



# Ethical considerations in global HIV phylogenetic research

Cordelia E M Coltart\*, Anne Hoppe\*, Michael Parker, Liza Dawson, Joseph J Amon, Musonda Simwanga, Gail Geller, Gail Henderson, Oliver Laeyendecker, Joseph D Tucker, Patrick Eba, Vladimir Novitsky, Anne-Mieke Vandamme, Janet Seeley, Gina Dallabetta, Guy Harling, M Kate Grabowski, Peter Godfrey-Faussett, Christophe Fraser, Myron S Cohen†, Deenan Pillay†; on behalf of the Ethics in HIV Phylogenetics Working Group‡

Phylogenetic analysis of pathogens is an increasingly powerful way to reduce the spread of epidemics, including HIV. As a result, phylogenetic approaches are becoming embedded in public health and research programmes, as well as outbreak responses, presenting unique ethical, legal, and social issues that are not adequately addressed by existing bioethics literature. We formed a multidisciplinary working group to explore the ethical issues arising from the design of, conduct in, and use of results from HIV phylogenetic studies, and to propose recommendations to minimise the associated risks to both individuals and groups. We identified eight key ethical domains, within which we highlighted factors that make HIV phylogenetic research unique. In this Review, we endeavoured to provide a framework to assist researchers, public health practitioners, and funding institutions to ensure that HIV phylogenetic studies are designed, done, and disseminated in an ethical manner. Our conclusions also have broader relevance for pathogen phylogenetics.

## Introduction

Understanding of the transmission dynamics of infectious agents is essential for development of effective public health interventions. Historically, transmission dynamics were investigated with epidemiology: tracking of the evolution of epidemics through time, place, and person, primarily with observation and self-reports of exposure and risk behaviours. However, despite advances in HIV prevention, incidence remains high, notably in sub-Saharan Africa, which accounts for 75% of all new HIV infections worldwide.<sup>1</sup> In this context, an understanding of who is most likely to infect whom remains important to the development of targeted prevention strategies.

In phylogenetic analyses, historical relationships between individuals or groups are deduced by comparison of pathogen genomes to establish how closely related viruses from two individuals are. In combination with traditional epidemiological data, viral genetic sequence data can help to infer transmission patterns. This combination has the potential to answer key questions that are not easily addressed by traditional or molecular approaches alone.<sup>2-4</sup>

The focus of HIV phylogenetic studies has extended from concentrated to generalised epidemics, and increasingly involves large datasets.<sup>5</sup> Funding bodies, such as the Wellcome Trust, the Bill & Melinda Gates Foundation, and the National Institutes of Health, are committed to sharing such datasets to maximise the benefit of HIV research. Additionally, sequence data analysed for publications in scientific journals are usually required to be submitted to open public sequence repositories. The collection, storage, sharing, and research use of such data raises important ethical, legal, and social challenges.

International ethical guidelines for research with human participants, such as the Helsinki Declaration and the Council for International Organizations of Medical Sciences guidelines, address several of these

issues, including the need for informed consent, community engagement, risk minimisation, and consideration of the risks and benefits of research for groups and communities.<sup>6,7</sup> Additionally, a large and diverse collection of academic and policy literature exists that addresses the ethical, legal, and social implications of the research and clinical uses of human genomics in high-income countries.<sup>8,9</sup> This literature has been accompanied by a growing assortment of bioethics and social science literature on the implications of genomic research in low-income and middle-income countries (LMICs),<sup>10,11</sup> and by stakeholder engagement initiatives in these settings (appendix p 3-4).<sup>12-14</sup>

HIV phylogenetic research presents complex ethical issues, including two specific challenges. First, (like contact-tracing data) phylogenetic analyses are fundamentally relational: analysis of data from one person might affect other people (eg, by identifying them as potential sources of infection). Second, as next-generation sequencing (NGS) produces richer sequence data, true anonymisation of viral sequence data becomes difficult because virus isolated from another timepoint in another study could be used to reidentify an individual. In recognition of these issues and the need for the development of good ethical practice models in this area, we held a multidisciplinary (scientists, bioethicists, lawyers, human rights advocates, HIV activists, and community engagement members from Africa) workshop in London (UK) in May, 2017.<sup>15</sup> This meeting focused on identifying the key issues arising from study design and conduct or use of results of HIV phylogenetic studies and on making recommendations regarding the public release and publishing of data obtained from HIV phylogenetic studies in an ethical manner.

