




Article

# Phylodynamics Helps to Evaluate the Impact of an HIV Prevention Intervention

Tetyana I. Vasylyeva <sup>1,2,\*</sup> , Alexander Zarebski <sup>1</sup>, Pavlo Smyrnov <sup>3</sup>, Leslie D. Williams <sup>4</sup>, Ania Korobchuk <sup>3</sup>, Mariia Liulchuk <sup>5</sup> , Viktoriia Zadorozhna <sup>5</sup>, Georgios Nikolopoulos <sup>6</sup>, Dimitrios Paraskevis <sup>7</sup> , John Schneider <sup>8</sup>, Britt Skaathun <sup>9</sup>, Angelos Hatzakis <sup>7</sup>, Oliver G. Pybus <sup>1</sup> and Samuel R. Friedman <sup>10</sup>

<sup>1</sup> Department of Zoology, University of Oxford, OX1 3SY Oxford, UK; alexander.zarebski@zoo.ox.ac.uk (A.Z.); oliver.pybus@zoo.ox.ac.uk (O.G.P.)

<sup>2</sup> New College, University of Oxford, OX1 3BN Oxford, UK

<sup>3</sup> Alliance for Public Health, Kyiv 03150, Ukraine; smyrnov@aph.org.ua (P.S.); korobchuk@aph.org.ua (A.K.)

<sup>4</sup> Division of Community Health Sciences, University of Illinois at Chicago School of Public Health, Chicago, IL 60612, USA; lwilliams@ndri.org

<sup>5</sup> State Institution “The L.V. Gromashevsky Institute of Epidemiology and Infectious Diseases of NAMS of Ukraine”, Kyiv 03038, Ukraine; mliulchuk@gmail.com (M.L.); viz2010@ukr.net (V.Z.)

<sup>6</sup> Medical School, University of Cyprus, Nicosia 1678, Cyprus; gknikolopoulos@gmail.com

<sup>7</sup> Department of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, 157 72 Athens, Greece; dparask@med.uoa.gr (D.P.); ahatzak@med.uoa.gr (A.H.)

<sup>8</sup> Department of Medicine, University of Chicago, Chicago, IL 60637, USA; jschnei1@medicine.bsd.uchicago.edu

<sup>9</sup> Department of Medicine, University of California San Diego, San Diego, CA 92093, USA; bskaathun@health.ucsd.edu

<sup>10</sup> Department of Population Health, New York University, New York, NY 10003, USA; samuel.friedman@nyulangone.org

\* Correspondence: Correspondence: tetyana.vasylyeva@zoo.ox.ac.uk

Received: 30 January 2020; Accepted: 15 April 2020; Published: 20 April 2020



**Abstract:** Assessment of the long-term population-level effects of HIV interventions is an ongoing public health challenge. Following the implementation of a Transmission Reduction Intervention Project (TRIP) in Odessa, Ukraine, in 2013–2016, we obtained HIV *pol* gene sequences and used phylogenetics to identify HIV transmission clusters. We further applied the birth-death skyline model to the sequences from Odessa ( $n = 275$ ) and Kyiv ( $n = 92$ ) in order to estimate changes in the epidemic's effective reproductive number ( $R_e$ ) and rate of becoming uninfected ( $\delta$ ). We identified 12 transmission clusters in Odessa; phylogenetic clustering was correlated with younger age and higher average viral load at the time of sampling. Estimated  $R_e$  were similar in Odessa and Kyiv before the initiation of TRIP;  $R_e$  started to decline in 2013 and is now below  $R_e = 1$  in Odessa ( $R_e = 0.4$ , 95%HPD 0.06–0.75), but not in Kyiv ( $R_e = 2.3$ , 95%HPD 0.2–5.4). Similarly, estimates of  $\delta$  increased in Odessa after the initiation of TRIP. Given that both cities shared the same HIV prevention programs in 2013–2019, apart from TRIP, the observed changes in transmission parameters are likely attributable to the TRIP intervention. We propose that molecular epidemiology analysis can be used as a post-intervention effectiveness assessment tool.

**Keywords:** HIV; phylodynamics; birth-death model; prevention; intervention

## 1. Introduction

In recent years phylodynamic analyses have become increasingly widely applied in the field of HIV surveillance and prevention. Phylodynamics can estimate changes in viral population size over time and be used to explore how various factors (epidemiological, evolutionary, behavioural, etc.) affect these changes [1]. Such analyses are often performed on sets of HIV genetic sequences collected through national or regional HIV drug resistance surveillance efforts. They can be used for identifying and monitoring of HIV transmission clusters, identification of local HIV outbreaks and for targeting subsequent public health interventions [2–4]. When the location from which the HIV genomes were obtained is known, phylogeographic methods can be used to describe HIV migration within various geographic regions [4–6], or among different epidemiological risk groups.

Phylodynamic models based on birth-death processes enable the estimation of three important epidemiological parameters: (i) the effective reproductive number,  $R_e$ , which is the average number of secondary infections arising from an infected individual at any given time during an epidemic, (ii) the proportion of the population that has been sampled and (iii) the rate of becoming uninfected,  $\delta$ , which is the rate with which individuals cease transmitting infections (i.e., due to behavioural change, treatment or death) [7–9]. The effective reproductive number is often used to describe transmission dynamics in a population;  $R_e > 1$  means the epidemic is growing, while  $R_e < 1$  means the epidemic is declining. The becoming-uninfected rate is inversely related to the average duration of the infectious period, assuming that length of this period is exponentially distributed. The becoming-uninfected rate is important, because it is known that Treatment as Prevention (TasP) is most effective if people are diagnosed and treated as soon as possible after they are infected [10]. Thus, increasing the rate at which individuals cease to be infectious can reduce onward transmission and is clinically beneficial for the infected individual.

Over the last 10 years the scope of HIV prevention interventions has grown to incorporate new clinical insights and technological advances; it is now common for interventions to involve the collection of virus genetic sequences. Often these studies involve marginalised groups, such as people who inject drugs (PWID), men who have sex with men (MSM), female sex workers and undocumented migrants, as usually HIV prevalence is higher in these groups than in so-called “general population”. Importantly, it is often unrealistic to try to obtain a random sample from certain populations due to their hidden (and in some locations criminalized) nature, since a random sampling assumes that each individual is selected by chance and everyone has an equal chance to participate in the study. Social-network based approaches have proved to be an effective alternative to recruit participants for HIV research studies and to deliver HIV preventive measures in hard-to-reach populations [11]. The Transmission Reduction Intervention Project (TRIP) is one example of such a network-based intervention that was implemented in various settings and transmission groups from 2011 to 2016 [12,13]. TRIP aimed to evaluate a network-based approach to locating recently (within the last 6 months) HIV-infected people and to preventing ongoing transmission. Previously published results have shown that the TRIP intervention was more effective in finding recently HIV-infected people, those with high HIV viral load and those previously undiagnosed, than the standard of care options [13–15].

Odessa, Ukraine, was one of the first locations where TRIP was implemented. Ukraine has a large HIV-infected population (estimated to be 240,000 people as of 2019 [16]), but only 4 (out of 27 total) administrative regions account for >50% of HIV cases in the country: Odessa, Kyiv, Dnipro and Donetsk. Odessa is thought to be the “birthplace” of the HIV epidemic in Ukraine and the post-soviet territories [17], and it still has one of the highest HIV prevalence levels in various risk groups in Ukraine (i.e., 13% in MSM and 28% in PWID in 2015) [18,19]. Overall, Kyiv and Odessa share similar HIV epidemiological profiles and have experienced a similar number of HIV prevention interventions, other than TRIP. They were the first cities in which treatment started being offered to patients upon diagnosis in 2014, even though it is still not a universal practice in Ukraine. Since 2013, in both cities, a case management approach to linkage to care has been introduced for PWID: peer outreach workers help newly diagnosed HIV-positive PWID to register with local AIDS Centres and start antiretroviral

treatment. During 2012–2013 the “Project Protect” intervention was implemented in both cities, which offered the PWID community education about the danger of acute HIV infection [20]. According to our previous estimates, Odessa and Kyiv were also some of the main sink (receiving infections) locations for HIV spatial dissemination associated with the recent military conflict in the east of Ukraine [21]. Neither city displayed a decline in the annual number of newly HIV diagnosed people in 2016–2019, while the annual number of performed HIV tests also grew [22].

