aug	0	7	6	13
Sep	9	11	10	30
Oct	5	9	8	22
Nov	4	4	8	16
Dec	2	0	7	9
2012 Total	36	63	110	209
2013	**AFG**	**PAK**		**AFG/PAK**
	AFP	**AFP**	**ENV**	**TOTAL**
Jan	1	2	5	8
Feb	0	3	4	7
Mar	1	1	4	6
Apr	0	2	4	6
May	0	9	2	11
Jun	1	4	2	7
Jul	1	4	7	12
Aug	1	8	7	16
Sep	3	18	7	28
2013 Total (Through Sept.)	8	51	42	101

Table 5. Number of infected states in Afghanistan and provinces Pakistan by surveillance type (AFP, ENV or both) by month and year – January 2011 through September 2013

	Number of Infected Provinces					
	AFG – Positive states by:		PAK – Positive provinces by:			
	AFP		AFP	Both AFP and ENV	AFP Only, No ENV	ENV Only, No AFP
	n	States*	n	Provinces**	Provinces**	Provinces**
2011						
Jan	1	KAN	4	KP, PUN, SIN	FATA	BAL
Feb	0	-	4	BAL, KP, SIN	FATA	-
Mar	0	-	4	BAL, KP, SIN	FATA	PUN
Apr	1	FAR	4	BAL, KP, SIN	FATA	PUN
May	2	HIL, KAN	5	BAL, KP, SIN	GB, FATA	PUN
Jun	3	FAR, HIL, KAN	4	KP, SIN	FATA	PUN
Jul	1	KAN	4	BAL, KP, PUN	FATA	SIN
Aug	4	FAR, HIL, KAN, PAR	5	BAL, KP, PUN, SIN	FATA	-
Sep	8	BAD,HIL, KAN, KUN, NAN, PAR, URU, ZAB	5	BAL, KP, PUN, SIN	FATA	-
Oct	9	BAG, FRY, HIL, KAN, KUN, PAK, PAR, URU, ZAB	5	BAL, KP, PUN, SIN	FATA	-
Nov	5	HIL, KAN, KAP, PAR, URU	5	BAL, KP, PUN, SIN	FATA	-
Dec	5	BAG, FAR, FRY, HIL, KAN	5	BAL, KP, PUN, SIN	FATA	-
2011 Total	14	BAD, BAG, FAR, FRY, HIL, KAN, KAP, KUN, KND, NAN, PAK, PAR, URU, ZAB	6	BAL, KP, PUN, SIN	GB, FATA	-
2012						
Jan	2	HIL, KAN	5	BAL, KP, PUN, SIN	FATA	-
Feb	1	PAK	2	KP	FATA	BAL, PUN, SIN
Mar	1	KAN	1	-	FATA	KP, PUN
Apr	0	-	1	-	FATA	KP, PUN, SIN
May	2	HIL, KAN	4	KP, PUN, SIN	FATA	-
Jun	3	FAR, HIL, KUN	2	-	BAL, KP	PUN, SIN
Jul	2	HIL, KUN	2	KP	FATA	PUN, SIN
Aug	0	-	2	KP	FATA	PUN, SIN
Sep	5	DAY, GHO, KAN, KHO, KUN	4	KP, SIN	GB, FATA	PUN
Oct	4	DAY, HIL, KAN, PAK	3	KP	BAL, FATA	PUN, SIN
Nov	3	HIL, KAN, KHO	2	KP	FATA	BAL, PUN, SIN
Dec	1	NAN	0	-	-	BAL, KP, SIN
2012 Total	9	DAY, FAR, GHO, HIL, KAN, KHO, KUN, NAN, PAK	6	BAL, KP, PUN, SIN	GB, FATA	-
2013						
Jan	1	NAN	2	KP, SIN	-	KP
Feb	0	-	2	KP	PUN	BAL, SIN
Mar		KUN	1	SIN	-	KP, PUN
Apr	0	-	2	KP	FATA	SIN
May	0	-	2	-	FATA, PUN	KP, SIN
Jun	1	NAN	2	KP	FATA	SIN
Jul	1	KUN	2	SIN	FATA	KP, PUN
Aug	1	KUN	3	KP, SIN	FATA	PUN
Sep	1	KUN	4	KP, PUN, SIN	FATA	-
2013 Total	2	KUN, NAN	4	KP, PUN, SIN	FATA	BAL

***Afghanistan state abbreviations:** BAD: Badghis; BAG: Baghlhan; FAR: Farah; FRY: Faryab; GHO: Ghor; HIL: Hilmand; KAN: Kandahar; KAP: Kapisa; KUN: Kunar; KND: Kunduz; NAN: Nangarhar; PAK: Paktita; PAR: Parwan; URU: Uruzgan; ZAB: Zabul

****Pakistan province abbreviations:** BAL: Balochistan; FATA: Federally administered tribal areas; GB: Gilgit-Baltistan; KP: Khyber Pakhtunkhwa; PUN: Punjab; SIN: Sindh

Table 6. Number of infected districts in Afghanistan and Pakistan by surveillance type (AFP, ENV or both) by month and year – January 2011 through September 2013

	Number of Infected districts				
	AFG	PAK			
	AFP ONLY	AFP ONLY	ENV ONLY	BOTH AFP & ENV[1]	TOTAL
2011					
Jan	1	6	5	1	12
Feb	0	6	3	0	9
Mar	0	8	7	1	16
Apr	1	10	6	0	16
May	4	10	5	1	16
Jun	4	8	4	0	12
Jul	1	8	6	1	15
Aug	7	10	5	2	17
Sep	12	16	6	2	24
Oct	13	19	7	1	27
Nov	12	13	6	1	20
Dec	9	11	5	2	18
2011 Total	36	56	5	3	64
2012					
Jan	4	7	4	2	13
Feb	1	2	6	0	8
Mar	1	1	2	0	3
Apr	0	1	5	0	6
May	2	5	5	0	10
Jun	3	2	4	0	6
Jul	4	5	7	0	12
Aug	0	6	4	1	11
Sep	5	9	7	0	16
Oct	4	7	5	1	13
Nov	4	3	4	1	8
Dec	2	0	4	0	4
2012 Total	24	26	8	3	37
2013					
Jan	1	2	4	0	6
Feb	0	3	3	0	6
Mar	1	1	3	0	4
Apr	0	2	3	0	5
May	0	4	2	0	6
Jun	1	2	1	1	4
Jul	1	4	5	0	9
Aug	1	5	4	1	10
Sep	3	5	3	1	9
2013 Total	6	15	7	1	23

[1] Inherently limited by the number of districts conducting ENV. Maximum varies by year (2011: 8 districts; 2012 & 2013: 11 districts)

Table 7. Number of infected districts in Afghanistan and Pakistan by surveillance type (AFP, ENV or both) by month and year – January 2011 through September 2013

Number of Positive Isolates by Genetic Cluster								
	AFG		PAK					
	AFP n=129		AFP n=326		ENV n=348		Total: AFP or ENV n=674	
Genetic Cluster	n	%	n	%	n	%	n	%
H4	-	-	5	1.5%	-	-	5	0.7%
Q1	-	-	1	0.3%	-	-	1	0.1%
Q2	-	-	2	0.6%	6	1.7%	8	1.2%
Q3	32	24.8%	6	1.8%	8	2.3%	14	2.1%
Q4	2	1.6%	1	0.3%	-	-	1	0.1%
R1	-	-	8	2.5%	4	1.1%	12	1.8%
R2A	37	28.7%	19	5.8%	12	3.4%	31	4.6%
R2B	20	15.5%	50	15.3%	42	12.1%	92	13.6%
R3A	-	-	7	2.1%	52	14.9%	59	8.8%
R3B	-	-	10	3.1%	29	8.3%	39	5.8%
R4A	5	3.9%	49	15.0%	60	17.2%	109	16.2%
R4B	32	24.8%	144	44.2%	121	34.8%	265	39.3%
R5	1	0.8%	15	4.6%	12	3.4%	27	4.0%
R6	-	-	1	0.3%	-	-	1	0.1%
R7	-	-	8	2.5%	2	0.6%	10	1.5%

Table 8. Number of orphan lineages detected by AFP surveillance only, ENV surveillance only, and ENV and AFP surveillance in conjunction from Jan 2011 – Sept 2013

Surveillance considering[1]:	AFP Orphans Detected n=38	ENV Orphans Detected n=22	Total Orphans[3] n=60	Total Isolates	Percent Orphans[4]
AFP isolates only	38 (100%)	-	38 (63.3%)	455	8.4%
ENV isolates only	-	22 (100%)	22 (36.7%)	348	6.3%
Both AFP & ENV isolates[2]	30 (78.9%)	13 (59%)	43 (71.6%)	803	5.4%

[1] Each line in the table represents number of orphan lineages detected considering the listed type of isolates.
[2] Includes viruses that remain orphans because there is no AFP isolate or ENV isolate within 1.5% of the nucleotide sequence of the VP1 capsid of that virus. 17 viruses are no longer orphans overall when considering both ENV and AFP.
[3] Orphans detected as a percentage of total number of orphans detected by any surveillance system
[4] Orphans detected as a percentage of total isolates for each respective surveillance type

Table 9. AFP isolate orphan lineages with improved percent matches when considering both AFP & ENV surveillance Jan 2011 – Sept 2013

A.) AFP Orphans, accounted for by ENV[1]						
#	**AFP Orphan (Country, Year, Type)**	**Best match by AFP only (Country, Year, Type)**	**AFP only best match % identity**	**Best match by AFP & ENV (Country, Year, Type)**	**AFP & ENV Best Match % Identity**	**% Change by ENV**
1	**PAK 2011 AFP**	PAK 2010 AFP	98.1236	PAK 2011 ENV	99.117	0.9934
2	**PAK 2011 AFP**	AFG 2010 AFP	98.4547	PAK 2010 ENV	98.5651	0.1104
3	**PAK 2011 AFP**	PAK 2010 AFP	98.3444	PAK 2011 ENV	98.6755	0.3311
4	**PAK 2011 AFP**	PAK 2011 AFP	98.4547	PAK 2011 ENV	98.7859	0.3312
5	**PAK 2011 AFP**	PAK 2010 AFP	98.1236	PAK 2011 ENV	99.6689	1.5453
6	**PAK 2013 AFP**	PAK 2011 AFP	98.4547	PAK 2012 ENV	99.3377	0.883
7	**PAK 2013 AFP**	PAK 2013 AFP	97.7925	PAK 2013 ENV	99.6689	1.8764
8	**PAK 2013 AFP**	PAK 2013 AFP	98.0132	PAK 2013 ENV	98.7859	0.7727
B.) AFP Orphans, better matches by ENV than AFP[2]						
#	**AFP Orphan (Country, Year, Type)**	**Best match by AFP only (Country, Year, Type)**	**AFP only best match % identity**	**AFP & ENV Best Match**	**AFP & ENV Best Match % Identity**	**% Change by ENV**
1	**AFG 2011 AFP**	PAK 2011 AFP	97.9029	PAK 2010 ENV	98.3444	0.4415
2	**PAK 2012 AFP**	PAK 2011 AFP	98.1236	PAK 2011 ENV	98.4547	0.3311
3	**PAK 2012 AFP**	PAK 2012 AFP	98.0132	PAK 2011 ENV	98.3444	0.3312
4	**PAK 2012 AFP**	PAK 2011 AFP	97.7925	PAK 2010 ENV	98.3444	0.5519
5	**PAK 2013 AFP**	PAK 2013 AFP	98.234	PAK 2012 ENV	98.3444	0.1104
6	**PAK 2013 AFP**	AFG 2012 AFP	97.7925	PAK 2012 ENV	98.0132	0.2207