In this Review, we summarise the findings and recommendations from the workshop and follow-up discussions, and we set out a framework for researchers and funding bodies supporting HIV genetic studies. This Review also has relevance for the increasing use of

## Lancet HIV 2018

Published Online  
August 30, 2018  
[http://dx.doi.org/10.1016/S2352-3018\(18\)30134-6](http://dx.doi.org/10.1016/S2352-3018(18)30134-6)

\*Joint first authors

†Joint senior authors

‡Members are listed in the appendix

Institute for Global Health (C E M Coltart PhD, G Harling ScD) and Division of Infection and Immunity (A Hoppe PhD, Prof D Pillay PhD), University College London, London, UK; The Wellcome Centre for Ethics and Humanities (Ethox), Nuffield Department of Population Health (Prof M Parker PhD) and Big Data Institute, Nuffield Department of Medicine (Prof C Fraser PhD), University of Oxford, Oxford, UK; Division of AIDS (L Dawson PhD), National Institute of Allergy and Infectious Diseases (O Laeyendecker PhD), National Institutes of Health, Bethesda, MD, USA; Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ, USA (J J Amon PhD); Zambart, University of Zambia, Lusaka, Zambia (M Simwanga PhD); Berman Institute of Bioethics and School of Medicine, Johns Hopkins University, Baltimore, MD, USA (Prof G Geller ScD); Center for Genomics and Society, Department of Social Medicine (Prof G Henderson PhD), Institute for Global Health and Infectious Diseases (J D Tucker PhD, Prof M S Cohen MD), and Department of Medicine (Prof M S Cohen), University of North Carolina, Chapel Hill, NC, USA; Community Support, Social Justice and Inclusion Department (P Eba PhD), Joint United Nations Programme on HIV/AIDS (Prof P Godfrey-Faussett FRCP), Geneva, Switzerland; School of Law, University of KwaZulu-Natal, Pietermaritzburg, South Africa

(P Eba); Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA (V Novitsky PhD); Clinical and Epidemiological Virology, Rega Institute for Medical Research, Department of Microbiology and Immunology, KU Leuven—University of Leuven, Leuven, Belgium (Prof A-M Vandamme PhD); Center for Global Health and Tropical Medicine, Universidade de Microbiologia, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisbon, Portugal (Prof A-M Vandamme); Africa Health Research Institute, KwaZulu-Natal, South Africa (Prof J Seeley PhD, G Harling, Prof D Pillay); Department of Global Health and Development (Prof J Seeley) and Department of Clinical Research (Prof P Godfrey-Faussett), London School of Hygiene & Tropical Medicine, London, UK; Bill & Melinda Gates Foundation, Washington, DC, USA (G Dallabetta MD); Department of Pathology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA (M K Grabowski PhD);

phylogenetics for non-HIV pathogens, including within outbreak response situations.

### Phylogenetics and its role in HIV research

Over time and successive generations, mutations occur in the genetic code of species. Phylogenetic inference exploits these changes to determine the genetic similarity of two organisms, assuming that the more similar their genetic sequence, the closer in time they are to having a common ancestor. Application of this approach to HIV detected in blood samples from infected human populations helps in the understanding of HIV transmission patterns.

HIV is suitable for phylogenetic analysis because it is highly genetically variable.<sup>16</sup> Given that transmission of HIV involves two individuals (a couple), the variability of HIV can be used to infer linkages by forming phylogenetic clusters between couples and groups of people.<sup>17</sup> Furthermore, in many cases inferring, with a degree of uncertainty, the direction of transmission within clusters is possible by use of either additional epidemiological data<sup>18,19</sup> or data with sequences from multiple viruses sampled from each individual.<sup>20,21</sup>

Different techniques can be used to generate sequence data required for phylogenetic analysis. NGS increases both the potential power and potential risks of phylogenetic approaches compared with conventional Sanger sequencing methods<sup>22,23</sup> because it provides information on intrahost variation and returns richer sequence data for multiple viral particles per sample.

Additionally, many methods and associated assumptions in phylogenetic analyses exist for identifying clusters of genetically related viruses (appendix p 5). Phylogenetic clusters are generally thought to represent groups of infected individuals who are closer together in a transmission chain and can be identified with various methods.<sup>3,24–28</sup>

Caution is required when interpreting phylogenetic clusters for epidemiological purposes, because clusters are typically inferred from partially sampled transmission chains (ie, some infected individuals were not sampled). Unsourced cases can be either a common source of infection or an intermediary in a transmission chain for people with genetically similar viruses. Although historically proving transmission between two individuals has been difficult, the use of NGS with improved interpretation algorithms makes such inferences more likely to happen in future.