It has been proposed that the analysis of HIV gene sequences can be used to assess the impact of a specific preventive intervention in a specific location [23]. Here we aimed to assess the population-level impact of TRIP on HIV transmission in Odessa, Ukraine, and to compare it to the estimated HIV transmission dynamics in the capital of Ukraine, Kyiv. We used a data set of HIV sequences from Odessa and Kyiv, Ukraine from 2000–2019, compiled from several sources. We then describe the application of phylodynamic methods to estimate the effective reproductive number and the becoming-uninfectious rate from these data, along with an analysis of HIV transmission clusters. We find that following the TRIP implementation in Odessa, the HIV epidemic there is characterized by a lower  $R_e$  and a faster  $\delta$  compared to Kyiv. Finally, we discuss the limitations and the potential of phylodynamic analysis to evaluate HIV prevention efforts.

## 2. Materials and Methods

### 2.1. Data

Our data set comprises HIV-1 subtype A *pol* genetic sequences (corresponding to HXB2 positions 2568–3256) compiled from three sources: (i) the TRIP intervention, (ii) AIDS Centres in Odessa and Kyiv and (iii) the Los Alamos National Laboratory HIV sequence database, LANL (<https://www.hiv.lanl.gov>). This includes 161 newly generated sequences from the TRIP intervention conducted in Odessa in 2013–2016. In TRIP, people were screened for HIV and, if positive, further tested to determine whether the infection was recent (defined by a Limiting Antigen avidity test or by a negative HIV test result within the previous 6 months). Those with recent HIV infection were invited to participate in TRIP and were given coupons to invite their sexual and injecting partners into the intervention. Coupons were also given to their partners, but then the recruitment stopped unless another recently HIV infected person was found in their network. Thus, within the TRIP data, we differentiated between samples that (a) came from people screened for recent infection who were not recently infected, and thus were not invited to further participate in the network intervention component of TRIP (Non-participant TRIP samples, NPTS) and samples which (b) came from TRIP participants—either recently infected people or their 1st or 2nd degree network contacts (participant TRIP samples, PTS). Further details on the TRIP data collection and genetic sequencing procedures can be found in the original publications [13,24]. We obtained 197 sequences of the subtype A *pol* gene (84 newly generated) from patients of AIDS Centres (HIV treatment clinics) in Odessa ( $n = 105$ ) and Kyiv ( $n = 84$ ) collected in 2012–2019. These samples were sequenced at the L.V.Gromashevskij Institute of Epidemiology and Infectious Diseases in Kyiv using ViroSeq® HIV-1 Genotyping System v2.0 Kit (Celera Corporation, Abbott Laboratories, Abbott Park, IL, USA) [25]. Finally, we downloaded from the LANL database all publicly available HIV-1 subtype A *pol* sequences that (i) corresponded to the same genomic region, (ii) were sampled from Odessa ( $n = 25$ ) and Kyiv ( $n = 12$ ) and had available sampling year and city information. See the Table S1 for accession numbers of the sequences used in the analysis. Accession numbers for TRIP sequences are MT348782–MT348931.

We further grouped the sequences into the following datasets: *Odessa* and *Kyiv* datasets comprised sequences from Odessa and Kyiv from LANL and AIDS Centres; sequences from TRIP were added to the *Odessa* dataset. Duplicates in both datasets were removed using the ElimDupes tool from the LANL website (<https://www.hiv.lanl.gov/>). Additionally, most PTS sequences had information about the participant’s transmission risk group, HIV viral load, HIV recency status (as defined by the TRIP

intervention protocol) and information on co-infection with hepatitis C virus (HCV), hepatitis B virus (HBV) and syphilis (co-infections determined by a rapid multi-test).

All TRIP participants gave informed consent under protocols approved by the IRBs of the National Development and Research Institutes, USA, and the medical ethics committee at L.V. Gromashevsky Institute of Epidemiology and Infectious Diseases, Ukraine.

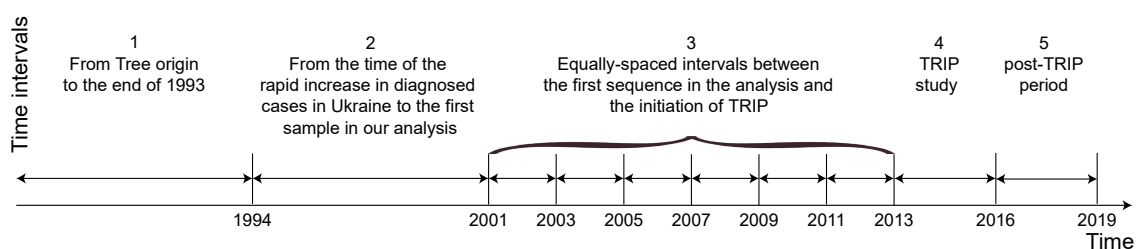
## 2.2. Phylogenetic Analysis

We used the REGAv3 [26] and COMET [27] subtyping tools to confirm that the sequences' subtypes were correctly labelled. We estimated maximum likelihood (ML) phylogenetic trees for both datasets using RAxML [28]. In all of the ML analyses we used an HKY nucleotide substitution model with gamma-distributed rate variation among sites and ran a bootstrap analysis with 100 replicates. We identified potential transmission clusters in the Odessa phylogenetic tree with ClusterPicker [29]. Any clade with two or more sequences, within-cluster genetic distance <1.5%, and bootstrap statistical support >90% was defined as a possible transmission cluster, in accordance with the literature [30]. To study differences between those individuals who were placed in clusters and those who were not, we used t-tests (for continuous variables) and Fisher's exact test (for categorical variables) to compare socio-demographic characteristics.

## 2.3. Phylodynamics and Estimation of Epidemiological Parameters

We first reconstructed time-calibrated phylogenetic trees from the Odessa dataset and estimated the time to the most recent common ancestors (TMRCA) of all identified transmission clusters using BEASTv1.10.4 [31]. First, we used path-sampling [32] and stepping stone [33] marginal likelihood estimators (MLE) to select the most appropriate coalescent model. The non-parametric Bayesian skyline plot (BSP) model [34] was selected over the constant population size and exponential growth models (see Table S2) and was therefore used in further analyses. For the subsequent BSP analyses, we used an HKY nucleotide substitution model with gamma-distributed rate variation among sites (HKY+G). Given the poor molecular clock signal in our dataset (as assessed using TempEst [35]), we relied on previously published estimates of the evolutionary rate to calibrate the molecular clock. Specifically, we used a lognormal relaxed molecular clock model with a continuous-time Markov chain (CTMC) prior on the mean (constrained to values between  $1 \times 10^{-3}$  –  $3 \times 10^{-3}$ ) and a normal prior on the standard deviation (mean =  $5 \times 10^{-4}$ , standard deviation =  $5 \times 10^{-4}$ ), as previously reported in [36,37]. We also used a normal distribution prior for the TMRCA of the tree (mean = 35, standard deviation = 5). We used the same nucleotide substitution and molecular clock models for all subsequent phylodynamic analyses. Markov chain Monte Carlo (MCMC) sampling was performed for  $300 \times 10^6$  iterations (with 20% removed as burn-in). We used Tracer [38] to visually inspect the convergence of the MCMC runs for all BEAST analyses and to ensure a minimum effective samples size of 200, as is standard practice [4,39].

We further applied a birth-death skyline model (BDSKY) [8] implemented in BEASTv2.6.0 [40] to estimate  $R_e$  and the becoming uninfected rate ( $\delta$ ) for *Odessa* and *Kyiv* independently. The BDSKY model assumes that  $R_e$  and  $\delta$  are piecewise constant functions through time; for this analysis we specified 10 time-intervals (Figure 1) that correspond to:



**Figure 1.** Time intervals specified for the birth-death skyline model (BDSKY) analysis.

(1) the time between the origin of the tree and 1993 (N of sequences = 0 in both Odessa and Kyiv);  
 (2) the time between 1994, which corresponds to the year when the number of newly diagnosed HIV cases dramatically increased in Ukraine [41,42]), and the first sequence in the analysis (N of sequences = 0 in both Odessa and Kyiv);

(3–8) multiple intervals of equal duration between the collection date of the first sequence in the analysis (2001 for Odessa, and 2000 for Kyiv) and the initiation of TRIP in 2013 (N of sequences = 22, 0, 0, 0, 3, 1 in Odessa and  $n = 2, 0, 1, 0, 9, 3$  in Kyiv in 2001–2002, 2003–2004, 2005–2006, 2007–2008, 2009–2010, 2011–2012, correspondingly);

(9) the time period during which the TRIP intervention has been implemented, corresponding to the date of sampling of the first and the last TRIP sequence (June 2013–February 2016) (N of sequences = 205 in Odessa and  $n = 54$  in Kyiv);

(10) the time after the TRIP intervention ended, corresponding to the time interval between the sampling date of the youngest TRIP sequence and the sampling date of the youngest sequence in the dataset (February 2016–February 2019) (N of sequences = 44 in Odessa and  $n = 23$  in Kyiv).