[1] With the addition of ENV isolates to AFP, these isolates are no longer orphans (>98.5% match)

[2] With the addition of ENV isolates to AFP surveillance, these isolates remain orphans (<98.5% match), but have better ENV matches than AFP matches

Table 10. ENV isolate orphan lineages detected by AFP& ENV surveillance Jan 2011 – Sept 2013

#	Orphan details			Best match details			Best match identity
	Country	Year	Type	Country	Year	Type	%
1	PAK	2011	ENV	PAK	2010	ENV	97.4614
2	PAK	2011	ENV	PAK	2011	ENV	98.1236
3	PAK	2011	ENV	PAK	2011	ENV	98.3444
4	PAK	2012	ENV	PAK	2011	ENV	95.3642
5	PAK	2012	ENV	PAK	2011	ENV	98.234
6	PAK	2012	ENV	PAK	2012	ENV	98.234
7	PAK	2012	ENV	AFG	2011	AFP	98.0132
8	PAK	2013	ENV	PAK	2012	AFP	98.234
9	PAK	2013	ENV	PAK	2013	AFP	98.3444
10	PAK	2013	ENV	PAK	2013	AFP	98.3444
11	PAK	2013	ENV	PAK	2013	AFP	98.4547
12	PAK	2013	ENV	PAK	2012	ENV	98.0132
13	PAK	2013	ENV	AFG	2012	AFP	97.9029

Table 11. Proportion of Pakistan AFP isolates by year and type of circulation preceding case – Jan 2011 – Sept 2013

Surveillance system detecting sub-lineage circulation[1]	2011 n=135	2012 n=40	2013 n=34	Overall N=209
No circulation before	25 (18.5%)	11 (27.5%)	10 (29.4%)	46 (22.0%)
ENV only before[2]	14 (10.4%)	2 (5.0%)	8 (23.5%)	24 (11.5%)
AFP only before[3]	18 (13.3%)	3 (7.5%)	5 (14.7%)	26 (12.4%)
AFP and ENV before	78 (57.8%)	24 (60.0%)	11 (32.4%)	113 (54.1%)

[1] Categories are mutually exclusive
[2] Circulation detected by ENV only, no AFP isolates circulating before sub-lineage median date
[3] Circulation detected by AFP only, no ENV isolates circulating before sub-lineage median date

Table 12. Proportion of Pakistan AFP sub-lineages by year and first type of circulating virus– Jan 2011 – Sept 2013

	TYPE OF FIRST CIRCULATING VIRUS SUB-LINEAGE WITHIN 99%			
YEAR	**ENV**	**AFP**	**NONE BEFORE**	**TOTAL**
2011	71 (52.6%)	39 (28.9%)	25 (18.5%)	135
2012	23 (57.5%)	6 (15.0%)	11 (27.5%)	40
2013	14 (41.2%)	10 (29.4%)	10 (29.4%)	34
Total	108 (51.7%)	55 (26.3%)	46 (22.0%)	209

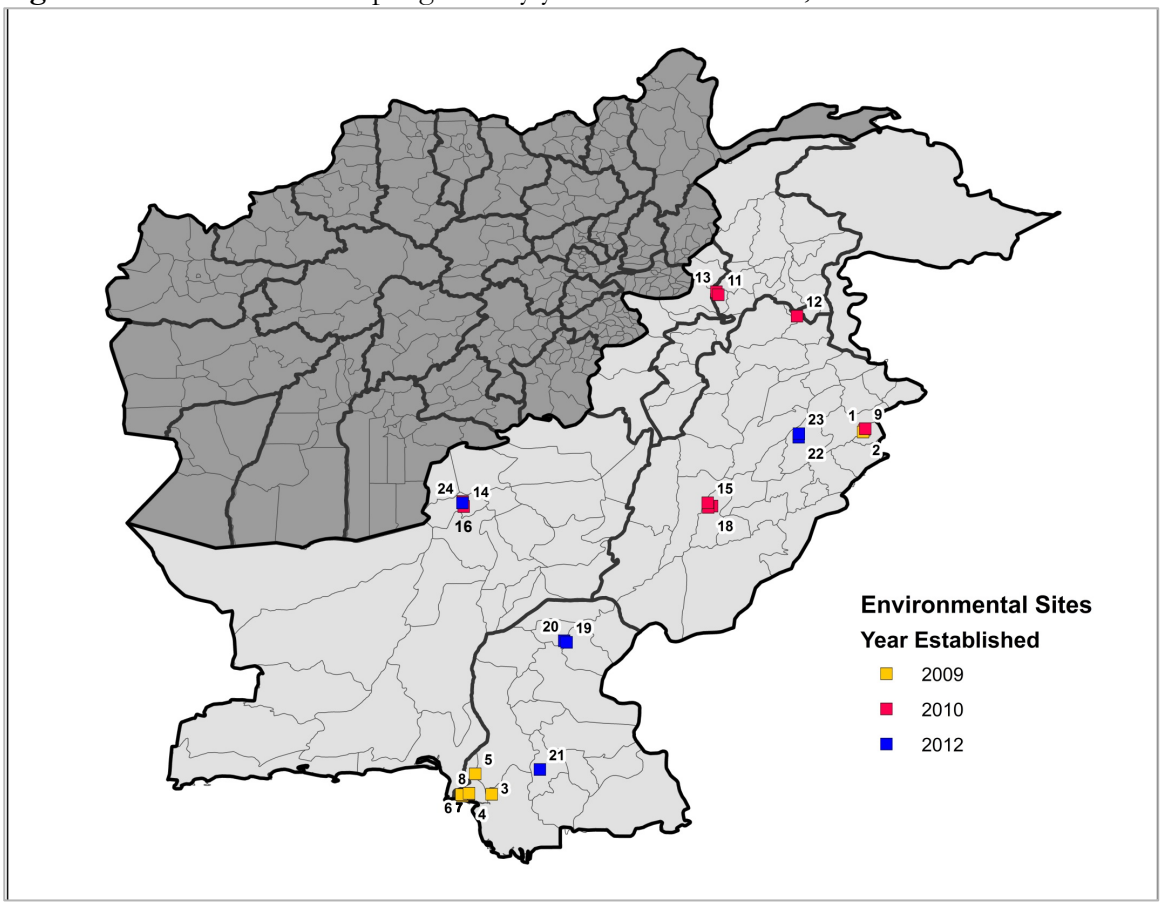
Table 13. Overall circulation time and circulation by isolate type (days) for PAK AFP sub-lineages – Jan 2011 to Sept 2013

A. Mean Circulation time before sub-lineage median date (Days)				
	n	**ENV** Mean (IQR)	**AFP** Mean (IQR)	**Difference (ENV vs. AFP)** Mean (IQR)
2011	135	192.82 (0, 328)	121.01 (0, 197)	71.81 (-34, 235)
2012	40	221.85 (0, 355)	176.18 (0, 310)	45.68 (0, 93)
2013	34	139.24 (0, 241)	88.18 (0, 188)	51.06 (-8, 93)
Total	209	189.66 (0,327)	126.22 (0, 209)	63.44 (-10, 192)
B. Mean circulation time after sub-lineage median date (Days)				
	n	**ENV** Mean (IQR)	**AFP** Mean (IQR)	**Difference (ENV vs. AFP)** Mean (IQR)
2011	135	75.13 (0, 122)	70.71 (0, 116)	4.42 (-10, 17)
2012	40	41.00 (0, 30.5)	40.33 (0, 30)	0.68 (0, 0)
2013	34	8.94 (0, 0)	20.00 (0, 17)	-11.06 (-14, 0)
Total (2011 & 2012)[1]	175	67.33 (0,110)	63.77 (0, 103)	3.57 (0, 11)
C. Total Circulation time for sub-lineage (days)				
	n	**ENV** Mean (IQR)	**AFP** Mean (IQR)	**Difference (ENV vs. AFP)** Mean (IQR)
2011	135	267.96 (0, 486)	191.72 (19, 323)	76.24 (-34, 239)
2012	40	262.85 (0, 393)	216.50 (0, 335)	46.35 (0, 107.5)
2013	34	148.18 (0, 272)	108.18 (0, 207)	40.00 (-34, 98)
Total (2011 & 2012)[1]	175	266.79 (0, 443)	197.38 (13, 323)	69.41 (-17, 189)

[1] Samples from only considered through Sept 2013. Therefore, 2013 data for circulation are biased due to restricted dates, with 2013 year end samples still in process. Only 2011-2012 samples considered.