Phylogenetic analyses can be used widely in HIV epidemiology (figure 1), for example, to study viral linkage and risk factors for epidemic spread (molecular epidemiology),<sup>4</sup> to examine the growth and decline of HIV epidemics (phylogenetics),<sup>3,29,30</sup> or to investigate the impact of migration on HIV spread and to identify hubs of transmission (phylogeography).<sup>31</sup>

HIV phylogenetics has been most applied in high-income countries with well developed scientific infrastructure, where HIV is characterised by smaller epidemics focused in specific risk groups (concentrated epidemics). In many high-income countries, sequencing of the HIV *pol* gene is used to monitor both transmitted drug resistance at time of diagnosis and emerging drug resistance on antiretroviral therapy.<sup>32</sup> This sequencing has led to the growth of national HIV genetic databases, such as those in the UK<sup>33</sup> and Switzerland.<sup>34</sup> If these datasets are linked to epidemiological surveillance and clinical cohort data, inferences can be made with regard to patterns of exposure and risk factors for onward transmission among infected individuals.<sup>18,29,35</sup> Unlike standard epidemiological data, molecular data can allow inferences to be made regarding the time of transmission relative to the time of sample collection. Furthermore, data obtained through phylogenetic analyses can be used to validate self-reported epidemiological data in relation to sexual and other behaviours. Combined with traditional epidemiological methods, phylogenetic research provides a more detailed and precise understanding of epidemic characteristics, thereby enabling improved public health policies, including more effective and better targeted programmes for prevention and treatment.<sup>36,37</sup>

Many African epidemics are much larger than those in the USA and Europe. Viral sequencing in Africa is not routinely done outside targeted programmes, such as WHO's HIV drug resistance surveillance,<sup>38</sup> and research projects. Although declining costs and progressively easier sequencing will increase the proportion of infected individuals represented within sequence databases, such as

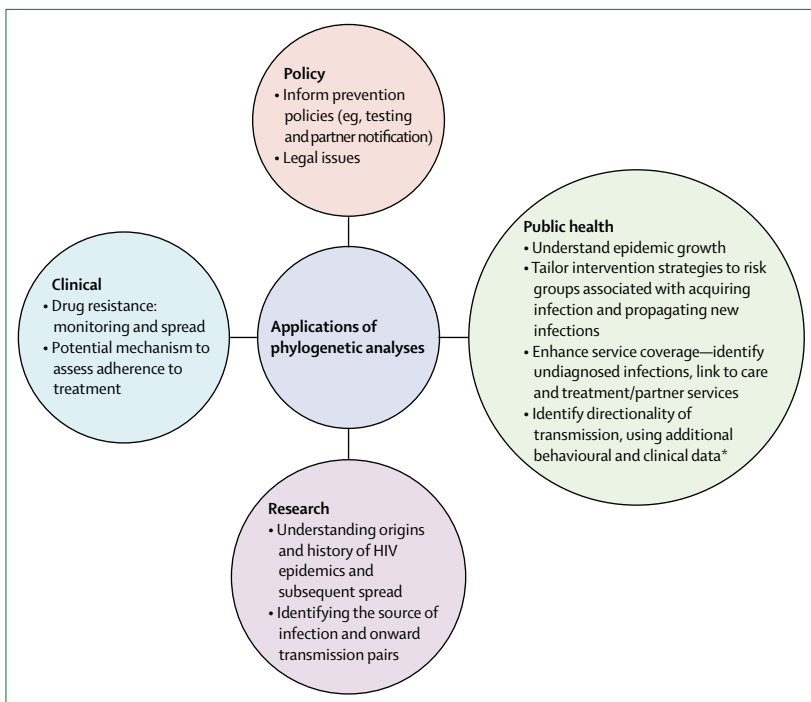


Figure 1: Applications of phylogenetic analyses

\*Denotes a potential future use of phylogenetic analyses.