The BDSKY model assumes individuals become non-infectious upon sampling (i.e., they cannot transmit HIV further) as they are provided treatment and reach viral suppression. The prior distributions for the BDSKY model consisted of: (a) a uniform distribution for the date of origin, which is the duration of the epidemic represented by the sequences in the analysis, with an upper bound of 60 and lower bound 19 (the age of the oldest sequence in the analysis); (b) a lognormal distribution for  $R_e$  (mean = 1.0, standard deviation = 1.25) with an upper bound of 10; (c) a lognormal distribution for  $\delta$  (mean = 0.0, standard deviation = 1.0) constrained to take values between 0.05 - 6, corresponding to the plausible values of the duration of infectious period (20 years–2 months); and (d) a beta prior for the sampling proportion ( $\alpha = 1.0$ ,  $\beta = 10.0$ , constrained to values between 0 and 0.01) for time intervals (3)–(10) as defined above; the sampling proportion was set to 0 before the first sampling date (time intervals (1) and (2) as defined above). We ran 4 independent MCMC chains for  $300 \times 10^6$  generations (with the initial 3% removed as burn-in) for the *Odessa* dataset. For the analysis of the *Kyiv* dataset we used the same analyses specifications and ran 2 independent MCMC analyses for  $300 \times 10^6$  generations (10% burn-in). We also used the same BEAST specifications in our consequent sensitivity analyses. We used LogCombiner [40] to combine results from multiple chains.

XML files specifying all the analyses and data described above are available at the GitHub repository [https://github.com/HIVMolEpi/TRIP\\_Ukraine](https://github.com/HIVMolEpi/TRIP_Ukraine).

### 3. Results

#### 3.1. Data and Phylogenetic Clustering

Between 2013–2016, 1252 participants were screened for recent HIV infection as part of the TRIP project in Odessa, Ukraine. For 67 non-recently infected participants screened at TRIP an HIV sequence was generated (NPTS sequences). Forty individuals were identified as recently infected according to the TRIP definition of “recency” [13], and a further 369 HIV-positive individuals were recruited from their networks [15], these individuals contributed an additional 94 HIV *pol* sequences (the PTS sequences). After removing duplicates and non-subtype A sequences, 65 NPTS, 85 PTS, 25 LANL and 100 sequences available from the Odessa AIDS Centre were included in the *Odessa* dataset, giving a total of 275 sequences collected from 2001–2019. Table 1 reports the socio-demographic and clinical characteristics of the individuals from whom NPTS and PTS sequences were obtained ( $n = 150$ ).

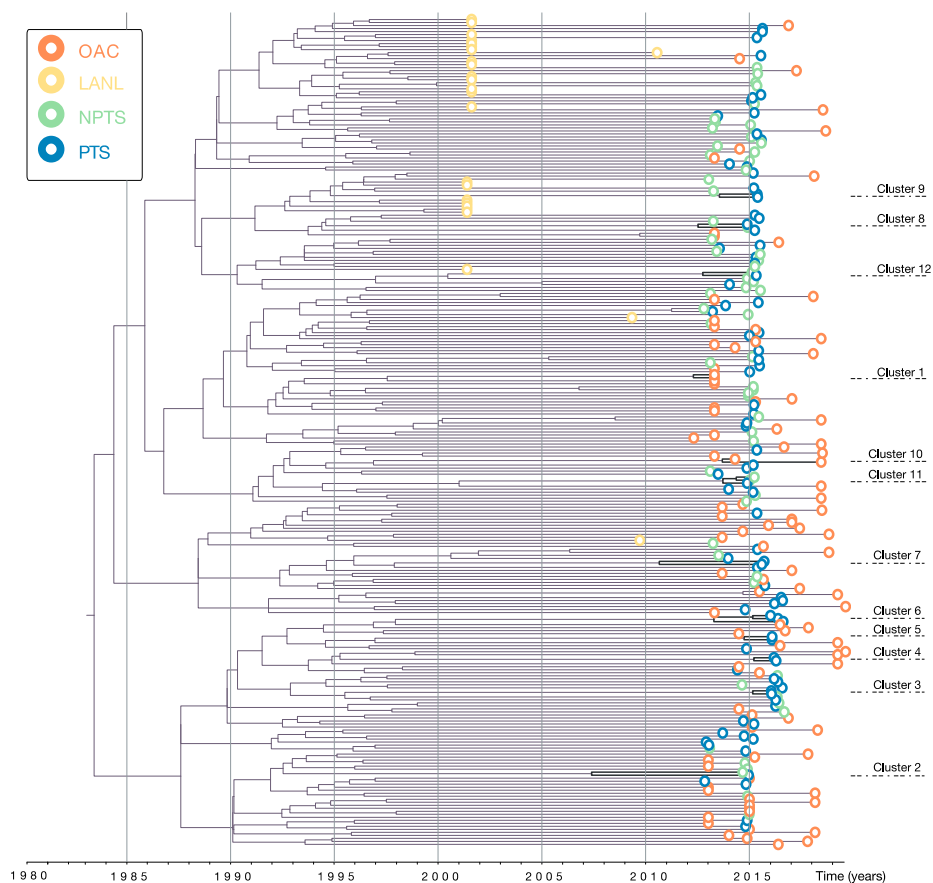
Based on the ML phylogenetic tree reconstructed from the *Odessa* dataset (See Figure S1), we identified 12 probable transmission clusters: 10 pairs and 2 clusters of three sequences. These clusters also had high posterior probability (>0.9) support on the maximum clade credibility tree reconstructed in the BSP analysis (Figure 2).



**Table 1.** Characteristics associated with the HIV sequences collected within the Transmission Reduction Intervention Project (TRIP) intervention.

Variables	Total (n = 150)	Clustered (n = 22)	Non-Clustered (n = 128)	Fisher's Exact Test/t-Test p-Value ***
Age (mean)	35.1	31.6	36.9	<b>0.05</b>
Women (N, % *)	46 (33.1)	6 (27.3)	40 (34.5)	0.63
PWID (N, %)	52 (40.6)	12 (54.5)	40 (37.7)	0.16
Years since 1st injection (for PWID, mean)	13.3	11.5	15.12	0.28
Recently infected (N, %)	28 (18.7)	6 (27.3)	22 (17.3)	0.37
HIV Viral load (mean)	142,740.6	245,157.6	113,935.8	<b>0.04</b>
HCV co-infection (N, %)	61 (53.5)	14 (63.6)	47 (51.1)	0.35
HBV co-infection (N, %)	5 (3.9)	0	5 (4.9)	0.59
Syphilis co-infection (N, %)	15 (11.6)	3 (13.6)	12 (11.2)	0.72
Index participants ** (N, %)	84 (56.4)	9 (40.9)	75 (59.1)	0.16

\* % excludes missing values. \*\* Participants NOT recruited through coupons. \*\*\* values marked in bold reached statistical significance at 0.05 level. Fisher's exact test was used to compare groups within variables.

**Figure 2.** Molecular clock tree reconstructed from sequences in *Odessa* dataset. The thick black lines indicate the identified probable transmission clusters. OAC – sequences from the Odessa AIDS Centre patients, LANL – sequences from the Los Alamos National Laboratory HIV sequence database, NPTS - Non-participant TRIP samples, PTS - participant TRIP sample.

Clusters 1 and 10 consisted of sequences isolated from AIDS Centre patients, thus no additional information was available for these sequences (e.g., age and HCV status). The sequences in the remaining ten clusters came from TRIP. The individuals from whom these sequences were obtained were younger on average than their unclustered counterparts (t-test  $p$ -value = 0.05,  $H_0$  = both groups have the same average age) and had approximately twice their viral load ( $p$ -value = 0.04,  $H_0$  = both groups have the same average viral load) (Table 1). However, these associations were not statistically significant after adjustment for multiple comparisons using the false discovery rate method (adjusted t-test  $p$ -value = 0.25 for age and  $p$ -value = 0.25 for viral load) [43].