FIGURES AND FIGURE LEGENDS

Figure 1. Environmental sampling sites by year of establishment, Pakistan



#	ENV SITE NAME	PROVINCE	DISTRICT	START YEAR	END YEAR
1	GULSHAN RAVI STATION	PUNJAB	LAHORE	2009	-
2	MULTAN ROAD STATION	PUNJAB	LAHORE	2009	-
3	CHAKORA NULLA	SINDH	KHI.GULSHAN-E-IQBAL TOWN	2009	-
4	RASHID MINHAS RD LAY	SINDH	KHI.GULSHAN-E-IQBAL TOWN	2009	-
5	COMPOSITE SAMPLE	SINDH	KHI.GADAP TOWN	2009	-
6	SAJJAN GOTH	SINDH	KHI.BALDIA TOWN	2009	-
7	KUMHAR WARRA HUB RIV	SINDH	KHI.BALDIA TOWN	2009	2010
8	SOHRAB GOTH	SINDH	KHI.GADAP TOWN	2009	-
9	OUTFALL STATION	PUNJAB	LAHORE	2010	-
10	BALDIA COMPOSITE	SINDH	KHI.GADAP TOWN	2010	-
11	SHAHEEN TOWN	KHYBER PAKHTUNKHWA	PESHAWAR	2010	-
12	SAFDAR ABAD	PUNJAB	RAWALPINDI	2010	-
13	LARA MA	KHYBER PAKHTUNKHWA	PESHAWAR	2010	-
14	JATAK KILLI & TAKHTHANI BY PASS	BALOCHISTAN	QUETTA	2010	-
15	SURAJ MIANI	PUNJAB	MULTAN	2010	-
16	JAM-E-SALFIA	BALOCHISTAN	QUETTA	2010	-

#	ENV SITE NAME	PROVINCE	DISTRICT	START YEAR	END YEAR
17	ALI TOWN	PUNJAB	MULTAN	2010	-
18	KOTLA ABDUL FATAH	PUNJAB	MULTAN	2010	-
19	MIANI PUMPING STATION	SINDH	SUKKUR	2012	-
20	MAKKA PUMPING STATION	SINDH	SUKKUR	2012	-
21	TULSIDAS PUMPING STATION	SINDH	HYDERABAD	2012	-
22	FAISALABAD COMPOSITE SITE	PUNJAB	FAISALABAD	2012	-
23	ISMAIL PUMPING STATION	PUNJAB	FAISALABAD	2012	-
24	SUR PUL	BALOCHISTAN	QUETTA	2012	-

Figure 2. Number of positive isolates detected by AFP and ENV surveillance in Pakistan and Afghanistan – January 2011-September 2013

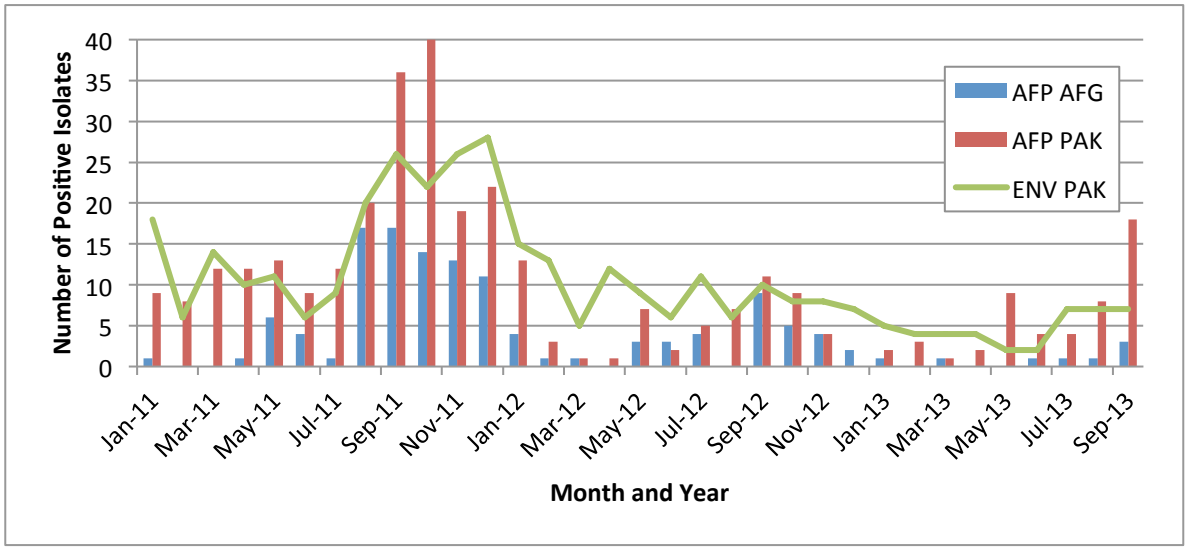


Figure 3. Number of districts with positive WPV isolates detected by AFP and ENV surveillance in Pakistan and Afghanistan – January 2011-September 2013

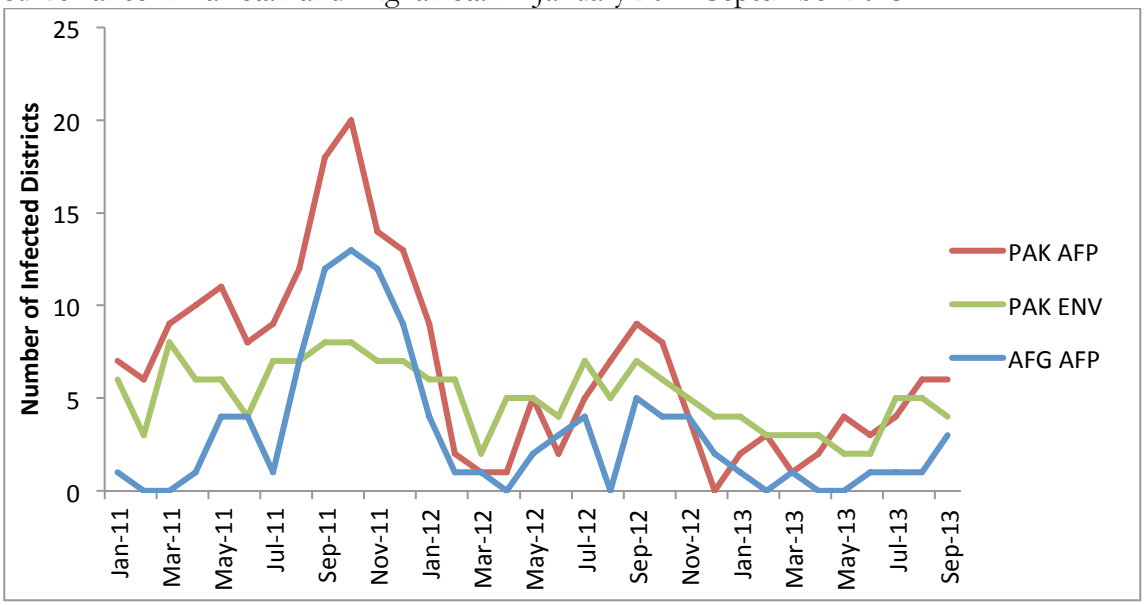
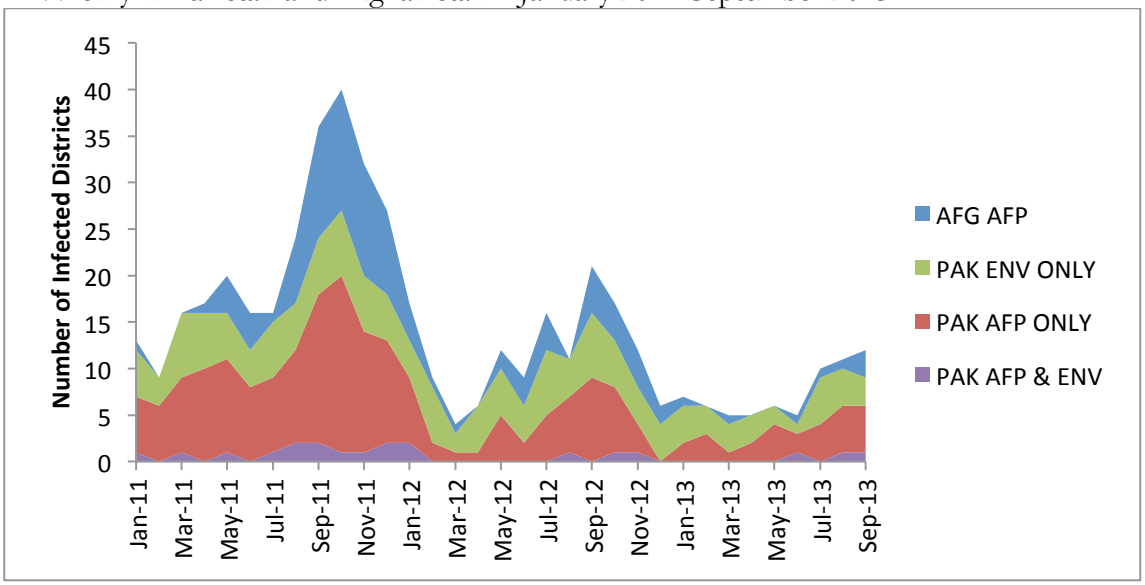
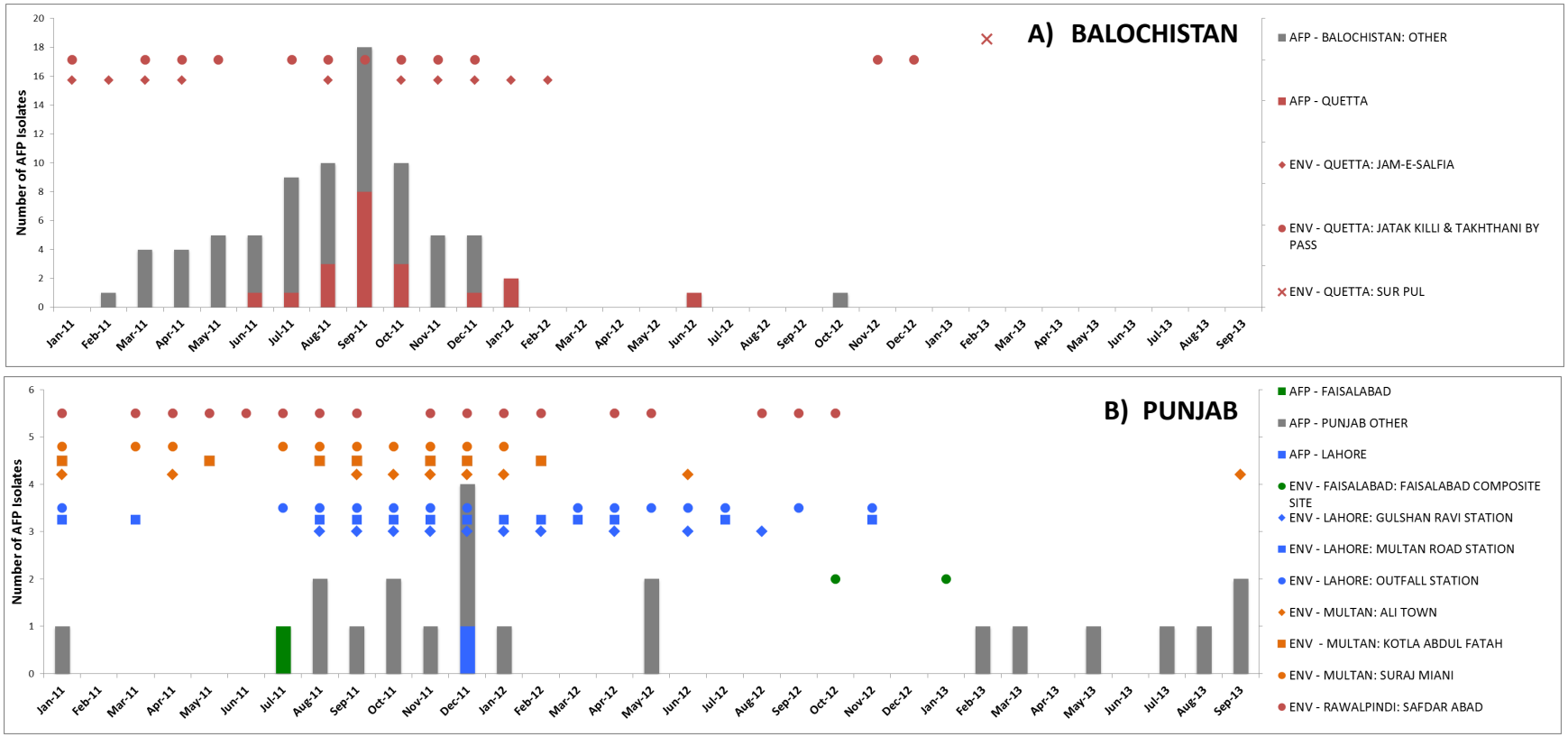


Figure 4. Number of districts with positive WPV isolates detected by both AFP and ENV, districts with WPV isolates detected by AFP only, and districts with WPV isolates detected ENV only in Pakistan and Afghanistan – January 2011-September 2013.



