

# Combining traditional and molecular epidemiology methods to quantify local HIV transmission among foreign-born residents

Diana M. Tordoff<sup>a,b</sup>, Susan Buskin<sup>a,c</sup>, Richard Lechtenberg<sup>c</sup>,  
Matthew R. Golden<sup>a,c,d</sup>, Roxanne P. Kerani<sup>a,c,d,\*</sup>  
and Joshua T. Herbeck<sup>b,\*</sup>

**Objectives:** We evaluated the ability for molecular epidemiology to augment traditional HIV surveillance beyond the detection of clusters for outbreak investigation. To do this, we address a question of interest to Public Health – Seattle and King County: what proportion of HIV diagnoses among people born outside of the United States are acquired locally?

**Design:** King County residents diagnosed with HIV, 2010–2018.

**Methods:** We linked HIV-1 *pol* gene sequences to demographic information obtained from the National HIV Surveillance System, Public Health – Seattle and King County case investigation and partner services interviews. We determined the likely location of HIV acquisition based on HIV testing, travel histories and cluster-based molecular analyses.

**Results:** Among 2409 people diagnosed with HIV, 798 (33%) were born outside of the United States. We inferred the location of acquisition for 77% of people born outside of the United States: 26% likely acquired HIV locally in King County (of whom 69% were MSM, 16% heterosexual), and 51% likely acquired HIV outside of King County (primarily outside of the United States). Of this 77% of people for whom we inferred the location of HIV acquisition, 45% were determined using traditional epidemiology methods and an additional 32% were inferred using molecular epidemiology methods.

**Conclusion:** We found that the National HIV Surveillance System misclassified the majority of HIV-infected foreign-born residents as ‘new’ local infections, and that these cases contribute to an overestimate of local incidence. Our findings highlight how molecular epidemiology can augment traditional HIV surveillance activities and provide useful information to local health jurisdictions beyond molecular cluster detection.

Copyright © 2020 The Author(s). Published by Wolters Kluwer Health, Inc.

*AIDS* 2021, **35**:655–664

**Keywords:** HIV, immigrants, molecular epidemiology, phylogenetics, surveillance

<sup>a</sup>Department of Epidemiology, <sup>b</sup>International Clinical Research Center, Department of Global Health, University of Washington, <sup>c</sup>HIV/STD Program, Public Health – Seattle & King County, and <sup>d</sup>Department of Medicine, University of Washington, Seattle, Washington, USA.

Correspondence to Joshua T. Herbeck, International Clinical Research Center, Department of Global Health, University of Washington, Seattle, Washington, USA.

E-mail: herbeck@uw.edu, rkerani@uw.edu

\* Roxanne P. Kerani and Joshua T. Herbeck share senior authorship.

Received: 9 June 2020; revised: 15 November 2020; accepted: 23 November 2020.

DOI:10.1097/QAD.0000000000002783

ISSN 0269-9370 Copyright © 2020 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

## Introduction

Molecular epidemiology and viral sequence analyses are now commonly used to identify HIV infection clusters and study patterns of HIV transmission. In approaches based on pairwise genetic similarity, sequences that cluster together are inferred to have epidemiological links and potentially represent a shared transmission network, and can be used to guide public health responses or targeted prevention measures. HIV gene sequence data are reportable in many parts of the United States, and the Centers for Disease Control and Prevention (CDC) aggregates these data at a national level as part of the HIV Cluster Detection and Response (CDR) program within the National HIV Surveillance System (NHSS). The CDC requires that all state and selected local public health agencies use molecular cluster detection to improve the identification of emerging HIV outbreaks within their jurisdictions [1]. However, the marginal benefit of these and other types of molecular epidemiological analyses over traditional HIV surveillance for local public health agencies is unclear.

We aimed to assess the ability of molecular epidemiology to improve traditional surveillance beyond the detection of clusters for outbreak investigation. We used data from the NHSS and supplementary data from HIV case investigations and partner services interviews in King County, Washington to address a key question of interest to Public Health – Seattle and King County (PHSKC): What proportion of new HIV diagnoses among residents who were born outside of the United States were acquired locally in King County? King County, which includes the city of Seattle, has a larger proportion of residents who were born outside of the United States (22%) than the United States overall (13%) [2,3]. Since 2006, King County has experienced an increase in the proportion of HIV cases among people born outside of the United States, despite an overall decline in HIV incidence over that same period [3–5]. While some HIV-infected individuals certainly acquired HIV prior to arriving in the United States, there is also local acquisition [6–8]. The extent to which HIV infections among foreign-born people are locally acquired is unclear. For individuals who acquired HIV outside of the United States, these infections contribute to an overestimate of local incidence. These diagnoses are still of public health importance, with the goal of diagnosis and linkage to care as soon as possible after arrival to the United States. In addition, understanding the degree of local HIV acquisition is important for HIV prevention, and for assessing which populations are not being reached by current prevention efforts.

Several recent studies have used molecular epidemiology to examine transmission dynamics among foreign-born residents of the United States [6,7,9,10]. However, they only considered individuals from a single region of birth

(e.g. Africa and Latin America) [6,10], or did not utilize linked clinical and epidemiological data, such as HIV testing or travel histories [7,9]. In this study, we combine data collected through traditional HIV surveillance activities with information from HIV sequences collected from drug resistance tests, and develop an algorithmic approach to understanding the location of HIV acquisition for foreign-born United States residents. Our primary aim was to estimate the proportion of foreign-born individuals who likely acquired HIV locally.

## Methods

### Study population and data collection

Our study population includes all residents of King County, Washington who were diagnosed with HIV between 1 January 2010 and 31 December 2018. Our analysis used linked, deidentified data from four sources: NHSS; HIV surveillance case investigations that review the medical records of persons who test positive; HIV partner services interviews conducted by PHSKC; and HIV-1 *pol* gene sequences collected through the CDR program. In Washington State, providers and laboratories are required to report HIV diagnoses to the local health jurisdiction, and laboratories are required to report all positive HIV antibody tests, HIV RNA test results, and HIV genotype results. Genotype results include HIV *pol* gene sequences and are the basis of the CDR program within NHSS. All HIV gene sequences are obtained as part of routine clinical care and not for the sole purpose of CDR. For all positive HIV test results received by PHSKC, staff conduct case investigations to determine if cases have been previously reported in other states or if persons were diagnosed prior to arrival in the United States. PHSKC also attempts to offer partner services to all individuals newly diagnosed with HIV. The NHSS includes demographic information and data on likely mode of HIV acquisition, country of birth, and clinical measures (such as CD4<sup>+</sup> cell counts and viral load). Partner services interviews collect additional information on HIV testing history, date of arrival in the United States, and sexual behavior, including sexual histories related to travel abroad (Supplemental Digital Content). Non-English-speaking patients are interviewed with in-person or phone interpreters.