Most of the sequences in the clusters came from individuals that (a) share a route of transmission (apart from clusters 2, 7 and 9); (b) were HCV-positive; (c) were HBV negative; (d) did not have syphilis (apart from clusters 6 and 11); and (e) were male (Table 2). Half of the clusters included at least one recently infected individual. Two clusters, 6 and 11, included individuals co-infected with syphilis, suggesting a possible sexual transmission route.

**Table 2.** Potential transmission clusters identified in Odessa.

Cluster	N	Year of Cluster Origin		Transm. Route	HCV	HBV	Syph-ilis	Includes Recently Infected Person	Gender	TRIP Recruitment Group
		Year	95% HPD							
1	2	2012.7	2010.8–2014.0	Unk	Unk	Unk	Unk	Unk	Unk	N/A
2	2	2008.0	2003.1–2012.9	Drug use/sexual	Pos	Neg	Neg	Yes	Male	PTS/NPTS
3	2	2014.5	2013.0–2015.6	Drug use	Pos	Neg	Neg	Yes	Male	PTS
4	2	2014.5	2012.9–2015.7	Drug use	Neg	Neg	Neg	No	Male	PTS
5	2	2014.1	2012.2–2015.5	Drug use	Pos	Neg	Neg	No	Male	PTS
6	3	2012.5	2009.7–2014.9	Drug use	Pos	Neg	Neg/Pos	No	Male/Female	PTS
7	2	2010.7	2007.0–2014.0	Drug use/sexual	Neg	Neg	Neg	Yes	Male/Female	PTS
8	2	2012.9	2010.3–2015.1	Sexual	Neg	Neg	Neg	Yes	Male/Female	PTS
9	2	2013.9	2011.4–2016.0	Drug use/sexual	Pos	Neg	Neg	No	Male	PTS
10	2	2014.1	2012.4–2015.0	Unk	Unk	Unk	Unk	Unk	Unk	N/A
11	3	2014.2	2012.5–2015.5	Sexual	Pos	Neg	Neg/Pos	Yes	Female	PTS/NPTS
12	2	2013.1	2010.3–2015.3	Sexual	Neg	Neg	Neg	No	Male	PTS/NPTS

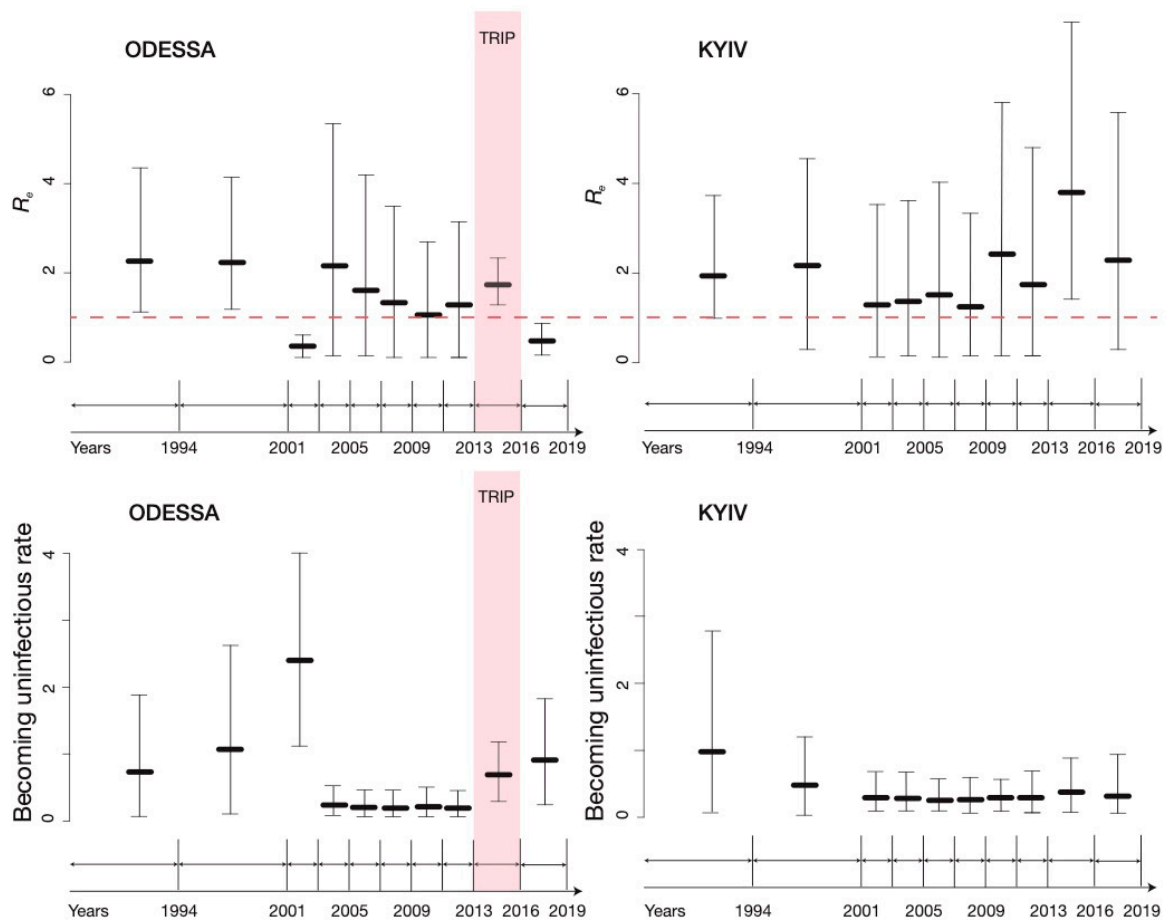
Unk—unknown, PTS—participant TRIP samples, NPTS—non-participants TRIP samples.

After removing duplicates and non-subtype A sequences, the *Kyiv* dataset comprised 92 sequences, those sampled at the AIDS Centres and available through the LANL database. The ML phylogenetic tree reconstructed from *Kyiv* sequences is presented in Figure S2. Using our definition of phylogenetic clusters, no clusters were found on the tree reconstructed from the *Kyiv* dataset.

### 3.2. BDSKY Analysis

The BDSKY analysis of the *Odessa* dataset showed that from the beginning of the epidemic and until early 2016 the  $R_e$  in Odessa remained consistently above 1 with the exception of 2001–2002 when it dropped to  $R_e = 0.27$  (95% Highest Posterior Density (HPD) 0.04–0.52, Figure 3). With the exception of 2001–2002, there was little difference in  $R_e$  in Odessa and *Kyiv* before the TRIP intervention: in 2011–2012 the  $R_e$  was estimated to be 1.19 (95% HPD 0.03–3.07) in Odessa and 1.66 (95% HPD 0.04–4.7) in *Kyiv* (Figure 3). In 2013–early 2016, during the implementation of TRIP  $R_e$  was estimated to be 1.66 (95% HPD 1.19–2.2) in Odessa and  $R_e = 3.7$  (95% HPD 1.32–7.5) in *Kyiv*. After TRIP, in 2016–2019, it declined to below the epidemiological threshold of 1 necessary to maintain growth of the epidemic

in Odessa ( $R_e = 0.4$ , 95% HPD 0.06–0.75). During the same time  $R_e$  has likely remained high in Kyiv, estimated to be 2.32 (95% HPD 0.2–5.4).



**Figure 3.** Estimates of the effective reproductive number,  $R_e$ , in Odessa and Kyiv (top), where the red dotted line represents the epidemiological threshold of  $R_e = 1$ ; and the becoming uninfected rate (bottom) obtained from the *Odessa* and *Kyiv* datasets by means of the BDSKY analysis.