Note: Number of districts in Pakistan with both AFP and ENV detected inherently limited by the number of districts conducting ENV. Maximum varies by year (2011: 8 districts; 2012 & 2013: 11 districts)

Figure 5. Number of polio AFP cases and detection by ENV by Province – January 2011-September 2013.
Panels represent individual provinces in Pakistan, with month number of polio AFP cases on the dependent axis. Dots represent months where WPV was detected by ENV Detection at a site in the province. Corresponding colors represent AFP or ENV isolates detected in the same district. Legend items labeled 'AFP' are depicted as bars on the depended axis, while legend items labeled 'ENV' are depicted as points.



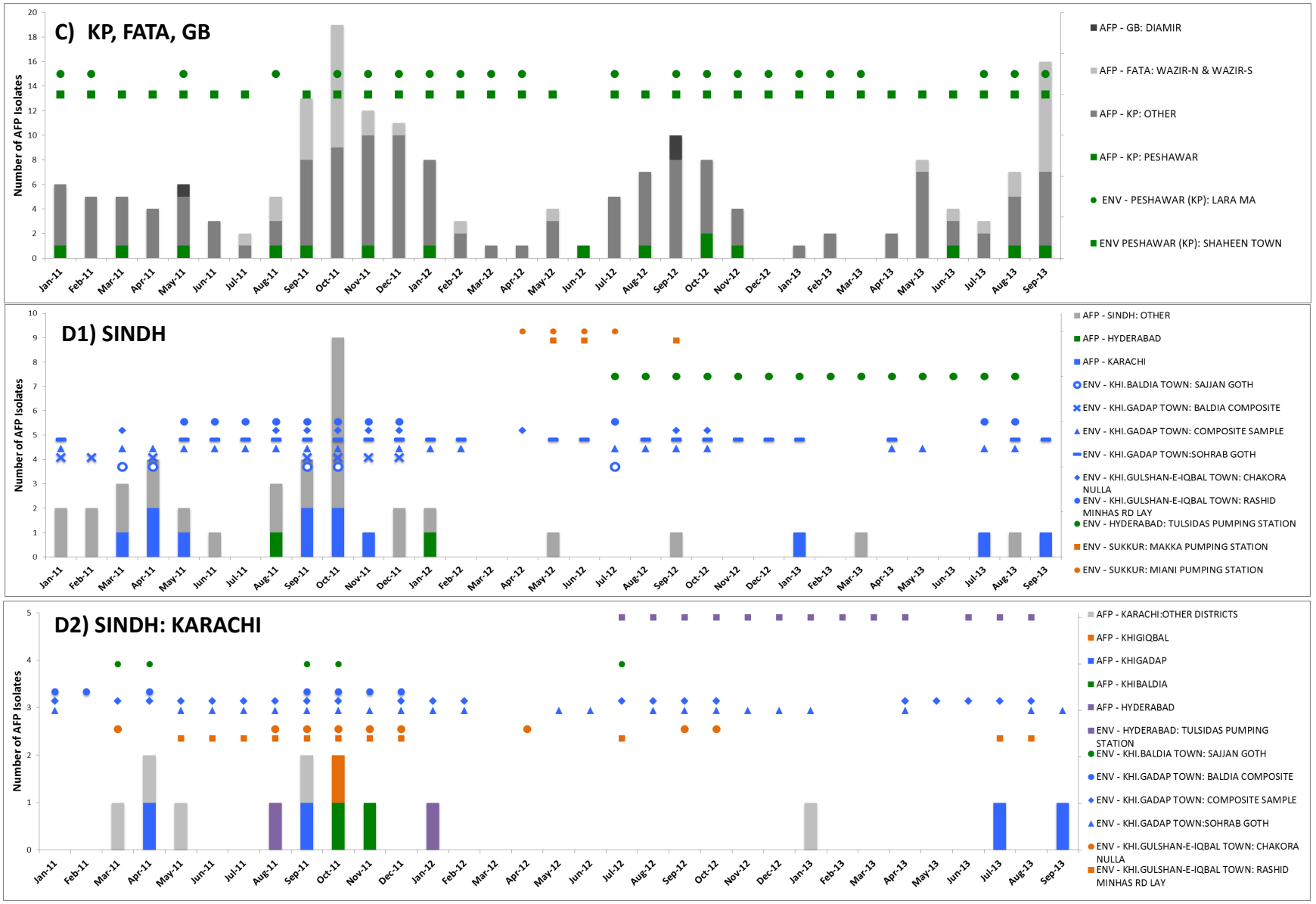

C) KP, FATA, GB
Number of AFP Isolates
AFP - GB: DIAMIR
AFP - FATA: WAZIR-N & WAZIR-S
AFP - KP: OTHER
AFP - KP: PESHAWAR
ENV - PESHAWAR (KP): LARA MA
ENV PESHAWAR (KP): SHAHEEN TOWN
D1) SINDH
Number of AFP Isolates
AFP - SINDH: OTHER
AFP - HYDERABAD
AFP - KARACHI
ENV - KHI.BALDIA TOWN: SAJJAN GOTH
ENV - KHI.GADAP TOWN: BALDIA COMPOSITE
ENV - KHI.GADAP TOWN: COMPOSITE SAMPLE
ENV - KHI.GADAP TOWN:SOHRAB GOTH
ENV - KHI.GULSHAN-E-IQBAL TOWN: CHAKORA NULLA
ENV - KHI.GULSHAN-E-IQBAL TOWN: RASHID MINHAS RD LAY
ENV - HYDERABAD: TULSIDAS PUMPING STATION
ENV - SUKKUR: MAKKA PUMPING STATION
ENV - SUKKUR: MIANI PUMPING STATION
D2) SINDH: KARACHI
Number of AFP Isolates
AFP - KARACHI:OTHER DISTRICTS
AFP - KHIGIQBAL
AFP - KHIGADAP
AFP - KHIBALDIA
AFP - HYDERABAD
ENV - HYDERABAD: TULSIDAS PUMPING STATION
ENV - KHI.BALDIA TOWN: SAJJAN GOTH
ENV - KHI.GADAP TOWN: BALDIA COMPOSITE
ENV - KHI.GADAP TOWN: COMPOSITE SAMPLE
ENV - KHI.GADAP TOWN:SOHRAB GOTH
ENV - KHI.GULSHAN-E-IQBAL TOWN: CHAKORA NULLA
ENV - KHI.GULSHAN-E-IQBAL TOWN: RASHID MINHAS RD LAY

Figure 6. Number of isolates detected by year, surveillance system type, and genetic cluster in Pakistan and Afghanistan – January 2011 through September 2013