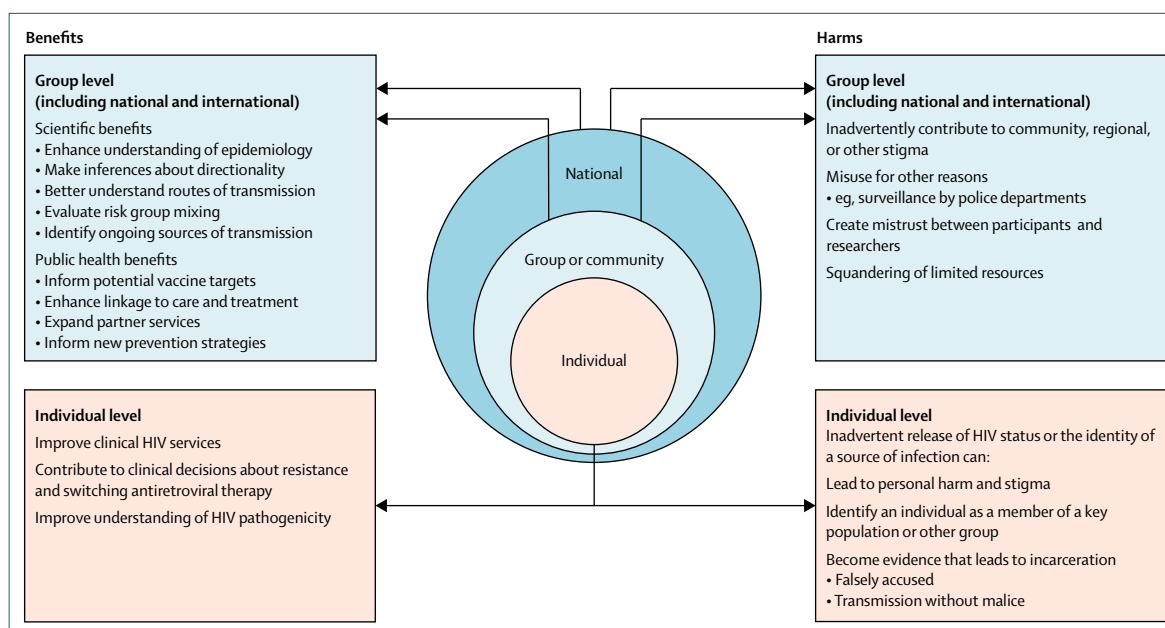


Figure 2: Potential benefits and harms associated with HIV phylogenetic analysis

PANGEA-HIV,<sup>5</sup> Africa does not yet have the extensive and comprehensive datasets seen in high-income countries. Additionally, community and patient mobilisation around HIV takes very different forms compared with that in high-income countries, and the social, political, and economic context is considerably different and varies among African countries. Ethical analysis of phylogenetic work will need to take account of international variation in both epidemic characteristics and local economic, legal, and social contexts.

### Key ethical issues arising in phylogenetic studies of HIV transmission

Some of the ethical issues raised by HIV phylogenetic research are similar to those in traditional epidemiology studies. These issues include the potential for stigmatisation and risk of social harm to individuals or groups, and concerns about privacy, confidentiality, and security of data. However, certain risks are particularly salient in phylogenetic research, which we discuss over eight key ethical domains.

### Risk and benefit assessments

The harms and benefits of phylogenetic research will vary depending on whether they are assessed at an individual, group, or societal level (figure 2). Information obtained through phylogenetic analyses should be used to advance socially valuable goals, such as reducing the spread of HIV, while at the same time minimising the risks to individuals, groups, and populations. Of particular concern to phylogenetics is the possible deduction of complex social and sexual relationships, should phylogenetic data be combined with minimal

clinical and demographic information. Conversely, traditional epidemiological studies require far more information to infer transmission of HIV between individuals, particularly with respect to directionality of infection.

Risks to individuals principally arise from inadvertent or intentional disclosure of HIV status or transmission events, or from demands for these data for judicial or extrajudicial targeting of individuals or groups. In several countries, phylogenetic evidence is being used in criminal cases of alleged HIV transmission.<sup>39–42</sup> Breaches of confidentiality could occur through inadequate anonymisation or deductive disclosure, through misinterpretation, miscommunication, or misuse of the analytic results, or through legal action.

These risks will increase if more data are generated and made publicly available, a requirement of many funding agencies and publishers. Although publication might maximise the scientific research value of a dataset, it raises concerns about how the data are used, appropriate consent for such use, confidentiality, and stigma. Furthermore, contrary to epidemiological studies, in which individuals can choose what information they disclose to investigators, inferences made from viral genetic sequences are not controlled by participants.