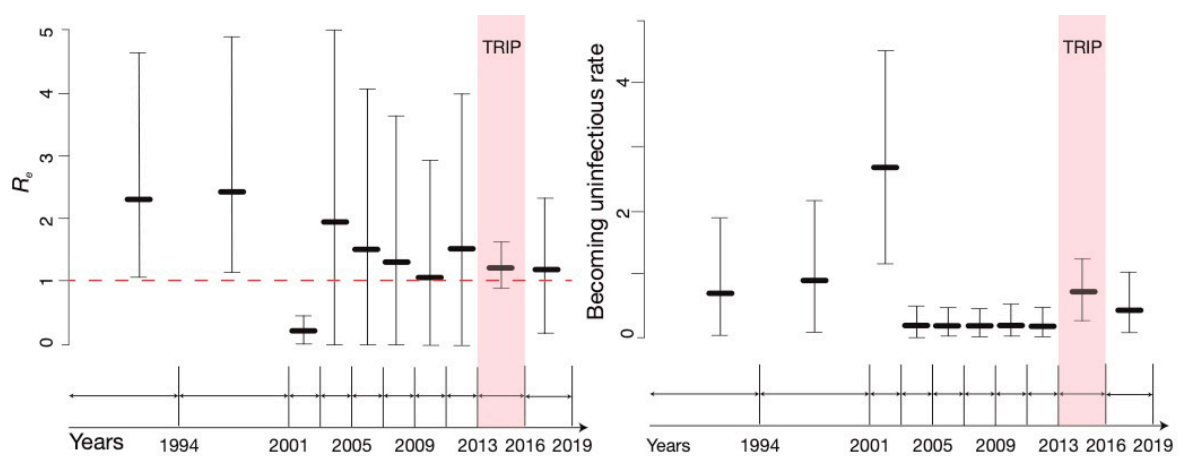
The becoming-uninfected rate for Odessa grew between the mid-1980s and 2000 from 0.7 before 1994 (corresponding to an infectious period of 1.4 years) to 2.65 in 2001–2002 (corresponding to infectious period of 5 months (Figure 3). This rate then dropped substantially to 0.2 in 2003–2004 (corresponding to an infectious period of 5 years) and remained low until the beginning of the TRIP project in 2013. In 2013–early 2016 the becoming-uninfected rate was estimated to be 0.72 (corresponding to an average infectious period of 1.4 years) and it continued to grow after the end of TRIP to be 0.94 (corresponding to an average infectious period of 1.1 years) in 2016–2019. Analysis of the *Kyiv* dataset revealed that the becoming-uninfected rate has gradually decreased over time, from 0.97 (corresponding to an average infectious period of 1 year) before 1994 to 0.33 (an average infectious period of 3 years) in 2016–2019 (Figure 3. Importantly, no difference in the becoming-uninfected rate between Odessa and Kyiv was observed for 10 years before the TRIP intervention (2003–2013), when the average duration of infectious period in Odessa started to decline.

After the initial epidemic growth period in late-90s/early 2000s, the BDSKY did not provide precise estimates for  $R_e$  or detect any changes in the becoming uninfected rate for the period when no or very few sampling events occurred (2002–2012). For the duration of this period, both locations had an  $R_e$  estimate just above 1 with confidence intervals from 0 to 4. Similarly, in both Odessa and Kyiv the becoming uninfected rate stabilized at around 0.2 between 2003–2013, corresponding to an average infectious period of 5 years.



### 3.3. Sensitivity

We performed several sensitivity analyses to explore the robustness of our estimates. First, we used a less strict definition of clusters for ClusterPicker analysis (bootstrap >80%, using the same tree and number of bootstrap replicates as for the analysis with higher bootstrap cut-off value), but this resulted in only one additional cluster. Second, we removed 65 PTS sequences generated by network-based recruitment in TRIP from the *Odessa* dataset, keeping the 65 NPTS sequences and 20 PTS sequences obtained from recently infected individuals at screening (thus, not obtained through network recruitment) in the analysis. This resulted in a dataset of  $n = 210$ . The ML phylogenetic tree reconstructed from this dataset can be found in Figure S3. After the network-derived sequences were eliminated, only 4 phylogenetic clusters were left on the *Odessa* phylogenetic tree, which presented a more similar figure to that in *Kyiv*, where no transmission clusters were identified. We then reran the BDSKY analysis with this dataset to make sure that no bias was introduced by these potentially-correlated samples (Figure 4).



**Figure 4.** Temporal estimates of the effective reproductive number,  $R_e$ , and the becoming uninfected rate, obtained from *Odessa* datasets after removal of all sequences generated by network-based recruitment of TRIP participants. The red dotted line represents the epidemiological threshold of  $R_e = 1$ .

This analysis yielded a lower mean estimate of  $R_e$ , 1.23 (95% HPD 0.9–1.64), during the implementation of the TRIP in *Odessa* in 2013–early 2016, and a larger mean estimate,  $R_e = 1.2$  (95% HPD 0.2–2.33), for the period after the TRIP intervention, in 2016–2019. Importantly, the  $R_e$  estimates in *Odessa* since the initiation of TRIP in 2013 remained substantially lower than the corresponding estimates for *Kyiv*, though the difference between the two estimates was no longer significant in 2016–2019. For the becoming uninfected rate in this analysis, the estimate was the same for the duration of TRIP as when the full dataset was analysed. For the years after the intervention (2016–2019), the becoming uninfected rate was lower, at 0.4, corresponding to an average infectious period of 2.4 years for *Odessa*, which was only a little higher than our estimates in *Kyiv* of 0.33, corresponding to an average infectious period of 3 years.

Finally, to make sure that our findings cannot be simply explained by a larger number of sequences in *Odessa* dataset compared to *Kyiv* dataset ( $n = 275$  vs.  $n = 92$ ), we have rerun the BDSKY analysis with 92 randomly selected sequences from the *Odessa* dataset. The mean estimates and 95% HPD obtained from this analysis did not substantially differ from the estimates obtained in the analysis of the full dataset (Figure S4).

## 4. Discussion

In the last five years, the amount of virus genetic data available for phylogenetic and phylodynamic analysis has substantially increased; in *Ukraine* alone, the number of HIV genetic sequences has

quadrupled since 2018. This presents new opportunities to apply phylogenetic and phylodynamic techniques to describe changes in virus populations and in HIV transmission dynamics in the country. Here, we examined factors associated with HIV phylogenetic clustering and applied a birth-death phylodynamic model to examine potential differences in HIV transmission dynamics between two cities (Odessa and Kyiv), following the introduction in one city (Odessa) of an HIV prevention intervention (TRIP). We found a reduction in the HIV effective reproduction number (i.e., reduced epidemic growth) in Odessa relative to Kyiv, where no TRIP intervention was enacted.

We have identified and described HIV genetic clustering patterns, as these are known to be informative for public health planning [44]. Interestingly, we did not observe many of the clustering patterns reported in other settings. Specifically, in our study, clustering status was not associated with gender, risk group or recent HIV infection. The latter is particularly surprising given that network contacts of recently-infected people were actively recruited in TRIP, and given that multiple previous reports have found that recently-infected individuals are more likely to be included in genetic clusters [45–47]. Moreover, modelling studies have showed that such clustering can be expected and is explained by intensive transmission during the early stages of HIV infection [48]. Absence of such an association in our analysis can be explained potentially by the low number of observed clusters. Clustering was associated with higher viral load in our sample (though this association lost significance after multiple test correction), which is similar to the findings from the TRIP intervention undertaken in Athens and can be interpreted as a marker of recent infection [24], but could also be explained by other factors (e.g., co-infections, late-stage HIV infection). Being in a phylogenetic cluster in our study was correlated with younger age, which might be a proxy for more recent transmission events; this is similar to the result of a previous report from Belgium [49], but unlike the results reported for some other European countries and Canada [50]. Unfortunately, we were not able to differentiate between heterosexual and MSM transmission in our analysis due to the way the data were collected and all sexually-acquired infections were analysed together. This prevents us from making direct comparisons with other studies that showed that MSM infections are more likely to cluster compared to infections acquired through other transmission routes [46,49].

The additional meta-data associated with the sequences in our study helped us to question some of the more obvious clustering patterns. For example, a co-infection with syphilis might suggest a possible sexual route for HIV transmission, or if clustered sequences came from PWID then a parenteral route of transmission might be suspected, as in cluster 6 in our analysis. In some cases, combining information about recent infection and estimates of TMRCA allowed us to rule out the potential for direct transmission between individuals whose sequences were clustered. For instance, a recently-infected individual in cluster 2 was recruited in 2016, which means that individual could not have been infected earlier than 2015, while the TMRCA of cluster 2 is much older (2008.0, 95% HPD 2003.1–2012.9) precluding direct transmission within this cluster. Similarly, in cluster 7, a recently infected individual was recruited in 2016 while the TMRCA of the cluster was estimated to be 2010.7 (95% HPD 2007.0–2014.0).