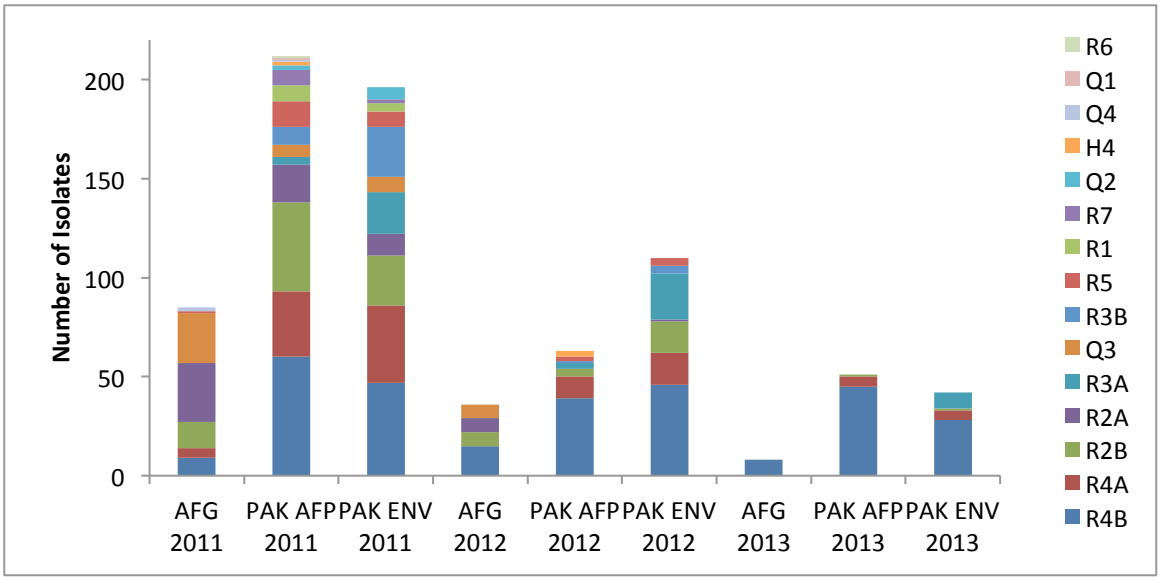
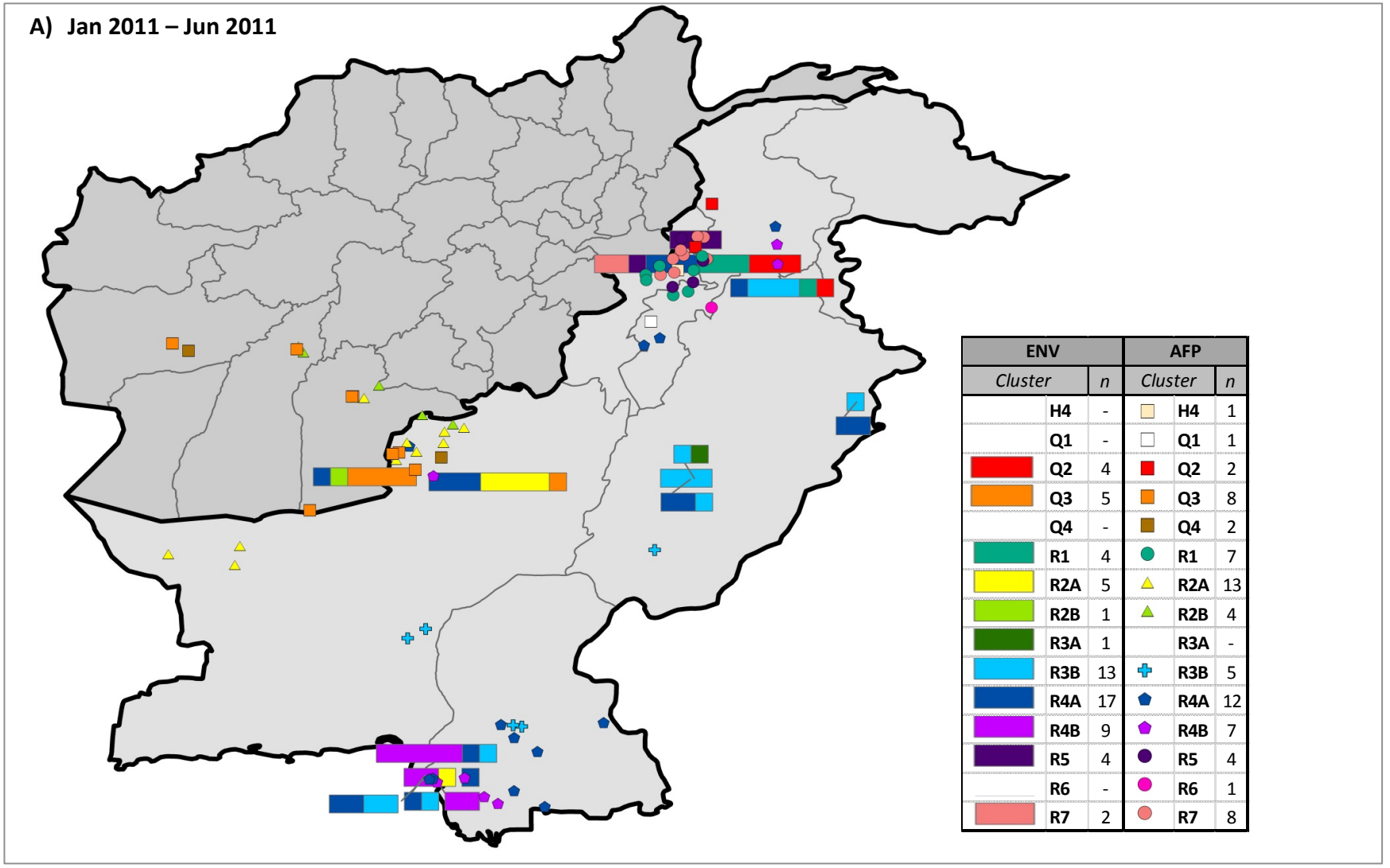
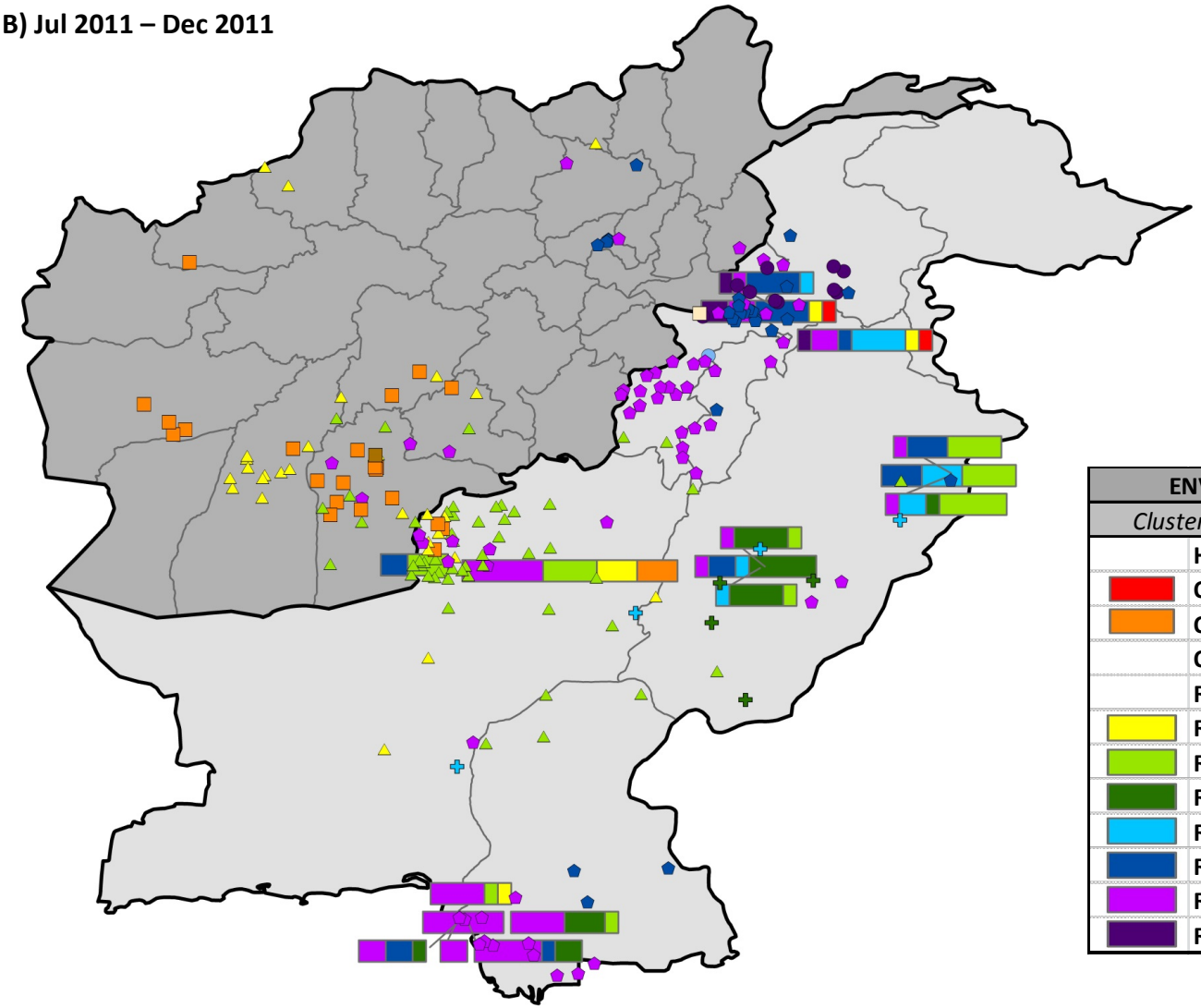


Figure 7. AFP polio cases and ENV isolates by genetic cluster**.** Corresponding colors indicate isolates of the same cluster. Bars represent isolates detected at each ENV site, while dots represent AFP cases, shown as random dots within the reporting district.