Anonymisation can provide some protection to individuals. However, even with anonymisation, deductive disclosure of identities from HIV sequences and other corresponding data remains theoretically possible. Through the use of rich NGS data applied at successive timepoints, a virus sequenced from an individual could in principle be used to relink that individual to an earlier study with high reliability. Small fragments of human

and Department of Rakai Community Cohort Study, Rakai Health Sciences Program, Kalisizo, Uganda (M K Grabowski)

Correspondence to: Dr Anne Hoppe, Division of Infection and Immunity, University College London, London WC1E 6BT, UK [a.hoppe@ucl.ac.uk](mailto:a.hoppe@ucl.ac.uk)

See Online for appendix

**Panel 1: Migration in Botswana**

Studying migrants is often fraught with both logistical and ethical problems. Migration has been identified as a key risk factor for the spread of HIV possibly because of the lack of access to culturally and linguistically appropriate prevention information and clinical care, disruption of established social relationships, and the potential for increased risky sexual practices when people are away from home.<sup>44,45</sup> However, a more nuanced approach to migration and the link to HIV is needed to take differences in migration flows, risk environments, and characteristics of the areas between which migrants move into account.

Migratory populations in Botswana (documented and undocumented, skilled and unskilled) face challenges in accessing health-care services and are prohibited from receiving free government-provided antiretroviral therapy.<sup>46</sup> According to the 2014 national census,<sup>47</sup> migrants accounted for approximately 14% of employed and 9% of unemployed populations in Francistown, Botswana. A large proportion of HIV-infected migrants in Botswana are unaware of their positive HIV status and are not on antiretroviral therapy, and might therefore disproportionately contribute to new HIV transmissions. Phylogenetic studies are able to assess the amount of integration of migrants in the generalised HIV epidemic in Botswana. However, ethical issues and risks associated with the participation of migrants in phylogenetic research are substantial. Migrants might face additional stigmatisation and marginalisation. They might also become subject to deportation, imprisonment, or extortion in exchange for short-term visas.<sup>48</sup> Obtaining consent can inhibit participation due to privacy and identification issues related to their undocumented status. Therefore, anonymous enrolling of migrants is ethically preferable to avoid these social harms.

DNA sequence contained in NGS data could also be accidentally released. Furthermore, HLA of the infected individual is imprinted on the virus because of immune selection, which might assist identification of individuals in the future.<sup>43</sup> The probability of relinking individuals to anonymised data can be minimised by processing sequence data before release (eg, by only including consensus sequences, which suffice for many phylogenetic methods).

Despite these risks, maintaining links to individuals' identities might allow for direct benefits to individuals. Sequence information can provide clinical guidance (eg, by allowing treatment optimisation following detection of drug resistance mutations). Most HIV phylogenetic studies to date have used data obtained for clinical drug resistance testing, from resistance surveillance programmes, or as part of broader research studies.

At the population level, phylogenetic analysis can allow individuals' data to be linked in a network, enabling inference about the characteristics of networks and identification of risk groups. This information could be used to focus public health interventions towards specific groups at high risk of both acquiring and transmitting the infection.

The choice of metadata variables used in phylogenetic analysis is an important ethical decision. Phylogenetic analyses are often based on individual-level demographic, behavioural, or clinical variables, ignoring structural and environmental factors. The perception that certain groups (eg, key populations, such as men who have sex with men [MSM] or people who inject drugs) are

responsible for infecting others and sustaining the HIV epidemic might be reinforced by only focusing on these variables. Conversely, other structural factors, such as those highlighted in the case study of migration in Botswana (panel 1), as well as sexual violence, lack of access to prevention and treatment, and having experienced discrimination, can have a prominent role in HIV transmission.<sup>49</sup> Studying these factors and their effect on HIV transmission risk can decrease the blame mentality and create an alternative understanding of how to reduce HIV transmission and which individuals or groups are most at risk and why.

Plans to address the risks to individuals and groups should be developed in the planning stages of research projects. For protection of individuals, particularly the risk of criminal prosecution or other targeting based on either HIV status or HIV transmission events, anonymisation of data provides substantial protection. Although individuals could theoretically be identified through reanalysing and relinking anonymised data from different sources, it would be difficult and require specialised expertise. By contrast, datasets linked with individual identifiers could be subpoenaed or obtained through unauthorised means, putting individuals at risk.

Researchers need to assess carefully the possibility of specific individuals or groups of people being identified from their data, whether this identification could provide benefit in informing targeted treatment or interventions, and whether these benefits outweigh the risks to individuals and groups by being identified. Preference should be given to other approaches that achieve the same research objective, but involve less risk. An ongoing monitoring of anticipated and unanticipated risks should be built into HIV phylogenetic research and mitigation strategies identified as early as possible.