Our analysis of changes in transmission dynamics in Odessa following the TRIP intervention suggested that epidemic growth rate was reduced relative to that in Kyiv during the same period and that the beginning of the TRIP intervention in Odessa coincides with the slowing down of HIV transmission in that city. We observe that estimates of the reproductive number decreased and estimates of the rate with which people become diagnosed (and within the assumption of our model, treated) increase. Crucially, no such trend was observed in Kyiv during the same time. This difference might be attributed to the fact that TRIP incorporated active network recruitment as part of its procedures and was able to reach to comparatively high number of people with undiagnosed HIV infection [14]. Thus, large networks of people were screened and linked to care in Odessa, while also raising awareness in the community. This could have led to the changes in  $R_e$  and the becoming-uninfectious rate, both during and after the implementation of TRIP.

Our estimates of  $R_e$  both in Odessa and Kyiv prior to the initiation of TRIP are different from previously reported estimates [42,51]. For the mid-90s, our BDSKY estimates ( $R_e = 2.08$ , 95%HPD 0.21–4.41, for Kyiv and  $R_e = 2.17$ , 95%HPD 1.11–4.07, for Odessa) are much lower than the estimates of the basic reproductive number that we have previously obtained through a coalescent-based analysis of HIV-1 subtype A *env* sequences from Russia and Ukraine ( $R_e = 7$ ) [42]. Even higher  $R_e$  estimates were obtained through other birth-death models applied to the same *env* sequences (sampled in 1996–2011) by Magee et al. [51]. For the mid-2000s our estimates are relatively high, though this is accompanied by a large degree of uncertainty, i.e., large HPD intervals, which means that the true value might be included in the broad confidence intervals and is not very different from the previous reports that suggest that  $R_e$  in Ukraine was below 1 after 2000 [51]. However, we believe that these discrepancies are largely due to the different number and sampling dates of sequences in the analysis. As suggested by Magee et al. [51], when no birth events (i.e., “infection events”) were observed in the tree during the 2000s, a very low  $R_e$  was inferred. Similarly, in our analysis the  $R_e$  in Odessa was estimated to be very low in 2001–2002, probably attributed to the absence of sampled sequences and birth events in the years after (2003–2009). In our analysis, the majority of sequences were sampled after 2011 and multiple birth events were observed in the tree during mid-2000s, which resulted in higher  $R_e$  estimates after 2002. The broad confidence intervals in our analysis are not surprising, as it has been reported before that the BDSKY approach returns particularly imprecise estimates if the analysis goes further into the past than the oldest sampling date in the analysis [19,50]. However, as we are primarily interested in the dynamics about the time of the TRIP intervention, the broad HPDs at different time periods are of lesser concern. In the recent years, the HPDs were wider in Kyiv compared to Odessa, however our sensitivity analysis showed that it is unlikely to be attributed solely to the lower number of sequences from Kyiv, as our reduced *Odessa* dataset resulted in the values and HPDs comparable to those obtained from the full *Odessa* dataset.

Importantly, we analysed sequences from Kyiv and Odessa separately as two unrelated epidemics, even though they do not form two independent clades on the Ukrainian HIV phylogenetic tree. Given that the two epidemics share the same ancestry and, thus, parts of the phylogenetic tree, we might have introduced bias for our estimates that go further in the past. In this analysis, apart from 2001–2003, our estimates for both  $R_e$  and the becoming uninfected rate are very similar for Odessa and Kyiv for the period before TRIP. While it is not crucial for the time periods during and after the TRIP intervention, which are relevant in our analysis, this might to some extent explain why our  $R_e$  estimates differ from other studies that focused on the Ukrainian HIV-1 epidemic in mid-90s–2000s.

We want to emphasise that currently there is no way to account for non-random sampling approaches in a phylodynamic context and we do not know how network-based sampling might be affecting estimates obtained in phylodynamic analyses of sequence data, which presents a particular theoretical challenge for the molecular epidemiology field [1]. In the context of HIV transmission, as well as many other infectious diseases, including viral hepatitises, a random sample of an infected population is often unrealistic and epidemiologists proceed with non-probability sampling methods. Consequently, more research is needed to develop phylodynamic frameworks that would be able to account for and accommodate these sampling techniques. Indeed, the sampling approach employed in this work might have biased some of our estimates. The fact that TRIP was a social network-based intervention, with an active contact recruitment approach means that some potential transmission events might be captured in our tree, that otherwise would not be captured under random population sampling. Specifically, the tree may be enriched for infection events among individuals with close or direct epidemiological linkage. It is possible that this would only bias our estimates towards a higher  $R_e$  in Odessa, as *observing* more recent transmission events would be misinterpreted as increased transmission by the model. This would particularly strongly affect any analyses carried out in a coalescent framework, while birth-death models implemented here provide slightly more control through the sampling parameter. Nevertheless, we undertook a sensitivity analysis to investigate this concern. Removing any network-derived TRIP sequences, we still found a reduction in  $R_e$  in Odessa

since 2013 relative to that in Kyiv (Figures 3 and 4), and an even more substantial reduction for the period when the TRIP intervention was implemented.

Crucially, if better surveillance data were available, it could be used to inform priors on some of the birth-death model parameters—particularly on the sampling proportion parameter—which could result in a significant reduction in the uncertainty of other parameters in the analysis. However, this could only be useful when reliable surveillance data are available and would need to be implemented with caution, as we have previously argued that if sampling assumptions of the model are violated, then discrepancy between the observed sampling times and the sampling priors can lead to biased  $R_e$  estimates [52].

The sensitive nature of the questions in the study precluded us from obtaining personal identification information at the point of recruitment. This could have led to individuals appearing multiple times in the dataset, which would introduce biases. ElimDupes was used in an attempt to ameliorate potential repeated recruitment. For example, cluster 3 includes 2 sequences isolated from individuals of the same gender, age, risk group and very similar viral load. Thus, we cannot fully rule out the possibility that cluster 3 includes 2 subsequently sampled strains from one individual, seen by different interviewers. Our sensitivity analyses suggest that if that was the case, and we removed cluster 3 from our analysis, we have found that no statistically significant difference in average viral load between clustered and unclustered individuals would be observed. However, using ElimDupes to ensure that no patient was represented twice in our sample might have reduced our power to detect phylogenetic clusters, as it might happen that two patients have identical sequences.

## 5. Conclusions

The use of HIV genetic sequences allowed us to investigate the impact that an intervention might have had on a specific community and epidemic dynamics, during and after the years it was implemented. While there might have been other factors that have contributed to the reduction in  $R_e$  and the decrease in the average infectious period in Odessa, this exercise allowed us to monitor transmission dynamics in certain communities and further promote preventive strategies associated with a positive dynamics in monitored epidemiological parameters. Importantly, molecular epidemiology methods are a useful tool when investigating the effect of certain prevention strategies and should be widely applied in public health settings.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1999-4915/12/4/469/s1>, Figure S1: Maximum likelihood phylogeny estimated from the Odessa dataset, mid-point rooting, Figure S2: Maximum likelihood phylogeny estimated from the Kyiv dataset, mid-point routing, Figure S3: Maximum likelihood phylogeny estimated from the Odessa dataset after removing the network-derived sequences, mid-point rooting. Remaining potential transmission clusters are marked in yellow, Figure S4: Temporal estimates of the effective reproductive number,  $R_e$ , and the becoming uninfected rate, obtained from the reduced (N = 92) Odessa dataset, Table S1: Sequence accession numbers, Table S2: Model selection procedure with the path-sampling (PS) and stepping-stone (SS) marginal likelihood estimators (MLE).

**Author Contributions:** Conceptualization, T.I.V. and S.R.F.; methodology, T.I.V., S.R.F., A.Z.; formal analysis, T.I.V.; data curation, L.D.W. and A.K.; data interpretation—T.I.V., G.N., D.P., J.S., B.S.; resources, M.L. and V.Z.; A.H.—Generation of some of the DNA sequences; writing—original draft preparation, T.I.V.; writing—review and editing, T.I.V., S.R.F., O.G.P., P.S.; project administration, T.I.V., L.D.W., A.K., P.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** The TRIP intervention was supported by the United States (US) National Institute on Drug Abuse (NIDA) grants DP1 DA034989 and P30DA011041. T.I.V. is supported by the Branco Weiss Fellowship, Junior Research Fellowship at New College, University of Oxford, and Wellcome Institutional Strategic Support Fund 0007785. J.S. is supported by R01AI136056. B.S. is supported by P30 AI036214. A.E.Z. and O.G.P. acknowledge the support of the Oxford Martin School.