### Sequence analysis

HIV-1 sequences from the protease and reverse transcriptase (PR/RT) region of the *pol* gene were aligned with the HXB2 reference genome using the MAFFT algorithm [11]. We chose to conduct distance-based clustering analyses (in lieu of other phylogenetic or phylogeographic methods) as these are most easily reproducible by local health departments. We identified genetic similarity clusters of two or more individuals using Tamura-Nei (TN93) pairwise genetic distance with a 0.02 substitutions/site threshold [12]. At this distance

threshold, belonging to a cluster is consistent with belonging to the same transmission chain and being epidemiologically linked [13,14]. In addition, we conducted sensitivity analyses using distance thresholds of 0.025, 0.015 and 0.01. HIV subtypes were determined using the COMET and REGA typing tools [15,16]. A phylogenetic tree was constructed using the FastTree approximate maximum likelihood method [17].

We only considered sequences obtained in King County (i.e. we did not include sequences that were not sampled in King County, which might be obtained via GenBank or the Los Alamos National Laboratory HIV Database). Any clusters, then, were suggestive of local transmission or acquisition. To increase the likelihood of detecting putative transmission clusters, we used all available PR/RT pol gene sequences from people who were residents of King County when they were diagnosed with HIV or AIDS (11 625 sequences from 6458 people) when identifying clusters. Variation in rates of clustering among subpopulations can be due to variation in transmission rates or variation in the time between infection and sequence collection, as sequences sampled in early infection are more likely to cluster [18,19]. Thus, we assessed the odds of clustering using logistic regression and adjusted for year of diagnosis and early HIV infection (defined as CD4<sup>+</sup> cell count >500 cells/ $\mu$ l at diagnosis).

### Determining location of HIV acquisition

The NHSS defines new HIV diagnoses based on residence in the United States at the time of a first documented positive HIV test. Therefore, some foreign-born individuals may have been diagnosed with HIV prior to arriving in the United States, but are counted as 'new' diagnoses in the NHSS if the individual does not have documentation of a previous positive test result. Using data collected through case investigation, PHSKC determines if individuals without such documentation were diagnosed prior to being a resident of King County (either outside of the United States or elsewhere within the United States). Thus, PHSKC's definition of local HIV cases excludes individuals for whom case investigation revealed they had a prior diagnosis (but are counted as a 'new' diagnosis in the NHSS).

Our primary analysis, described below, aims to determine the location of HIV acquisition for all individuals that PHSKC determined to be a local case of HIV (e.g. individuals without a prior diagnosis outside of King County). In addition, we estimate the proportion of all NHSS defined new HIV diagnosis that were locally acquired, abroad or outside of King County.

To determine the likely location of HIV acquisition for PHSKC-defined new diagnoses, we first used information collected from partner services interviews on date of last HIV negative test, date of arrival to the United States, and sex during travel outside of the United States (Fig. 1). We

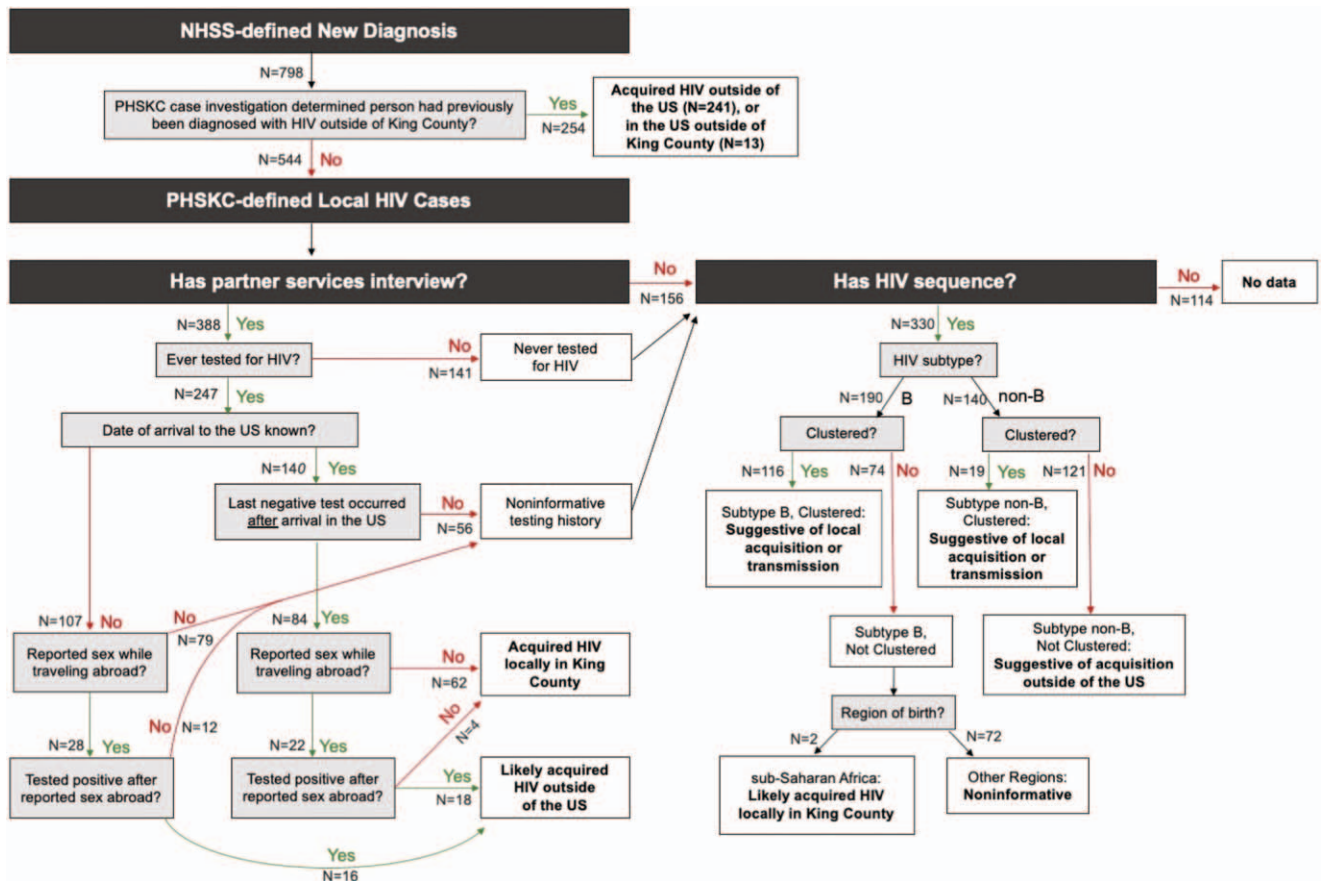
categorized individuals as likely having acquired HIV outside of the United States if they report a negative HIV test result after arrival in the United States and subsequently tested positive for HIV after having sex while traveling to their country of birth. We categorized individuals as likely having acquired HIV in King County (local acquisition) if they report a negative HIV test after arrival in the United States and no sex during subsequent travel abroad. We were unable to use partner services data to determine the location of HIV acquisition for individuals who never tested for HIV prior to their diagnosis; were missing a date of arrival to the United States, or; whose most recent negative HIV test was *prior* to their arrival in the United States (noninformative testing history).