**Protection of the rights and interests of study participants**

Effective phylogenetic work often occurs at the interface between research and public health practice because the same data can be used for both purposes. Researchers are typically viewed as obliged to protect individuals who enrol in a study from risk of harm, as far as possible, while pursuing valuable knowledge. Conversely, public health agencies have the mission of protecting the health of the public, which sometimes involves overruling individuals' privacy interests to use data for public-health decision making. When research also has implications for specific population groups, further considerations relating to group harm are important (panel 2).<sup>52-54</sup>

The obligation of researchers to communicate results to study participants needs to be evaluated for each phylogenetic study. When clinical action is required, there is an obligation to make results available. In general, this goal (and benefit) is theoretical because phylogenetic results are produced with a substantial delay from sampling; therefore, any result is unlikely to be timely in informing clinical care. However, with the evolution of



real-time phylogenetics, reporting of drug resistance data to study participants could result in changes to clinical management. A second potential issue is the source of HIV acquisition in discordant couples. Whether HIV acquisition events are linked to the known infected partner might be crucial for interpretation of the efficacy of prevention strategies (panel 2). However, disclosure of these results to study participants could lead to adverse consequences for individuals involved.<sup>50</sup>

### Local social and legal context, including human rights violations

Comprehension of the local social and legal context is crucial to understanding the risks and benefits associated with phylogenetic research. For example, although certificates of confidentiality are legally binding tools to protect research participants and researchers from being compelled to reveal personal data in the USA,<sup>55</sup> similar protocols do not exist in many other high-income countries or in LMICs. Knowledge of local legal proceedings is essential for ensuring that research data are unavailable to subpoena, reducing the individual risk. Furthermore, changes to the social and legal environment need to be regularly monitored to ensure that new risks are not introduced during the course of the study. Researchers need to be clear with policy makers about how proposed laws or policies could negatively impact HIV research efforts and interventions.

The global human right to health<sup>6,56</sup> encompasses prevention and treatment, and a right to privacy, consent, and freedom from discrimination and violence, yet persecution of key populations in Africa remains widespread.<sup>57,58</sup> Research methods have been used to violate individuals' rights, including the use of key population mapping by police to arrest and harass sex workers and MSM in Nigeria in 2014,<sup>59</sup> impose travel bans on foreigners, enforce restrictions on access to housing, schooling, and employment, and trigger violent attacks, including murder. The use of phylogenetic methods for public health should incorporate key legal and human rights prerequisites (panel 3) that are based on standards provided under international human rights treaties as well as national constitutions and legislation.

Globally, 72 countries (a third of them in Africa) have laws specifically allowing for HIV criminalisation,<sup>60</sup> which has led to the use of phylogenetic analyses in criminal convictions (panel 4). Government officials, or other actors, can misinterpret, or wilfully misconstrue, the results of phylogenetic research in support of political agendas or criminal convictions, putting individuals at risk of criminal prosecution for HIV transmission or broader human rights abuses. Realisation of the possible consequences for privacy and prosecution can lead to a reluctance to test, failure to disclose contacts, or refusal of resistance testing in these communities.<sup>60,64,65</sup> The likelihood of misuse and abuse of these data is high, particularly for stigmatised populations. Researchers should alert ethical review

### Panel 2: Examples of research and clinical challenges that are prevalent in phylogenetic analyses

#### Serodiscordant couples

HIV transmission often occurs within discordant couples, in which one person is HIV positive and the other is not. Phylogenetic analysis allows identification of linked transmission between the members of the couple. In one HIV study enrolling discordant couples, 30% of the HIV acquisition events were not linked to the known infected partner.<sup>50</sup> Although this information was essential for interpreting the efficacy of the prevention strategy tested in this particular trial (the interventions to reduce transmission were focused on the HIV-positive partner), these results were not provided to study participants for fear of adverse consequences for the individuals involved.<sup>51</sup> For example, domestic violence, loss of trust in relationships, and relationship break-ups might all result from disclosure of partnerships. It might not be possible to mitigate these risks with counselling or follow-up.

#### Detection of transmission events in non-treatment-compliant patients

If antiretroviral therapy does not suppress viral replication, some new HIV transmissions from those patients receiving treatment could occur. Linked to epidemiological data, phylogenetics have the potential to differentiate between transmitted drug resistance and poor adherence to antiretroviral therapy, thereby allowing health-care professionals to initiate an appropriate intervention (eg, treatment switch vs adherence counselling). However, if non-adherence results in transmission of HIV, then additional ethical issues can arise, particularly in relation to transmission laws.