**Acknowledgments:** We thank Louis du Plessis (Department of Zoology, University of Oxford) for his input in early discussions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Volz, E.M.; Koelle, K.; Bedford, T. Viral phylodynamics. *PLoS Comput. Biol.* **2013**, *9*, e1002947. [[CrossRef](#)] [[PubMed](#)]
2. Poon, A.F.; Gustafson, R.; Daly, P.; Zerr, L.; Demlow, S.E.; Wong, J.; Woods, C.K.; Hogg, R.S.; Kraiden, M.; Moore, D.; et al. Near real-time monitoring of HIV transmission hotspots from routine HIV genotyping: An implementation case study. *Lancet HIV* **2016**, *3*, e231–e238. [[CrossRef](#)]
3. Wertheim, J.O.; Murrell, B.; Mehta, S.R.; Forgione, L.A.; Kosakovsky Pond, S.L.; Smith, D.M.; Torian, L.V. Growth of HIV-1 Molecular Transmission Clusters in New York City. *J. Infect. Dis.* **2018**, *218*, 1943–1953. [[CrossRef](#)]
4. Vasylyeva, T.I.; du Plessis, L.; Pineda-Pena, A.C.; Kuhnert, D.; Lemey, P.; Vandamme, A.M.; Gomes, P.; Camacho, R.J.; Pybus, O.G.; Abecasis, A.B.; et al. Tracing the Impact of Public Health Interventions on HIV-1 Transmission in Portugal Using Molecular Epidemiology. *J. Infect. Dis.* **2019**, *220*, 233–243. [[CrossRef](#)]
5. Faria, N.R.; Rambaut, A.; Suchard, M.A.; Baele, G.; Bedford, T.; Ward, M.J.; Tatem, A.J.; Sousa, J.D.; Arinaminpathy, N.; Pepin, J.; et al. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science* **2014**, *346*, 56–61. [[CrossRef](#)] [[PubMed](#)]
6. Bbosa, N.; Ssemwanga, D.; Nsubuga, R.N.; Salazar-Gonzalez, J.F.; Salazar, M.G.; Nanyonjo, M.; Kuteesa, M.; Seeley, J.; Kiwanuka, N.; Bagaya, B.S.; et al. Phylogeography of HIV-1 suggests that Ugandan fishing communities are a sink for, not a source of, virus from general populations. *Sci. Rep.-UK* **2019**, *9*. [[CrossRef](#)] [[PubMed](#)]
7. Stadler, T.; Kouyos, R.; von Wyl, V.; Yerly, S.; Boni, J.; Burgisser, P.; Klimkait, T.; Joos, B.; Rieder, P.; Xie, D.; et al. Estimating the basic reproductive number from viral sequence data. *Mol. Biol. Evol.* **2012**, *29*, 347–357. [[CrossRef](#)]
8. Stadler, T.; Kuhnert, D.; Bonhoeffer, S.; Drummond, A.J. Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 228–233. [[CrossRef](#)]
9. Kuhnert, D.; Stadler, T.; Vaughan, T.G.; Drummond, A.J. Simultaneous reconstruction of evolutionary history and epidemiological dynamics from viral sequences with the birth-death SIR model. *J. R. Soc. Interface* **2014**, *11*, 20131106. [[CrossRef](#)]
10. Cohen, M.S.; Chen, Y.Q.; McCauley, M.; Gamble, T.; Hosseinipour, M.C.; Kumarasamy, N.; Hakim, J.G.; Kumwenda, J.; Grinsztejn, B.; Pilotto, J.H.; et al. Prevention of HIV-1 infection with early antiretroviral therapy. *N. Engl. J. Med.* **2011**, *365*, 493–505. [[CrossRef](#)]
11. Latkin, C.A.; Davey-Rothwell, M.A.; Knowlton, A.R.; Alexander, K.A.; Williams, C.T.; Boodram, B. Social network approaches to recruitment, HIV prevention, medical care, and medication adherence. *J. Acquir. Immune Defic. Syndr.* **2013**, *63*, S54–S58. [[CrossRef](#)] [[PubMed](#)]
12. Friedman, S.R.; Downing, M.J., Jr.; Smyrnov, P.; Nikolopoulos, G.; Schneider, J.A.; Livak, B.; Magiorkinis, G.; Slobodianyk, L.; Vasylyeva, T.I.; Paraskevis, D.; et al. Socially-integrated transdisciplinary HIV prevention. *AIDS Behav.* **2014**, *18*, 1821–1834. [[CrossRef](#)] [[PubMed](#)]
13. Nikolopoulos, G.K.; Pavlitina, E.; Muth, S.Q.; Schneider, J.; Psychogiou, M.; Williams, L.D.; Paraskevis, D.; Sypsa, V.; Magiorkinis, G.; Smyrnov, P.; et al. A network intervention that locates and intervenes with recently HIV-infected persons: The Transmission Reduction Intervention Project (TRIP). *Sci. Rep.* **2016**, *6*, 38100. [[CrossRef](#)] [[PubMed](#)]
14. Smyrnov, P.; Williams, L.D.; Korobchuk, A.; Sazonova, Y.; Nikolopoulos, G.K.; Skaathun, B.; Morgan, E.; Schneider, J.; Vasylyeva, T.I.; Friedman, S.R. Risk network approaches to locating undiagnosed HIV cases in Odessa, Ukraine. *J. Int. AIDS Soc.* **2018**, *21*. [[CrossRef](#)]
15. Korobchuk, A.; Davtyan, H.; Denisiuk, O.; Zachariah, R.; Nikolopoulos, G.K.; Paraskevis, D.; Skaathun, B.; Schneider, J.; Vasylyeva, T.I.; Williams, L.D.; et al. People with high HIV viral load within risk networks: Who are these people and who refers them best? *J. Infect. Dev. Ctries* **2019**, *13*, 1035–110S. [[CrossRef](#)]
16. Kovtun, O.; Lysenko, Y.; Sokolenko, O.; Mikhedko, A.; kurdu, O.; Korotych, O.; Kuzin, I. The medical information system “HIV-infection in Ukraine”: Development and implementation. *Public Health Panor.* **2019**, *5*, 1–121.