Next, we used sequence data for individuals for whom we were unable to determine the likely location of acquisition based on partner services information alone. Individuals who clustered (regardless of HIV subtype) were assumed to be associated with either local acquisition or transmission. Examining HIV subtypes among individuals who were born in regions where B is not the most common subtype is informative for the location of infection [20]. Individuals who did not cluster and had a non-B HIV subtype likely acquired HIV outside of the United States (or their local transmission source was unsampled). Among individuals who did not cluster, those who were born in sub-Saharan Africa and have a subtype B HIV virus likely acquired HIV in the United States because subtypes C, D, F and circulating recombinant forms 01AE and 02AG are dominant in sub-Saharan Africa. Lastly, we are unable to infer the location of acquisition among individuals with subtype B who did not cluster and who were born in Asia, Latin/South American, Europe/Canada, Oceania or the Middle East/ North Africa because subtype B circulates more frequently in those regions [20].

All analyses were conducted in R version 3.4.2 (R Foundation for Statistical Computing, Vienna, Austria) and our scripts are available at <https://github.com/dianatorloff/KClocalHIV>. This study received ethical approval from the Washington State and University of Washington Institutional Review Boards.

## Results

From 2010 to 2018, 2409 King County residents were diagnosed with HIV and categorized as new infections based on the NHSS definition. Of these 2409 individuals, 798 were born outside of the United States. Among King County residents who were born outside of the United States: 37% were born in Latin/Southern America, 29% in sub-Saharan Africa, 15% in Asia, 9% in Europe or Canada, 5% in Oceania and 3% were born in the Middle East or North Africa (Table 1). One hundred and fifty-five people had an unknown or unspecified country of



**Fig. 1.** Decision tree for determining the location of HIV acquisition for people born outside of the United States who were designated as new diagnoses in the National HIV Surveillance System.

birth and were excluded from our analysis. The proportion of King County cases that were foreign-born increased from 17% in 2010 to 33% in 2017 (Supplemental Digital Content).

### Public Health – Seattle and King County case investigation

Among the 798 foreign-born people with a newly reported positive HIV test in the NHSS, PHSKC identified 254 (32%) individuals with a prior HIV diagnosis outside of the United States ( $N=241$ ) or in the United States but outside of King County ( $N=13$ ; Fig. 2). The proportion of individuals with a prior diagnosis varied by region, and was highest among people born in sub-Saharan Africa (Table 2A). The remaining 544 (68%) foreign-born individuals were categorized as local cases of HIV based on the PHSKC definition (Fig. 2). In the next two sections, we only present data for these 544 PHSKC-defined local HIV cases among people born outside of the United States.

### Partner services interviews

From 2010 to 2018, 1383 of PHSKC-defined new cases completed a partner services interview (82% of United

States-born v. 71% of foreign-born individuals). Among the 388 foreign-born people who completed a partner services interview, we determined that 9% (34/388) likely acquired HIV outside of the United States because they reported testing positive for HIV for the first time after having sex while traveling to their country of birth (Table 2B). We also estimated that 17% (66/388) of people acquired HIV in the United States because they reported a negative HIV test after arrival in the United States and no sex during subsequent travel abroad. For the remaining individuals, 36% (141/388) had never tested for HIV prior to their diagnosis and 38% (147/388) had their last negative test result prior to their arrival in the United States.

### Molecular epidemiology

There were 1448 people who had an available HIV gene sequence (80% of US-born v. 74% of foreign-born individuals). HIV subtype distribution generally mirrored individuals' regions of birth (Table 2C) [20], except for MSM, in which a higher proportion of MSM had HIV subtype B compared with non-MSM (across all regions of birth) (86 vs. 32%,  $P$  value  $<0.001$ ). The phylogenetic

**Table 1. Demographics of Public Health – Seattle and King County-defined local HIV cases in King County by region of birth, 2010–2018.**