### Panel 3: Key legal and human rights prerequisites for the use of phylogenetic methods for public health research

- Informed consent for collection and dissemination of phylogenetic data and information.
- Confidentiality, safety, and prevention of unauthorised use of phylogenetic data and information.
- Non-stigmatisation and non-discrimination in collection and publication of phylogenetic data.
- Attention to criminalisation and other potentially negative consequences relating to collection and dissemination of phylogenetic data.
- Specific gender consideration and attention to the particular risks and concerns faced by women and key populations, due to coercive social and legal environments.
- Awareness that in some countries, collection and publication of phylogenetic data might require legislative or policy change.
- Community participation and accountability for collection and use of data.
- Legal redress in case of misuse of phylogenetic data.
- Phylogenetic experts need to be consistent in their statements that source attribution cannot be definitively determined from phylogenetics alone.

#### Legal associated risks of misuse of phylogenetic data

- Minimal information is required for self-identification, even with anonymisation, which might provide information about individuals in the same network, leading to attribution of blame for infections, which might increase prosecution episodes.
- Research data can be subject to subpoena because of laws on accessing public health data, which might result in misuse by governments and police to target vulnerable populations.

committees and suspend research when risks to study participants increase.

Misuse of phylogenetic data, including seizure by police and subpoena in criminal proceedings, or the perceptions of people with HIV and members of key

#### Panel 4: Use of phylogenetic analysis in criminal convictions

Since the Florida dentist case in the beginning of the 1990s,<sup>39</sup> phylogenetic analyses started to be used in court cases as a forensic tool in HIV transmission investigations (eg, cases in which one or more complainants allege that a defendant has unlawfully infected them with HIV).<sup>42</sup> Cases can be criminal (in countries where transmission of HIV infection is specifically criminalised) or civil (in the context of general civil laws, for example, by applying physical or sexual assault laws to HIV-related cases). Most HIV-specific laws are overly broad or vague, and do not require proof of transmission for conviction; prosecution is often based on potential or perceived exposure with allegations of non-disclosure. However, when general criminal laws (such as those relating to bodily harm) are applied to allegations of HIV transmission, proof of causality is often required.

Phylogenetic evidence cannot stand alone in court and should be used in the context of other evidence, such as full epidemiological investigation and contact tracing.<sup>41,61-63</sup>

Experts have worked with the Crown Prosecution Service for England and Wales to highlight the limitations and challenges of phylogenetics in prosecution cases including:

- Phylogenetic information based on Sanger sequencing alone cannot prove transmission beyond reasonable doubt; although an indirect link can never be ruled out. By contrast, substantially separated clustering can be used as evidence against direct transmission, provided the samples have been drawn close enough to the timing of transmission and do not get phylogenetically separated by onward transmission events.
- Communication of results to non-experts has challenges, such as the lack of certainty.
- Identification of the source of a transmission is not possible, as it would require for all strains of all patients ever infected with HIV to be available as controls, and for phylogenetic trees to flawlessly reconstruct a true epidemic history. Both assumptions are unrealistic.

populations that data might be misused against them can undermine trust in research projects and health-care systems, putting HIV prevention and treatment programmes at risk. Research done in countries where privileged information between medical practitioners and their patients can be seized in HIV-related criminal trials showed that people with HIV were more reluctant to speak openly with their practitioners about their sexual partners and practices.<sup>61,65</sup> These risks can be mitigated in phylogenetic studies through awareness of social and legal issues and ensuring they are addressed at the planning stages and monitored throughout the project.

#### Risk mitigation strategies to protect individual and group identities

Many of the risks associated with identification of individuals from phylogenetic information in environments with oppressive laws and policies can be reduced through use of anonymisation. Therefore, one default position is that anonymisation of data is preferred, if the scientific objectives can be accomplished. This position presumes no overriding interest in individuals receiving research results at the individual level. If data are not relevant for clinical care (eg, because of a substantial time delay between sample collection and generation of sequence information) then little rationale

exists for returning data to health-care workers. When sequence analysis is timely, resistance data should be returned to clinics before doing phylogenetic analyses on anonymised sequencing data.

If anonymisation is detrimental to the scientific objectives or public health, further ethical analysis must be done and specific steps taken to protect the data from use in harmful proceedings. These steps might be technical (eg, storage linkage to identifiers in coded, separate databases with controlled access) or legal (eg, legal agreements protecting data from disclosure for the duration of the study).

The deanonymisation risk for individuals by use of later samples, such as infected blood collected at a different timepoint, as part of a linked or unrelated study or as part of a criminal investigation, can be mitigated by restricting the amount of data shared. Such restriction could include limitation to one virus sequence per individual, storage of raw NGS data under managed access, or destruction of raw data. However, these actions could reduce the potential scientific benefits of the study (eg, because they limit the ability to apply newly developed bioinformatic algorithms to infer the direction of transmission).