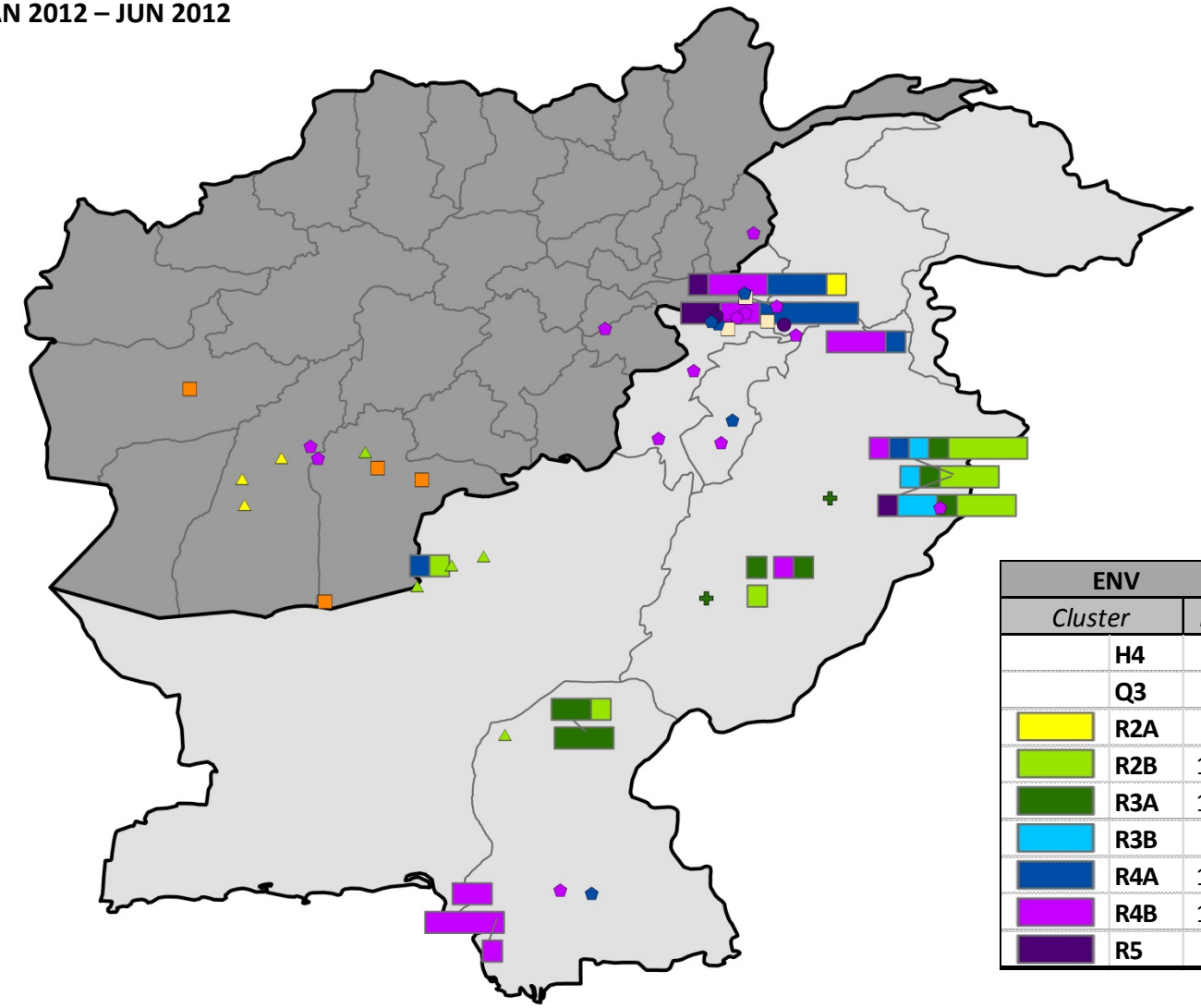


B) Jul 2011 – Dec 2011



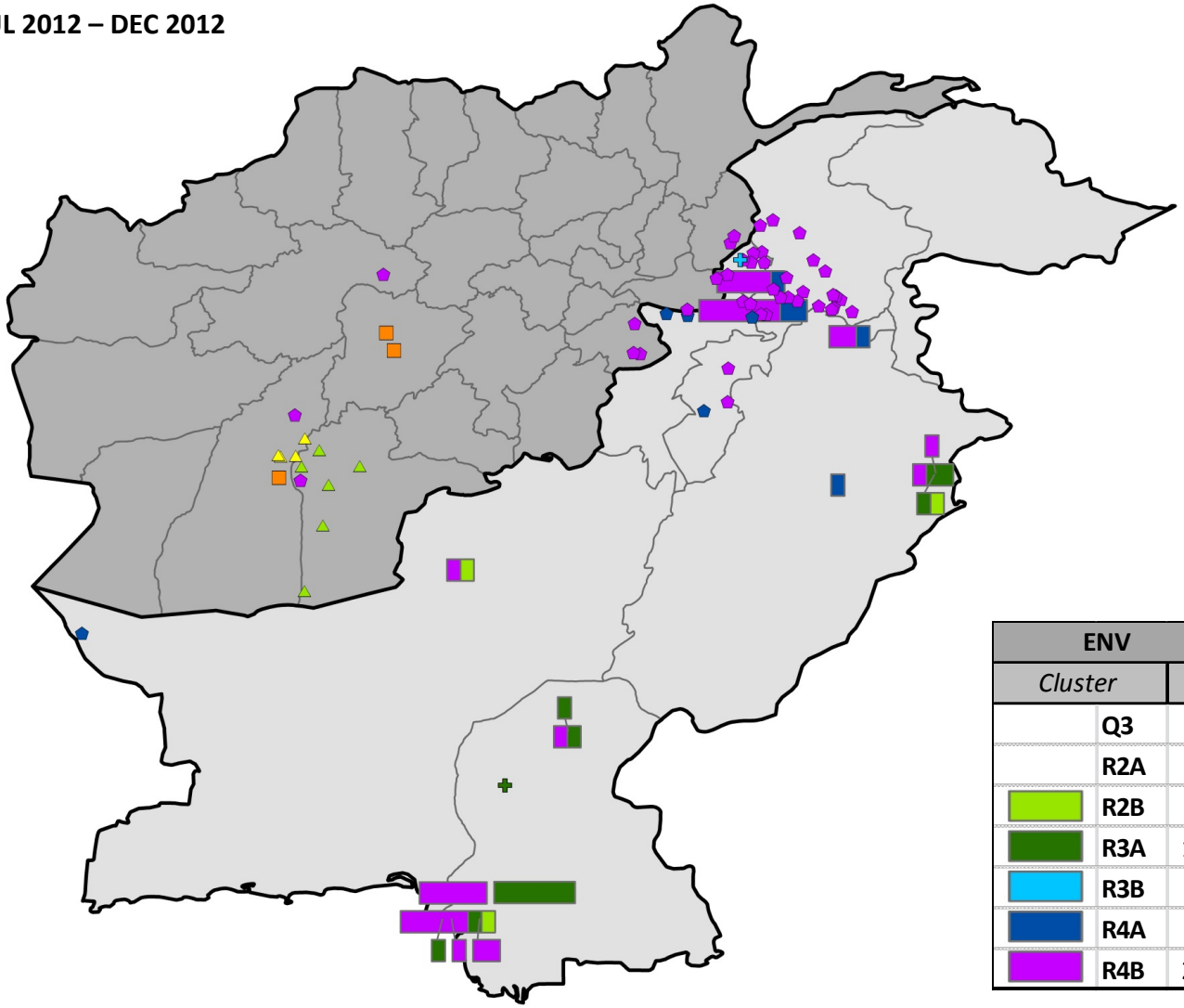
ENV			AFP		
Cluster		*n*	*Cluster*		*n*
	H4	-		**H4**	1
	Q2	2		**Q2**	-
	Q3	3		**Q3**	23
	Q4	-		**Q4**	1
	R1	-		**R1**	1
	R2A	6		**R2A**	36
	R2B	24		**R2B**	54
	R3A	20		**R3A**	4
	R3B	12		**R3B**	4
	R4A	22		**R4A**	26
	R4B	38		**R4B**	62
	R5	4		**R5**	10

C) JAN 2012 – JUN 2012



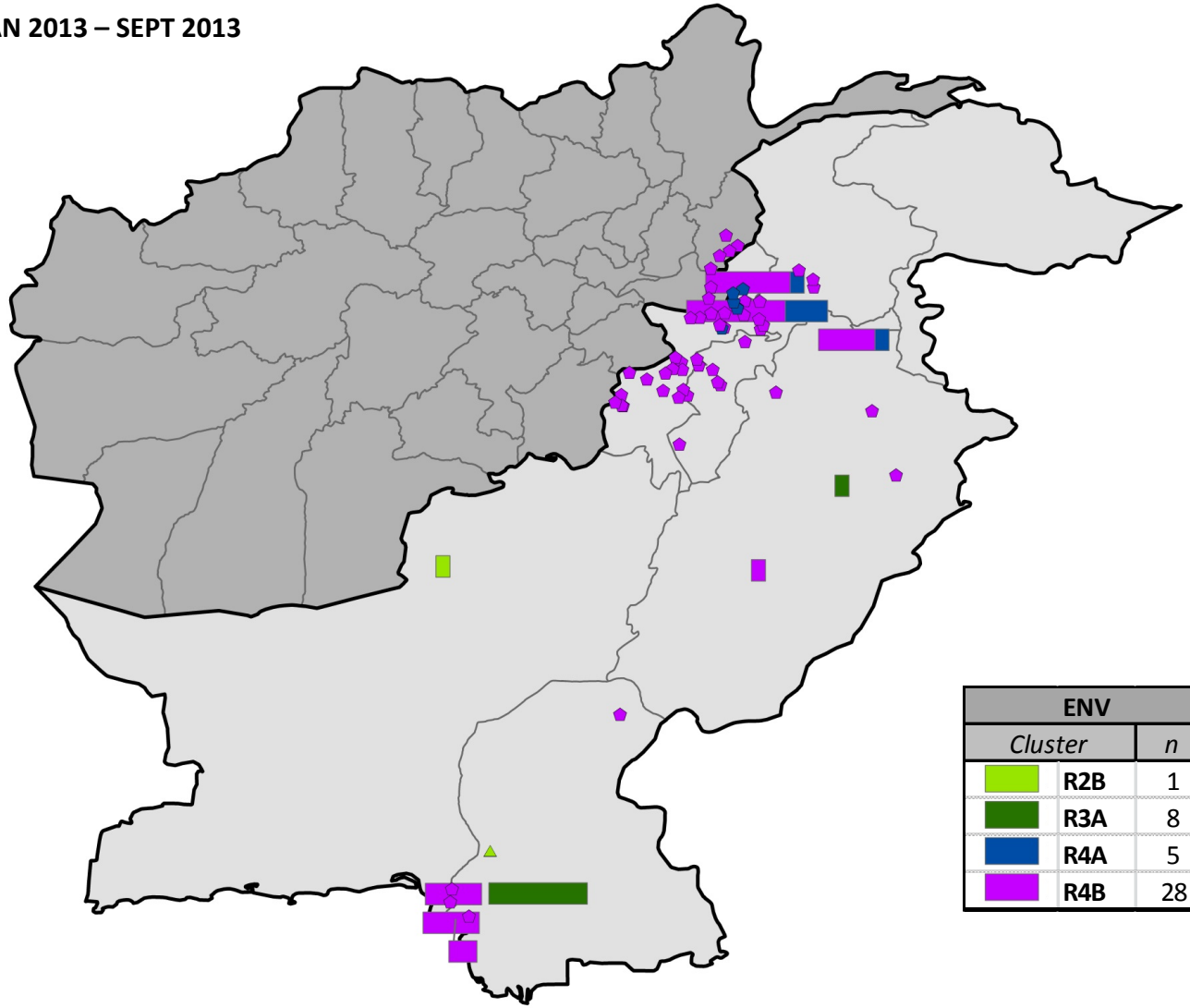
ENV			AFP		
Cluster		*n*	*Cluster*		*n*
	H4	-		**H4**	3
	Q3	-		**Q3**	4
	R2A	1		**R2A**	3
	R2B	13		**R2B**	5
	R3A	10		**R3A**	2
	R3B	4		**R3B**	-
	R4A	11		**R4A**	6
	R4B	17		**R4B**	14
	R5	4		**R5**	2