	Latin/ South America N=203	Sub- Saharan Africa N=160	Asia N=84	Europe/ Canada N=51	Oceania N=29	Middle East/North Africa N=17	All foreign- born N=544	US-born N=1304
<b>Sex, n (%)</b>								
Cisgender women	24 (11.8%)	94 (58.8%)	19 (22.6%)	7 (13.7%)	1 (3.4%)	3 (17.6%)	148 (27.2%)	109 (8.4%)
Cisgender men	176 (86.7%)	66 (41.2%)	65 (77.4%)	44 (86.3%)	26 (89.7%)	14 (82.4%)	391 (71.9%)	1180 (90.5%)
Transgender women	3 (1.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (6.9%)	0 (0.0%)	5 (0.9%)	13 (1.0%)
Transgender men	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (0.2%)
<b>Age at diagnosis (years), n (%)</b>								
<25	19 (9.4%)	17 (10.6%)	16 (19.0%)	3 (5.9%)	2 (6.9%)	0 (0.0%)	57 (10.5%)	203 (15.6%)
25–34	75 (36.9%)	32 (20.0%)	23 (27.4%)	17 (33.3%)	13 (44.8%)	6 (35.3%)	166 (30.5%)	482 (37.0%)
35–44	67 (33.0%)	40 (25.0%)	20 (23.8%)	19 (37.3%)	5 (17.2%)	8 (47.1%)	159 (29.2%)	281 (21.5%)
>45	42 (20.7%)	71 (44.4%)	25 (29.8%)	12 (23.5%)	9 (31.0%)	3 (17.6%)	162 (29.8%)	338 (25.9%)
<b>Transmission category, n (%)</b>								
PWID	4 (2.0%)	1 (0.6%)	4 (4.8%)	5 (9.8%)	0 (0.0%)	0 (0.0%)	14 (2.6%)	221 (16.9%)
MSM	140 (69.0%)	12 (7.5%)	51 (60.7%)	38 (74.5%)	23 (79.3%)	13 (76.5%)	277 (50.9%)	1053 (80.8%)
Heterosexual	24 (11.8%)	45 (28.1%)	8 (9.5%)	2 (3.9%)	0 (0.0%)	1 (5.9%)	80 (14.7%)	65 (5.0%)
Unknown	34 (16.7%)	102 (63.7%)	22 (26.2%)	9 (17.6%)	4 (13.8%)	3 (17.6%)	174 (32.0%)	90 (6.9%)
<b>CD4<sup>+</sup> cell count at diagnosis (cells/μl), n (%)</b>								
<200	60 (30.0%)	71 (45.2%)	25 (29.8%)	7 (14.3%)	6 (20.7%)	4 (23.5%)	173 (32.3%)	237 (18.7%)
200–500	97 (48.5%)	60 (38.2%)	39 (46.4%)	21 (42.9%)	18 (62.1%)	9 (52.9%)	244 (45.5%)	530 (41.7%)
>500	43 (21.5%)	26 (16.6%)	20 (23.8%)	21 (42.9%)	5 (17.2%)	4 (23.5%)	119 (22.2%)	503 (39.6%)
Completed PS interview, n (%)	163 (80.3%)	88 (55.0%)	59 (70.2%)	41 (80.4%)	26 (89.7%)	11 (64.7%)	388 (71.3%)	995 (81.5%)
Available PR/RT sequence, n (%)	167 (82.3%)	116 (72.5%)	63 (75.0%)	31 (60.8%)	15 (51.7%)	12 (70.6%)	404 (74.3%)	1044 (80.1%)

This table excludes 155 individuals with an unknown country of birth and 406 individuals with a date of diagnosis that precede the date of their first laboratory or medical visit while a resident of King County and who were thus not defined as a local cases of HIV by Public Health – Seattle & King County. PR/RT, protease/reverse transcriptase; PS, partner services; PWID, person who injects drugs.

tree highlights the subtype diversity in King County (Fig. 3).

Among all individuals with an available sequence, 1104 (76%) clustered into 295 genetic similarity clusters. Fewer foreign-born people clustered compared with US-born people (44 vs. 89%, *P* value <0.001). In a multivariate regression model adjusting for early infection and year of diagnosis, foreign-born individuals had a significantly lower odds of clustering [adjusted odds ratio (AOR) 0.23; 95% confidence interval (CI): 0.19, 0.28; Supplemental Digital Content]. This association was true for all regions, but varied substantially between regions of birth, such that the odds of clustering was lowest for sub-Saharan Africa (AOR 0.06; 95% CI: 0.04, 0.09) and highest, but still below unity, for Latin/South America (AOR 0.48; 95% CI: 0.36, 0.63). We also found that foreign-born individuals were more likely to cluster with other foreign-born individuals (AOR 1.88; 95% CI: 1.28, 2.75) compared with US-born individuals.

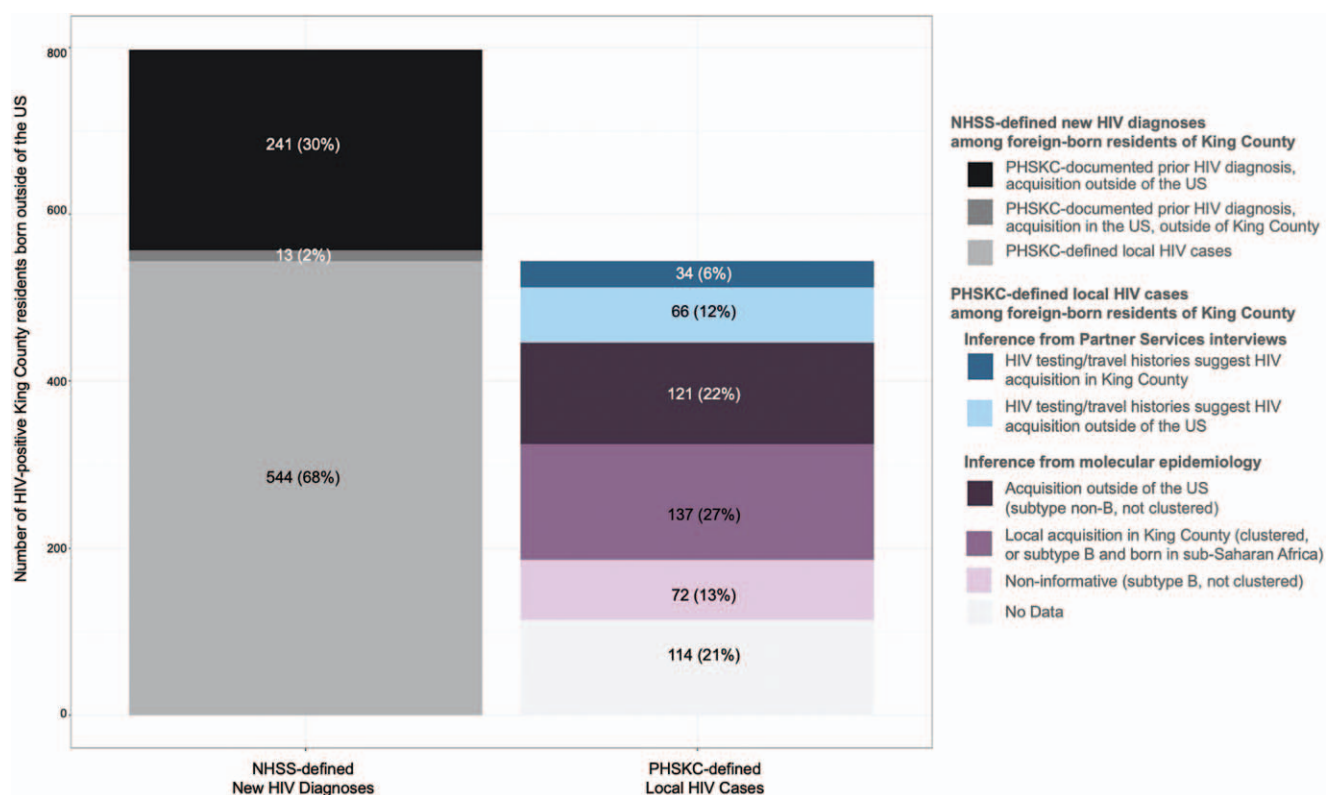
### Combined inference on the location of HIV acquisition