17. Saad, M.D.; Shcherbinskaya, A.M.; Nadai, Y.; Kruglov, Y.V.; Antonenko, S.V.; Lyullchuk, M.G.; Kravchenko, O.N.; Earhart, K.C.; Sanchez, J.L.; Birx, D.L.; et al. Molecular epidemiology of HIV Type 1 in Ukraine: Birthplace of an epidemic. *AIDS Res. Hum. Retrovir.* **2006**, *22*, 709–714. [[CrossRef](#)]
18. Barska, Y.G.; Yermolenko, N.Y.; Novak, H.Y.; Sazonova, Y.O.; Strembitska, A.S.; Tonkonog, L.M. *Monitoring Behavior and HIV Prevalence among Injecting Drug Users. Regional Profiles Based on the Results of the 2015 Bio-Behavioral Study*; Alliance for Public Health: Kyiv, Ukraine, 2016.
19. Sazonova, Y.; Dukach, Y. *Report on the Results of a Bio-Behavioral Study of Men who have Sex with Men in Ukraine*; Alliance for Public Health: Kyiv, Ukraine, 2019.
20. Vasylyeva, T.I.; Friedman, S.R.; Smyrnov, P.; Bondarenko, K. A new approach to prevent HIV transmission: Project Protect intervention for recently infected individuals. *AIDS Care* **2015**, *27*, 223–228. [[CrossRef](#)]
21. Vasylyeva, T.I.; Liulchuk, M.; Friedman, S.R.; Sazonova, I.; Faria, N.R.; Katzourakis, A.; Babii, N.; Scherbinska, A.; Theze, J.; Pybus, O.G.; et al. Molecular epidemiology reveals the role of war in the spread of HIV in Ukraine. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 1051–1056. [[CrossRef](#)]
22. Kuzin, I.M.V.; Antonenko, J. *HIV-infection in Ukraine: Information Bulletin N50*; Center for Public Health: Kyiv, Ukraine, 2019.
23. Magiorkinis, G.; Karamitros, T.; Vasylyeva, T.I.; Williams, L.D.; Mbisa, J.L.; Hatzakis, A.; Paraskevis, D.; Friedman, S.R. An Innovative Study Design to Assess the Community Effect of Interventions to Mitigate HIV Epidemics Using Transmission-Chain Phylodynamics. *Am. J. Epidemiol.* **2018**, *187*, 2615–2622. [[CrossRef](#)]
24. Kostaki, E.G.; Frampton, D.; Paraskevis, D.; Pantavou, K.; Ferns, B.; Raffle, J.; Grant, P.; Kozlakidis, Z.; Hadjickou, A.; Pavlitina, E.; et al. Near Full-length Genomic Sequencing and Molecular Analysis of HIV-Infected Individuals in a Network-based Intervention (TRIP) in Athens, Greece: Evidence that Transmissions Occur More Frequently from those with High HIV-RNA. *Curr. HIV Res.* **2018**, *16*, 345–353. [[CrossRef](#)]
25. Cunningham, S.; Ank, B.; Lewis, D.; Lu, W.; Wantman, M.; Dileanis, J.A.; Jackson, J.B.; Palumbo, P.; Krogstad, P.; Eshleman, S.H. Performance of the applied biosystems ViroSeq human immunodeficiency virus type 1 (HIV-1) genotyping system for sequence-based analysis of HIV-1 in pediatric plasma samples. *J. Clin. Microbiol.* **2001**, *39*, 1254–1257. [[CrossRef](#)]
26. Pineda-Pena, A.C.; Faria, N.R.; Imbrechts, S.; Libin, P.; Abecasis, A.B.; Deforche, K.; Gomez-Lopez, A.; Camacho, R.J.; de Oliveira, T.; Vandamme, A.M. 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. [[CrossRef](#)] [[PubMed](#)]
27. Struck, D.; Lawyer, G.; Ternes, A.M.; Schmit, J.C.; Bercoff, D.P. COMET: Adaptive context-based modeling for ultrafast HIV-1 subtype identification. *Nucleic Acids Res.* **2014**, *42*, e144. [[CrossRef](#)] [[PubMed](#)]
28. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)] [[PubMed](#)]
29. Ragonnet-Cronin, M.; Hodcroft, E.; Hue, S.; Fearnhill, E.; Delpech, V.; Brown, A.J.; Lycett, S.; Database, U.H.D.R. Automated analysis of phylogenetic clusters. *BMC Bioinform.* **2013**, *14*, 317. [[CrossRef](#)] [[PubMed](#)]
30. Hassan, A.S.; Pybus, O.G.; Sanders, E.J.; Albert, J.; Esbjornsson, J. Defining HIV-1 transmission clusters based on sequence data. *AIDS* **2017**, *31*, 1211–1222. [[CrossRef](#)] [[PubMed](#)]
31. Suchard, M.A.; Lemey, P.; Baele, G.; Ayres, D.L.; Drummond, A.J.; Rambaut, A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **2018**, *4*, vey016. [[CrossRef](#)] [[PubMed](#)]
32. Gelman, A.; Meng, X.L. Simulating normalizing constants: From importance sampling to bridge sampling to path sampling. *Stat. Sci.* **1998**, *13*, 163–185. [[CrossRef](#)]
33. Xie, W.; Lewis, P.O.; Fan, Y.; Kuo, L.; Chen, M.H. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst. Biol.* **2011**, *60*, 150–160. [[CrossRef](#)] [[PubMed](#)]
34. Drummond, A.J.; Rambaut, A.; Shapiro, B.; Pybus, O.G. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **2005**, *22*, 1185–1192. [[CrossRef](#)]
35. Rambaut, A.; Lam, T.T.; Max Carvalho, L.; Pybus, O.G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2016**, *2*, vew007. [[CrossRef](#)] [[PubMed](#)]
36. Abecasis, A.B.; Vandamme, A.M.; Lemey, P. Quantifying differences in the tempo of human immunodeficiency virus type 1 subtype evolution. *J. Virol.* **2009**, *83*, 12917–12924. [[CrossRef](#)] [[PubMed](#)]

37. Patino-Galindo, J.A.; Gonzalez-Candelas, F. The substitution rate of HIV-1 subtypes: A genomic approach. *Virus Evol.* **2017**, *3*, vex029. [[CrossRef](#)] [[PubMed](#)]
38. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **2018**, syy032. [[CrossRef](#)]
39. Kuhnert, D.; Stadler, T.; Vaughan, T.G.; Drummond, A.J. Phylodynamics with Migration: A Computational Framework to Quantify Population Structure from Genomic Data. *Mol. Biol. Evol.* **2016**, *33*, 2102–2116. [[CrossRef](#)]
40. Bouckaert, R.; Vaughan, T.G.; Barido-Sottani, J.; Duchene, S.; Fourment, M.; Gavryushkina, A.; Heled, J.; Jones, G.; Kuhnert, D.; De Maio, N.; et al. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **2019**, *15*, e1006650. [[CrossRef](#)]
41. Diez-Fuertes, F.; Cabello, M.; Thomson, M.M. Bayesian phylogeographic analyses clarify the origin of the HIV-1 subtype A variant circulating in former Soviet Union's countries. *Infect. Genet. Evol.* **2015**, *33*, 197–205. [[CrossRef](#)]
42. Vasylyeva, T.I.; Friedman, S.R.; Lourenco, J.; Gupta, S.; Hatzakis, A.; Pybus, O.G.; Katzourakis, A.; Smyrnov, P.; Karamitros, T.; Paraskevis, D.; et al. Reducing HIV infection in people who inject drugs is impossible without targeting recently-infected subjects. *AIDS* **2016**, *30*, 2885–2890. [[CrossRef](#)]
43. Benjamini, Y.; Drai, D.; Elmer, G.; Kafkafi, N.; Golani, I. Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* **2001**, *125*, 279–284. [[CrossRef](#)]
44. Grabowski, M.K.; Herbeck, J.T.; Poon, A.F.Y. Genetic Cluster Analysis for HIV Prevention. *Curr. HIV/AIDS Rep.* **2018**, *15*, 182–189. [[CrossRef](#)] [[PubMed](#)]
45. Bezemer, D.; van Sighem, A.; Lukashov, V.V.; van der Hoek, L.; Back, N.; Schuurman, R.; Boucher, C.A.; Claas, E.C.; Boerlijst, M.C.; Coutinho, R.A.; et al. Transmission networks of HIV-1 among men having sex with men in the Netherlands. *AIDS* **2010**, *24*, 271–282. [[CrossRef](#)] [[PubMed](#)]
46. Brenner, B.G.; Roger, M.; Routy, J.P.; Moisi, D.; Ntemgwa, M.; Matte, C.; Baril, J.G.; Thomas, R.; Rouleau, D.; Bruneau, J.; et al. High rates of forward transmission events after acute/early HIV-1 infection. *J. Infect. Dis.* **2007**, *195*, 951–959. [[CrossRef](#)] [[PubMed](#)]
47. Yerly, S.; Junier, T.; Gayet-Ageron, A.; Amari, E.B.; von Wyl, V.; Gunthard, H.F.; Hirschel, B.; Zdobnov, E.; Kaiser, L.; Swiss, H.I.V.C.S. The impact of transmission clusters on primary drug resistance in newly diagnosed HIV-1 infection. *AIDS* **2009**, *23*, 1415–1423. [[CrossRef](#)]
48. Volz, E.M.; Koopman, J.S.; Ward, M.J.; Brown, A.L.; Frost, S.D. Simple epidemiological dynamics explain phylogenetic clustering of HIV from patients with recent infection. *PLoS Comput. Biol.* **2012**, *8*, e1002552. [[CrossRef](#)]
49. Chalmet, K.; Staelens, D.; Blot, S.; Dinakis, S.; Pelgrom, J.; Plum, J.; Vogelaers, D.; Vandekerckhove, L.; Verhofstede, C. Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections. *BMC Infect. Dis.* **2010**, *10*, 262. [[CrossRef](#)]
50. Paraskevis, D.; Beloukas, A.; Stasinou, K.; Pantazis, N.; de Mendoza, C.; Bannert, N.; Meyer, L.; Zangerle, R.; Gill, J.; Prins, M.; et al. HIV-1 molecular transmission clusters in nine European countries and Canada: Association with demographic and clinical factors. *BMC Med.* **2019**, *17*, 4. [[CrossRef](#)] [[PubMed](#)]
51. Magee, A.F.; Höhna, S.; Vasylyeva, T.I.; Leache, A.; Minin, V.N. Locally adaptive Bayesian birth-death model successfully detects slow and rapid rate shifts. *BioRxiv* **2019**. [[CrossRef](#)]
52. Vasylyeva, T. *Using Molecular Epidemiology to Investigate HIV Transmission and Prevention Efforts*; University of Oxford: Oxford, UK, 2018; p. 250.