D) JUL 2012 – DEC 2012



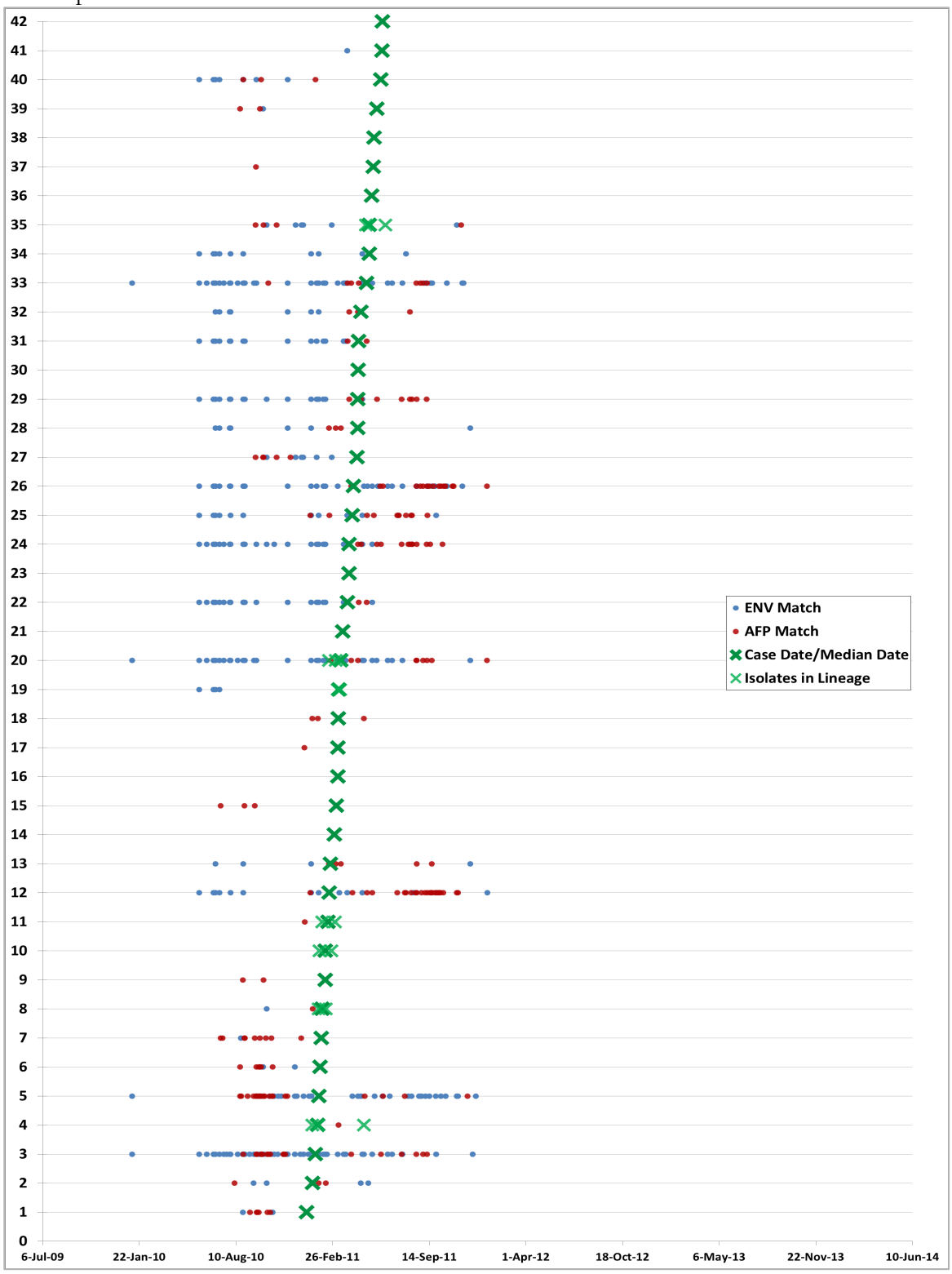
ENV			AFP		
Cluster		*n*	*Cluster*		*n*
	Q3	-		**Q3**	3
	R2A	-		**R2A**	4
	R2B	3		**R2B**	6
	R3A	13		**R3A**	1
	R3B	-		**R3B**	1
	R4A	5		**R4A**	5
	R4B	29		**R4B**	40