When we combine our inference from traditional surveillance data with additional information from molecular analyses, we are able to infer the location of HIV acquisition for 611 (77%) of all 798 HIV-infected foreign-born residents of King County categorized as new diagnoses by the NHSS (Fig. 2): PHSKC case investigation determined that 254 (32%) people acquired HIV outside of King County (primarily outside of the

United States); additional analysis of partner services data determined the location of HIV acquisition (either local or outside of the United States) for another 100 (13%) people; and the location of HIV acquisition for the remaining 258 (32%) people was inferred using molecular epidemiology methods. We were unable to determine the location of acquisition for 186 people because our inference was noninformative (*N*=72) or because they neither completed a partner services interview nor had an available PR/RT sequence (*N*=114).

Our combined inference suggests that 394 (49%) likely acquired HIV outside of the United States, 13 (2%) acquired HIV in the United States outside of King County, and 205 (26%) likely acquired HIV locally in King County (Table 2D). This varied by region and transmission group: Individuals from Latin/Southern America had the highest proportion of individuals who likely acquired HIV locally (43%), while individuals born in sub-Saharan African were most likely to acquire HIV outside of the United States (78%). This may reflect differences in transmission risk among individuals from these regions of birth, since 42% of MSM likely acquired HIV locally compared with 67% of heterosexuals and 66% of people with unknown transmission who likely acquired HIV outside of the United States.

Figures 1 and 2 show the incremental information gained over traditional epidemiologic data from molecular cluster analyses, reported as the additional proportion of individuals who did or did not cluster by HIV subtype.



**Fig. 2. Combined analysis of HIV testing and travel histories obtained during partner services interviews and molecular epidemiology among all 798 foreign-born residents of King County diagnosed with HIV, 2010–2018.** The National HIV Surveillance System defines new HIV diagnoses as a first documented positive HIV test. Therefore, some foreign-born individuals may have been previously diagnosed with HIV prior to arriving in the United States, but are counted as ‘new’ diagnoses in the National HIV Surveillance System if the individual does not have documentation of a previous positive test result. Public Health – Seattle and King County identified 254 individuals with a prior diagnosis date, thus only 544 foreign-born residents of King County were categorized as new diagnoses per the Public Health – Seattle and King County definition (left column). The right column indicates inference from self-reported HIV testing and travel histories as well as date of arrival to the United States using data from Partner Services interview in blue (see Fig. 1 for decision tree), and the additional information gained through molecular epidemiology analyses in purple.

Beyond the 66 individuals whose HIV testing and travel histories suggested they acquired HIV locally, 135 additional individuals clustered (116 subtype B, 19 non-B), suggesting they are linked with local HIV transmission or acquisition. We also observed that 121 had a non-B subtype and did not cluster, which is suggestive of acquisition outside of King County. Lastly, 74 had a subtype B virus and did not cluster: Two of these individuals were born in sub-Saharan Africa, suggesting they acquired HIV locally, while 72 were born in regions where subtype B is in circulation, and thus are noninformative.

Lastly, we observed concordance between clustering and inference from partner services interviews. There were 315 foreign-born individuals who had both an available sequence and partner services interview. Among individuals whose HIV testing and travel histories suggested that they likely acquired HIV locally, 68% belonged to a genetic cluster. This might suggest that 32% of persons who transmitted HIV to these foreign-born residents are

unsampled, a proportion that is concordant with the overall HIV sequence coverage in King County. In contrast, only 25% of individuals whose HIV testing and travel histories suggested that they acquired HIV outside of the United States belonged to a genetic cluster. This suggests that these individuals may have transmitted HIV to another individual who resides in King County, or immigrated to the United States with the person they acquired HIV from or transmitted HIV to (e.g. a spouse). Sensitivity analyses obtained comparable results.

## Discussion

Overall, we inferred the location of acquisition for 77% of people living with HIV (PLWH) in King County who were born outside of the United States. We estimate that a minimum of 26% of HIV-infected foreign-born residents likely acquired HIV locally in King County between

**Table 2. Inference from partner services and molecular epidemiology on the probable location of HIV acquisition among foreign-born residents of King County diagnosed 2010–2018, by region of birth and transmission category.**

	By region of birth										By transmission category				
	All Foreign-born					Sub-Saharan Africa					Middle East/North Africa				
	Latin/South America	Asia	Europe/Canada	Oceania		Latin/South America	Asia	Europe/Canada	Oceania		Middle East/North Africa	MSM	PWID	Heterosexual	Unknown
<b>A. All NHSS diagnoses, PHSKC case investigation</b>	N = 798	N = 245	N = 121	N = 65	N = 38	N = 19	N = 337	N = 18	N = 134	N = 310					
Prior diagnosis															
Acquisition outside of the United States	241 (30.2%)	39 (15.9%)	144 (46.5%)	35 (28.9%)	9 (23.7%)	2 (10.5%)	55 (16.3%)	4 (22.2%)	52 (38.8%)	130 (41.9%)					
Acquisition in the United States, outside of King County	13 (1.6%)	3 (1.2%)	6 (1.9%)	2 (1.7%)	0 (0.0%)	0 (0.0%)	5 (1.5%)	0 (0.0%)	2 (1.5%)	6 (1.9%)					
PHSKC-defined local case	544 (68.2%)	203 (82.9%)	160 (51.6%)	84 (69.4%)	29 (76.3%)	17 (89.5%)	277 (82.2%)	14 (77.8%)	80 (59.7%)	174 (56.1%)					
<b>B. PHSKC-defined local case with completed PS interview</b>	N = 388	N = 163	N = 88	N = 59	N = 26	N = 11	N = 220	N = 10	N = 62	N = 96					
Local acquisition	66 (17.0%)	34 (20.9%)	13 (14.8%)	10 (16.9%)	4 (9.8%)	2 (18.2%)	47 (21.4%)	1 (10.0%)	6 (9.7%)	12 (12.5%)					
Foreign acquisition	34 (8.8%)	15 (9.2%)	8 (9.1%)	7 (11.9%)	1 (2.4%)	1 (9.1%)	20 (9.1%)	1 (10.0%)	6 (9.7%)	6 (6.2%)					
Noninformative HIV testing history <sup>b</sup>	147 (37.9%)	63 (38.7%)	24 (27.3%)	18 (30.5%)	12 (46.2%)	5 (45.5%)	103 (46.8%)	5 (50.0%)	18 (29.0%)	23 (24.0%)					
Never tested for HIV prior to diagnosis	141 (36.3%)	51 (31.3%)	43 (48.9%)	24 (40.7%)	11 (26.8%)	3 (27.3%)	50 (22.7%)	3 (30.0%)	32 (51.6%)	55 (57.3%)					
<b>C. PHSKC-defined local cases with PR/RT sequence</b>	N = 404	N = 167	N = 116	N = 63	N = 31	N = 15	N = 209	N = 13	N = 68	N = 117					
HIV-1 subtype															
B	243 (60.1%)	159 (95.2%)	7 (6.0%)	33 (52.4%)	11 (73.3%)	7 (58.3%)	180 (86.1%)	10 (76.9%)	20 (29.4%)	36 (30.8%)					
A	41 (10.1%)	2 (1.2%)	31 (26.7%)	3 (4.8%)	0 (0.0%)	2 (16.7%)	8 (3.8%)	0 (0.0%)	15 (22.1%)	18 (15.4%)					
C	64 (15.8%)	1 (0.6%)	59 (50.9%)	3 (4.8%)	0 (0.0%)	0 (0.0%)	1 (0.5%)	1 (7.7%)	22 (32.4%)	40 (34.2%)					
CRF 01AE	21 (5.2%)	0 (0.0%)	0 (0.0%)	17 (27.0%)	1 (3.2%)	0 (0.0%)	9 (4.3%)	0 (0.0%)	5 (7.4%)	5 (5.1%)					
CRF 02AG	12 (3.0%)	0 (0.0%)	11 (9.5%)	0 (0.0%)	0 (0.0%)	1 (8.3%)	2 (1.0%)	0 (0.0%)	2 (2.9%)	8 (6.8%)					
Other	23 (5.7%)	5 (3.0%)	8 (6.9%)	7 (11.1%)	1 (3.2%)	2 (16.7%)	9 (4.3%)	2 (15.4%)	4 (5.9%)	9 (7.7%)					
Clustered	176 (43.6%)	91 (54.5%)	16 (13.8%)	33 (52.4%)	21 (67.7%)	6 (50.0%)	119 (56.9%)	7 (53.8%)	25 (36.8%)	28 (23.9%)					
<b>D. All NHSS diagnoses, combined inference</b>	N = 798	N = 245	N = 310	N = 121	N = 65	N = 38	N = 337	N = 18	N = 134	N = 310					
Acquisition outside of the United States <sup>b</sup>	394 (49.4%)	56 (22.9%)	243 (78.4%)	59 (48.8%)	16 (24.6%)	7 (36.8%)	91 (27.0%)	6 (33.3%)	90 (67.2%)	205 (66.1%)					
Acquisition in the United States, outside of King County	13 (1.6%)	3 (1.2%)	6 (1.9%)	2 (1.7%)	2 (3.1%)	0 (0.0%)	5 (1.5%)	0 (0.0%)	2 (1.5%)	6 (1.9%)					
Local acquisition in King County <sup>c</sup>	205 (25.7%)	105 (42.9%)	23 (7.4%)	37 (30.6%)	24 (36.9%)	6 (31.6%)	142 (42.1%)	7 (38.9%)	27 (20.1%)	33 (10.6%)					
Noninformative <sup>d</sup>	72 (9.0%)	56 (22.9%)	0 (0.0%)	5 (4.1%)	7 (10.8%)	2 (10.5%)	48 (14.2%)	4 (22.2%)	6 (4.5%)	15 (4.8%)					
No data <sup>e</sup>	114 (14.3%)	25 (10.2%)	38 (12.3%)	18 (14.9%)	16 (24.6%)	4 (21.1%)	51 (15.1%)	1 (5.6%)	9 (6.7%)	51 (16.5%)					

This table excludes US-born people and the 155 individuals with an unknown country of birth. CRF, circulating recombinant form; NHSS, National HIV Surveillance System; PHSKC, Public Health – Seattle & King County; PR/RT, protease/reverse transcriptase; PS, partner services; PWID, person who injects drugs.

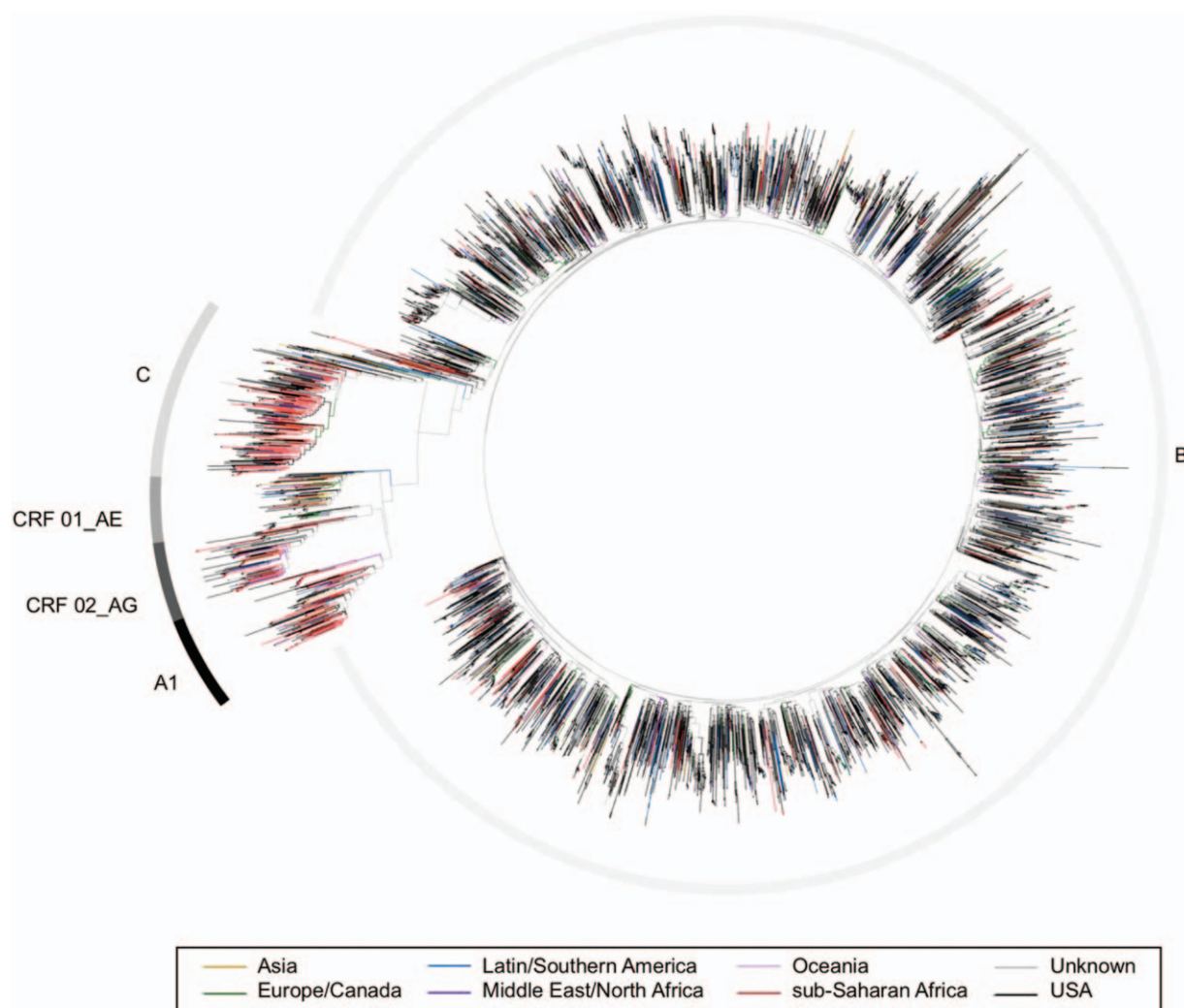
<sup>a</sup>Noninformative testing histories are among individuals for whom we are unable to determine location of HIV acquisition because their most recent negative HIV test was prior to their arrival in the United States.