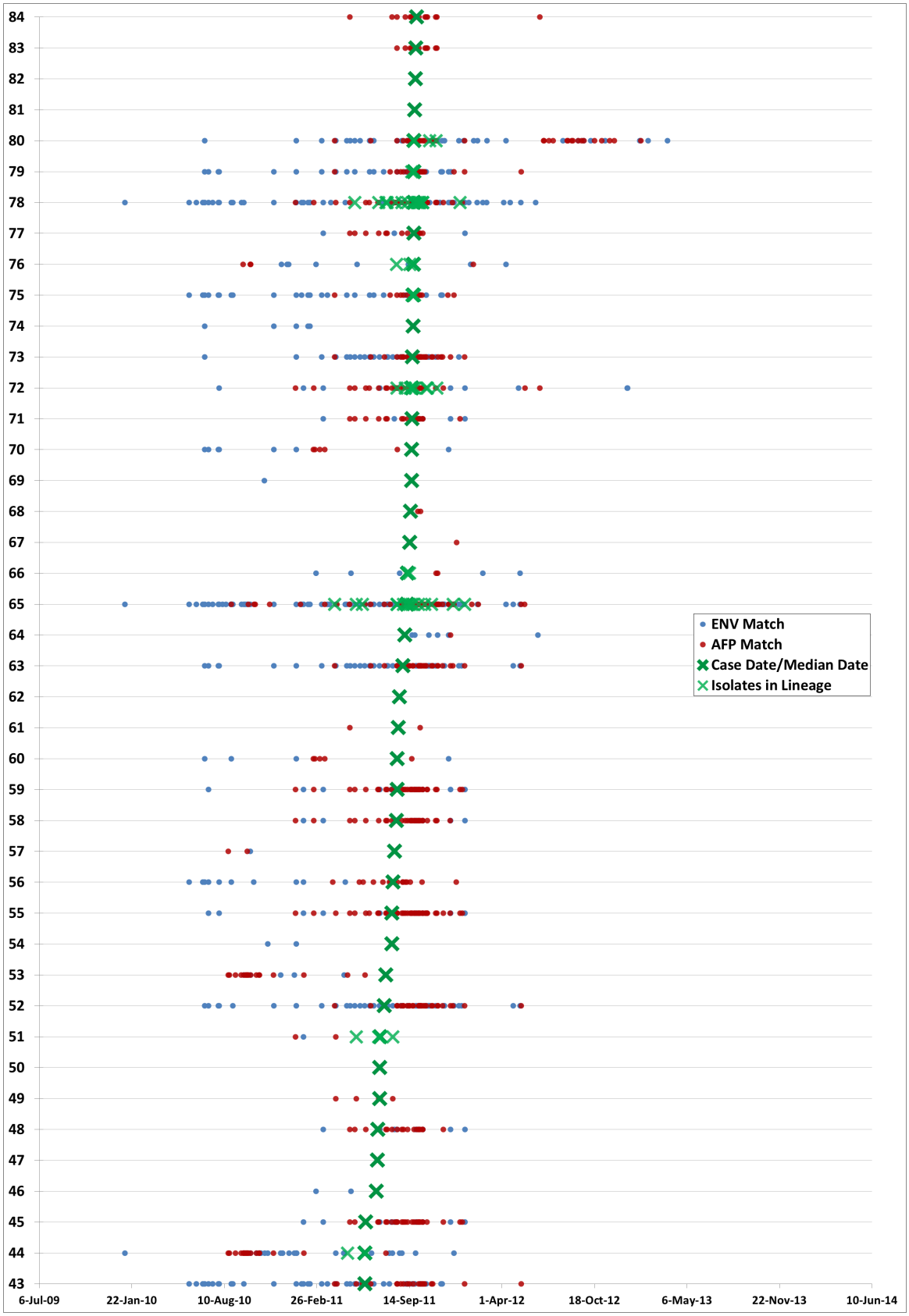
E) JAN 2013 – SEPT 2013

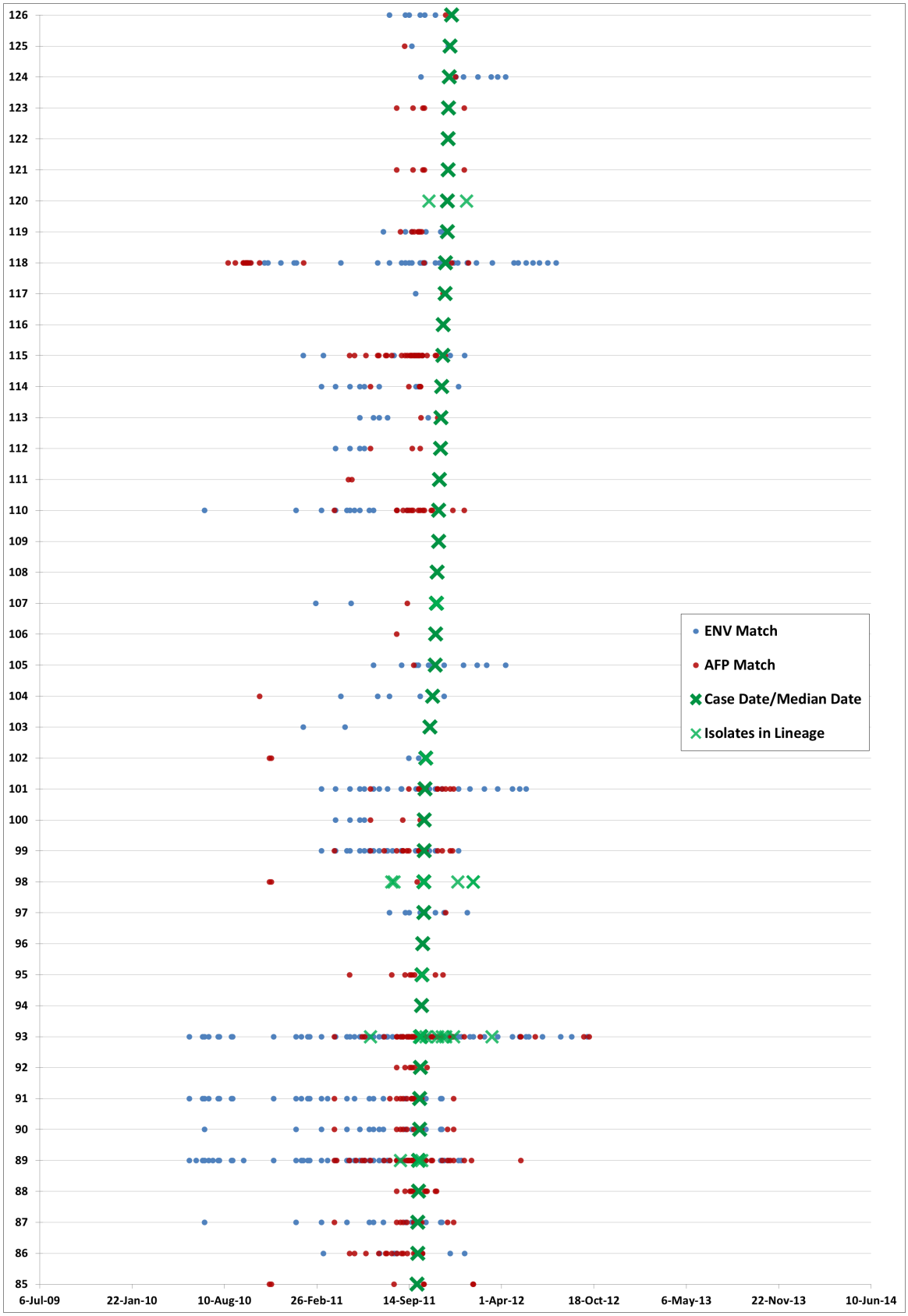


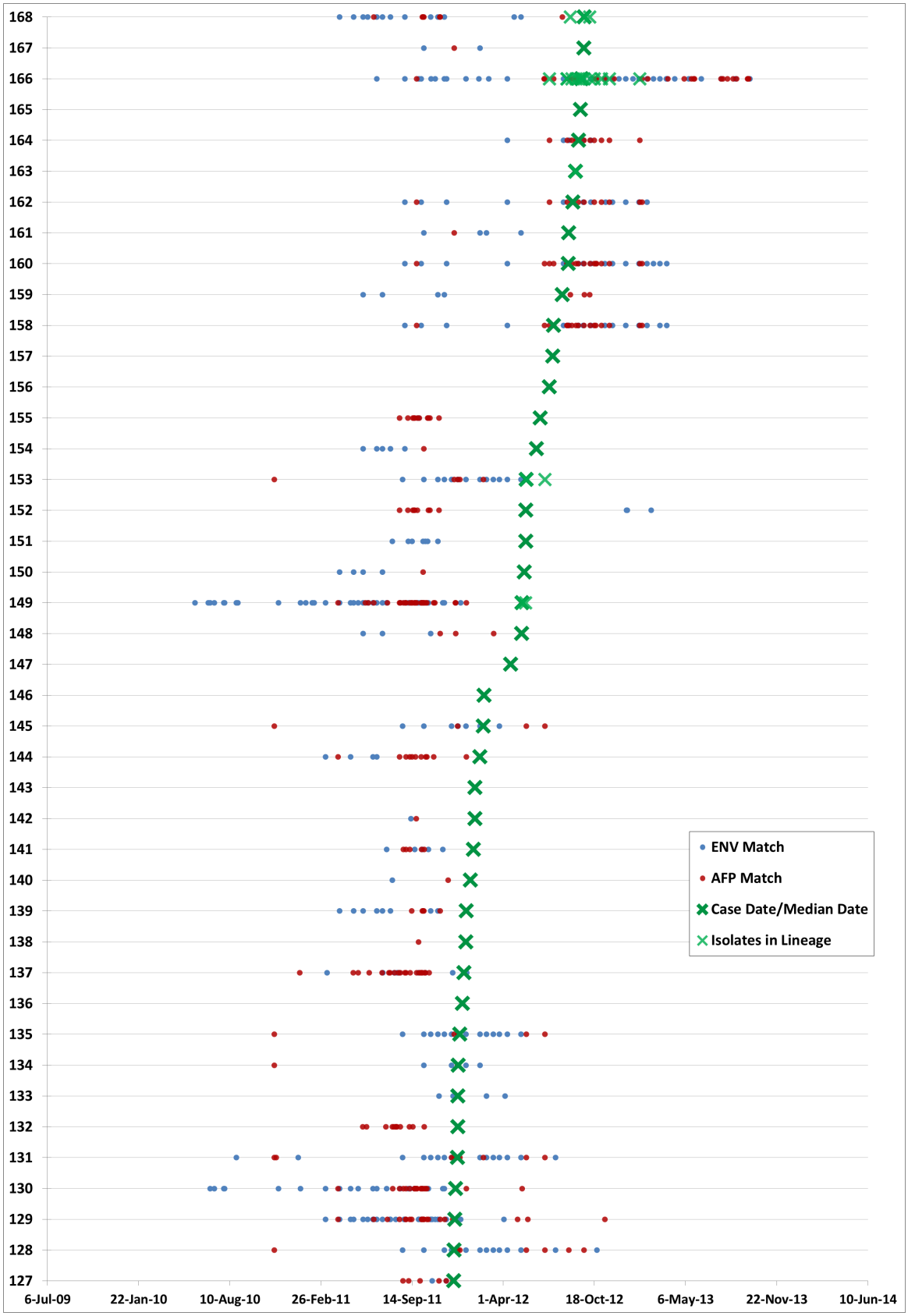
ENV		AFP	
Cluster	*n*	*Cluster*	*n*
R2B	1	**R2B**	1
R3A	8	**R3A**	
R4A	5	**R4A**	5
R4B	28	**R4B**	53

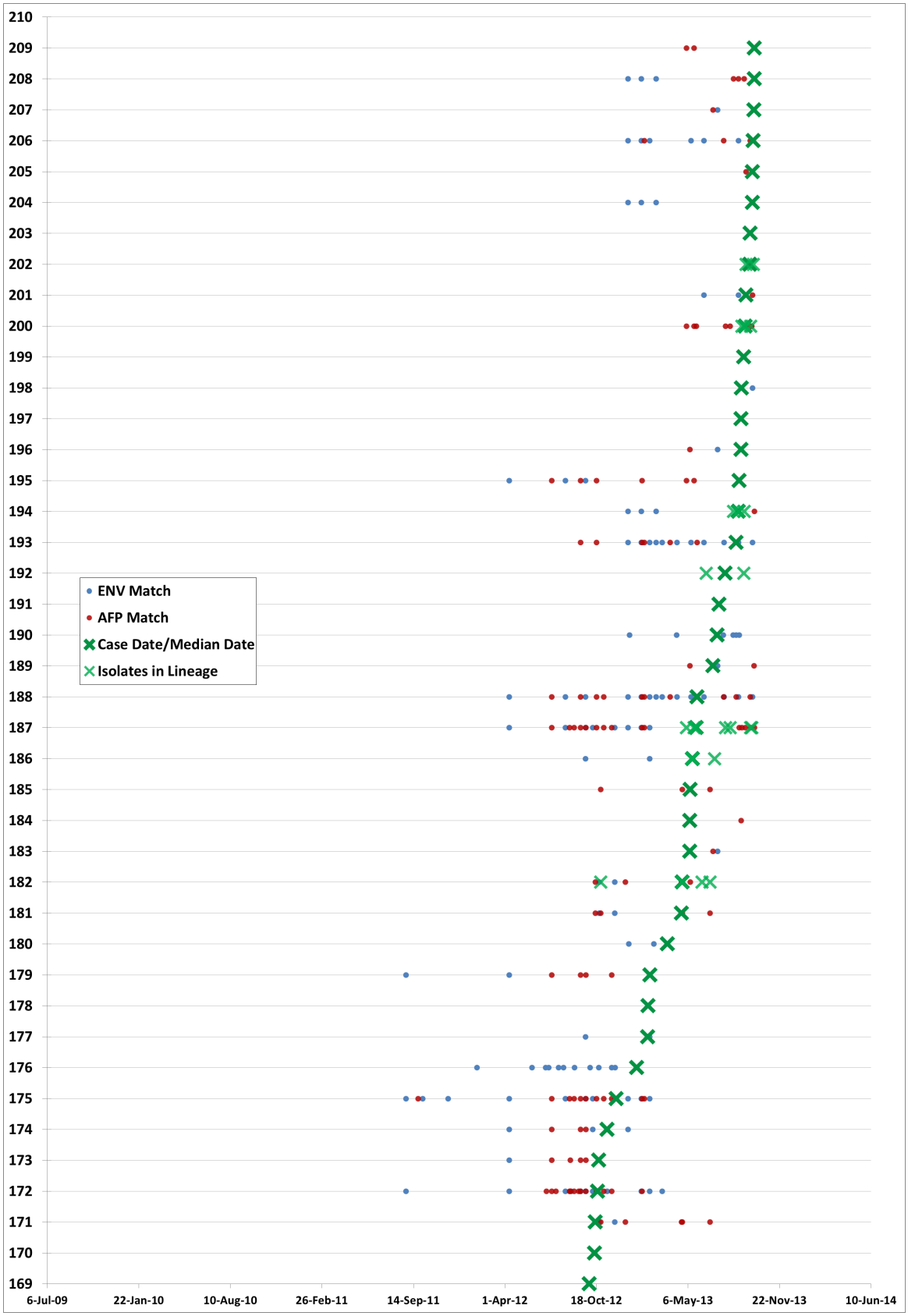
Figure 8. Circulation of sub-lineages and ≥99.0% ENV and matches by ID (y-axis) ordered by median sub-lineage date. Thin green X's depict AFP isolates in sub-lineage while thick green X depicts median date for sub-lineage. Red dots indicate an AFP isolate, while blue dots depict ENV isolates.




84
83
82
81
80
79
78
77
76
75
74
73
72
71
70
69
68
67
66
65
64
63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
47
46
45
44
43
6-Jul-09
22-Jan-10
10-Aug-10
26-Feb-11
14-Sep-11
1-Apr-12
18-Oct-12
6-May-13
22-Nov-13
10-Jun-14
ENV Match
AFP Match
Case Date/Median Date
Isolates in Lineage


126
125
124
123
122
121
120
119
118
117
116
115
114
113
112
111
110
109
108
107
106
105
104
103
102
101
100
99
98
97
96
95
94
93
92
91
90
89
88
87
86
85
6-Jul-09
22-Jan-10
10-Aug-10
26-Feb-11
14-Sep-11
1-Apr-12
18-Oct-12
6-May-13
22-Nov-13
10-Jun-14
ENV Match
AFP Match
Case Date/Median Date
Isolates in Lineage


168
167
166
165
164
163
162
161
160
159
158
157
156
155
154
153
152
151
150
149
148
147
146
145
144
143
142
141
140
139
138
137
136
135
134
133
132
131
130
129
128
127
6-Jul-09
22-Jan-10
10-Aug-10
26-Feb-11
14-Sep-11
1-Apr-12
18-Oct-12
6-May-13
22-Nov-13
10-Jun-14
ENV Match
AFP Match
Case Date/Median Date
Isolates in Lineage


210
209
208
207
206
205
204
203
202
201
200
199
198
197
196
195
194
193
192
191
190
189
188
187
186
185
184
183
182
181
180
179
178
177
176
175
174
173
172
171
170
169
ENV Match
AFP Match
Case Date/Median Date
Isolates in Lineage
6-Jul-09
22-Jan-10
10-Aug-10
26-Feb-11
14-Sep-11
1-Apr-12
18-Oct-12
6-May-13
22-Nov-13
10-Jun-14

APPENDICES

Appendix 1. Genetic network visualization of Pakistan AFP isolates. Connections depict ≥99.5% match. Connected Isolates were grouped together to form sub-lineages. AFP isolates are numbered by onset date, with lower numbers corresponding to earlier onset dates. Network visualization was conducted in Gephi® (Version 0.8.2). Isolates are colored by year: Blue, 2011; Yellow 2012; Red