<sup>b</sup>Per decision tree (Fig. 1): Individuals whose last reported negative test result was after date of arrival in the United States and who did not report any sex during travel abroad during partner services interviews. For individuals with uninformative HIV testing histories and who had an available HIV sequence, those who belonged to a genetic cluster.

<sup>c</sup>Per decision tree (Fig. 1): Individuals for whom PHSKC documented a prior diagnosis outside of the United States and individuals who either reported an HIV-positive test result after sex during travel abroad during partner services interviews. For individuals with uninformative HIV testing histories and who had an available HIV sequence, those who had a non-B subtype and did not cluster.

<sup>d</sup>Per decision tree (Fig. 1): Individuals who has a noninformative testing history or who had a B-subtype, didn't cluster and were born in Asia, Latin/South American, Europe/Canada, Oceania or the Middle East/North Africa because subtype B circulates more frequently in those regions.

<sup>e</sup>Individuals who neither completed a partner services interview nor had an available HIV sequence.



**Fig. 3. FastTree Phylogeny of All Sequences Sampled 2010–2018 in King County, Washington by region of birth and HIV subtype.**

2010 and 2018. However, most (51%) foreign-born HIV-infected residents likely acquired HIV outside of King County (primarily outside of the United States) and were misclassified by the NHSS as ‘new’ local infections and contribute to an overestimate of local HIV incidence. Traditional epidemiology methods – including case investigation and analysis of data collected through partner services interviews – identified the location of HIV acquisition for 45% of PLWH who were born outside of the United States and defined as a new diagnosis by the NHSS. Molecular epidemiology methods allowed us to infer the location of HIV acquisition for an additional 32% of foreign-born residents of King County newly diagnosed with HIV.

These findings suggest that there is some degree of local acquisition among people born outside of the United States, the majority of whom were MSM (69%) and a minority of whom were heterosexual (13%) or had unknown mode of transmission (16%). Location of HIV

acquisition also varied significantly by region of birth. Consistent with prior studies, sub-Saharan African residents newly diagnosed with HIV were mostly women, and few were MSM [2,5,7]. Although they were the least likely to cluster, our combined inference approach suggest that at least 7% of people born in sub-Saharan Africa acquired HIV locally, a proportion that is much lower than previous findings from King County, which estimated 21–26% of African-born people acquired HIV locally [6]. HIV among individuals born in other regions more closely reflected the epidemic in King County in that they were predominantly MSM infected with HIV subtype B.

Molecular epidemiology provided additional insight into HIV transmission patterns. Patterns of clustering and coclustering suggest that foreign-born individuals are not contributing substantially to either the heterosexual or MSM epidemics in King County. Few heterosexual foreign-born individuals belonged to a genetic cluster. In



addition, while the majority of foreign-born individuals who clustered were MSM, foreign-born MSM still had significantly lower odds of clustering with other MSM compared with people born in the United States. Thus, these data suggest that being born outside of the United States is associated with relatively lower local transmission rates. These dynamics have also been observed in European settings. A phylogenetic analysis from Belgium similarly found that sub-Saharan African immigrants had a limited contribution to ongoing HIV transmission, despite comprising a large proportion of the population of PLWH [21].

Our results are limited by the large proportion of foreign-born individuals who had not been tested for HIV since arriving in the United States or who had never tested for HIV prior to diagnosis; notably, over half of individuals reporting heterosexual or unknown transmission had never previously tested for HIV. This finding is consistent with a previous study of nation-wide data [22]. Another limitation is that genetic clustering methods rely on the choice of a distance threshold (for the current study we performed sensitivity analyses on these thresholds) and are unable to determine directionality of putative transmissions. Thus, our analyses are vulnerable to some degree of misclassification; while clustering is suggestive of an individual being linked to a local chain of transmission, this may be due to either acquisition or transmission. Thus, some individuals who acquired HIV abroad may be misclassified as local cases if they clustered, and others may not be identified as local cases because they did not cluster at the chosen distance threshold. However, at a population level, this is a reasonable proxy for the proportion of infections that are locally acquired. Lastly, clustering is also vulnerable to confounding by sampling coverage and time from infection to sample collection [18,19]. Our analysis is also limited to individuals residing in King County at the time of diagnosis due to jurisdictional boundaries that govern the management of partner services and surveillance data. Thus, we may not capture individuals who cluster with people in neighboring counties or elsewhere in the United States. Lastly, clustering analyses are not able to determine location of acquisition for marital partners who cluster and who immigrate to United States together.

In summary, the majority of PLWH in King County who were born outside of the United States likely acquired HIV abroad and efforts should be made to diagnose and link them to care as soon as possible after arrival to the United States. In addition, some foreign-born residents (e.g. MSM and persons born in Latin America) were more likely to have acquired HIV locally, suggesting these populations may not be reached by prevention efforts. Our findings also highlight how molecular epidemiology can complement traditional epidemiology and provide useful information to local public health jurisdictions beyond molecular cluster detection. This study demonstrates how molecular

epidemiology can fill gaps left by variations in the content and completeness of partner service interview and HIV surveillance data. Specifically, molecular analyses provided key information on the location of HIV acquisition that may provide more accurate estimates of local incidence and help tailor interventions.

## Acknowledgements

The current work was supported by NIAID NIH R01AI127232 to J.T.H. and R.P.K. This work was also supported by the University of Washington and Fred Hutch Center for AIDS Research, an NIH-funded program under award number AI027757 which is supported by the following NIH Institutes and Centers: NIAID, NCI, NIMH, NIDA, NICHD, NHLBI, NIA, NIGMS, NIDDK.

## Conflicts of interest

There are no conflicts of interest.

*Results previously presented at STI & HIV 2019 World Congress, Joint meetings of ISSTD and IUSTI, in Vancouver, CAN, on July 17, 2019.*

## References

1. National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Division of HIV/AIDS Prevention. *Detecting and responding to HIV transmission clusters: a guide for health department*. 2018, Available at: <https://www.cdc.gov/hiv/pdf/funding/announcements/ps18-1802/CDC-HIV-PS18-1802-AttachmentE-Detecting-Investigating-and-Responding-to-HIV-Transmission-Clusters.pdf>. [Accessed 9 October 2019].
2. Prosser AT, Tang T, Hall HI. **HIV in persons born outside the United States, 2007–2010**. *JAMA* 2012; **308**:601–607.
3. U.S. Census Bureau QuickFacts 2018. Available at: <https://www.census.gov/quickfacts/fact/table/US/PST045218>. [Accessed 20 September 2019].
4. Kerani R, Bennett AB, Golden M, Castillo J, Buskin SE. **Foreign-born individuals with HIV in King County, WA: a glimpse of the future of HIV?** *AIDS Behav* 2018; **22**:2181–2188.
5. Buskin S, Bennett A, Thibault C, Hood J, Kerani R, Golden M. **Increases in the proportion of HIV among foreign-born individuals in King County and their impact on HIV prevention**. *8th International AIDS Society Conference on HIV Pathology, Treatment and Prevention*. 2015.
6. Kerani R, Herbeck J, Buskin S, Dombrowski J, Bennett A, Barash E, et al. **Evidence of local HIV transmission in the African community of King County, Washington**. *8th International AIDS Society Conference on HIV Pathology, Treatment and Prevention*. 2015.
7. Valverde EE, Oster AM, Xu S, Wertheim JO, Hernandez AL. **HIV transmission dynamics among foreign-born persons in the United States**. *JAIDS J Acquir Immune Defic Syndr* 2017; **76**:445–452.
8. Wiewel EW, Torian LV, Hanna DB, Bocour A, Shepard CW. **Foreign-born persons diagnosed with HIV: where are they from and where were they infected?** *AIDS Behav* 2015; **19**:890–898.
9. Herbeck JT, Kerani RP, Poon AFY. **What proportion of HIV-infected foreign-born individuals in the United States have been infected after immigrating to the United States?** *JAIDS J Acquir Immune Defic Syndr* 2018; **77**:e35–e36.

10. Dennis AM, Hué S, Pasquale D, Napravnik S, Sebastian J, Miller WC, *et al.* **HIV transmission patterns among immigrant latinos illuminated by the integration of phylogenetic and migration data.** *AIDS Res Hum Retroviruses* 2015; **31**:973–980.
11. Katoh K, Standley DM. **MAFFT multiple sequence alignment software version 7: improvements in performance and usability.** *Mol Biol Evol* 2013; **30**:772–780.
12. Tamura K, Nei M. **Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees.** *Mol Biol Evol* 1993; **10**:512–526.
13. Wertheim JO, Kosakovsky Pond SL, Forgiione LA, Mehta SR, Murrell B, Shah S, *et al.* **Social and genetic networks of HIV-1 transmission in New York City.** *PLoS Pathog* 2017; **13**:e1006000.
14. Poon AFY. **Impacts and shortcomings of genetic clustering methods for infectious disease outbreaks.** *Virus Evol* 2016; **2**:vew031.
15. Struck D, Lawyer G, Ternes AM, Schmit JC, Bercoff DP. **CO-MET: adaptive context-based modeling for ultrafast HIV-1 subtype identification.** *Nucleic Acids Res* 2014; **42**:e144–e144.
16. Pineda-Peña AC, Faria NR, Imbrechts S, Libin P, Abecasis AB, Deforche K, *et al.* **Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: performance evaluation of the new REGA version 3 and seven other tools.** *Infect Genet Evol* 2013; **19**:337–348.
17. Price MN, Dehal PS, Arkin AP. **FastTree: computing large minimum evolution trees with profiles instead of a distance matrix.** *Mol Biol Evol* 2009; **26**:1641–1650.
18. Poon AF, Joy JB, Woods CK, Shurgold S, Colley G, Brumme CJ, *et al.* **The impact of clinical, demographic and risk factors on rates of HIV transmission: a population-based phylogenetic analysis in British Columbia, Canada.** *J Infect Dis* 2015; **211**:926–935.
19. Volz EM, Koopman JS, Ward MJ, Brown AL, Frost SDW. **Simple epidemiological dynamics explain phylogenetic clustering of HIV from patients with recent infection.** *PLoS Comput Biol* 2012; **8**:e1002552.
20. Hemelaar J, Elangovan R, Yun J, Dickson-Tetteh L, Fleminger I, Kirtley S, *et al.* **Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis.** *Lancet Infect Dis* 2019; **19**:143–155.
21. Verhofstede C, Dauwe K, Fransen K, Van Laethem K, Van den Wijngaert S, Ruelle J, *et al.* **Phylogenetic analysis of the Belgian HIV-1 epidemic reveals that local transmission is almost exclusively driven by men having sex with men despite presence of large African migrant communities.** *Infect Genet Evol* 2018; **61**:36–44.
22. Valverde E, DiNenno E, Oraka E, Bautista G, Chavez P. **HIV testing among foreign-born men and women in the United States: results from a nationally representative cross-sectional survey.** *J Immigr Minor Health* 2018; **20**:1118–1127.