Risks to groups cannot be addressed through anonymisation of individuals. Certain groups can be placed at risk through characterisation as high risk or likely to transmit virus, including geographically defined groups, sexual or gender minorities, or those defined by ethnicity, nationality, or migration status. Mitigation plans to address these risks need to include consultation with community representatives, consideration of the public health value of the findings, and development of communication plans in formats and venues that are least damaging to vulnerable groups. In some cases, detailed findings might need to be communicated confidentially, rather than publicly, and some group descriptors might need to be masked in research publications and press releases. Risk mitigation strategies must also provide for redress mechanisms in cases of abuse or misuse of phylogenetic data. These strategies might require the establishment of ties with local legal services, organisations working to protect people with HIV, and criminalised or stigmatised populations, to ensure that they have access to the means to protect their rights.

Training researchers and health-care professionals involved in phylogenetic investigations on the potential of harm to communities and individuals is an important risk mitigation strategy. Such training should aim to ensure that research staff are sensitive to the risk of harm and understand key issues of anonymity, confidentiality, informed consent, and protection of research participants and communities.

#### Valid informed consent and other safeguards

The formal requirements for the achievement of valid consent are well established in literature and

guidelines.<sup>67,66–68</sup> The issues arising in relation to consent for phylogenetic studies are likely to be multifaceted. The procurement of community assent (via community leaders) and individual informed consent is particularly challenging for complex scientific studies, such as phylogenetic research, which involve concepts that are hard both to explain and to understand, and have multiple possible risks and benefits.

The complexity of concepts involved in phylogenetic research can raise fears about the aims of the work and the implications of participation among research participants, front-line research staff, health-care professionals, and ethics committee members. Communication methods that increase the understanding of phylogenetic studies need to be designed and evaluated. These must emphasise potential harms, thoughtful mitigation of harms to risk groups, processes for monitoring risk, and clear protection procedures to minimise risks. Nevertheless, with ever-advancing technologies, a comprehensive consent model suitable for all circumstances will be hard to design.

Study participants and patients whose samples are being used for phylogenetic analysis should ideally have consented to such use. However, sequence data generated from drug resistance testing and other surveillance data typically do not include explicit consent to participate in large-scale phylogenetics analyses. When using data from previous studies, researchers must be aware that only broad consent for HIV-related research might have been obtained. In such situations, a waiver of specific consent might be obtainable from an ethics committee. Waivers of specific consent are allowable when samples are no longer linked to identifiers, or when consent was given for sample collection for research and storage in future studies, without specific consent for the current research.

Independent review of protocols for phylogenetic studies is also essential for the protection of research participants. The role of local ethics committees is essential for providing local, independent representation for research participants and others affected by the research, as well as ensuring that the local context in which researchers and participants are situated is taken into account.

### Community engagement

Community engagement should occur early in the research design process, ensuring that phylogenetic research is relevant to participating communities and that local perspectives are included in the design and overall conduct of research studies.<sup>67,68</sup> Meaningful community engagement is particularly challenging in research-naïve and low-income communities, and in criminalised or socially marginalised populations. Lack of authentic representative structures, poor literacy, or poverty place these communities at risk of being exploited,<sup>69,70</sup> especially when research involves highly technical elements, such as viral genomics.

Nevertheless, these challenges should not limit attempts to maximise engagement. The phylogenetics study team of the PopART study<sup>71</sup> in Zambia has performed extensive community engagement in communities in which the study takes place. The process involved obtaining community input in the design stages, as well as ongoing consultation and the development of a feedback protocol. Community representatives were consulted on the benefits and risks of informing and sharing results with entire communities and on measures of how to avoid stigmatisation of or within communities.

### Communication

Scientific inferences are based on probabilities. Comprehension and communication of uncertainty is key to understanding phylogenetic results; technical complexity or lack of familiarity with methods might easily generate a false sense of accuracy and precision. Researchers doing phylogenetic analyses must ensure that caveats, such as the fact that inferences are always based on probabilities and that methods are based on assumptions, are clearly highlighted in any dissemination, including interviews, publications, oral presentations, and posters. It is important to note that probabilities vary; an assignment of 50% to a transmission event is very different to an assignment of 99%. In both cases the analyst will report uncertainty, but the conclusions drawn by most observers will be different. An ethical framework in an area of rapid technological development should prepare for the possibility that, in some cases at least, probabilistic assignments will probably improve over time.

Mass media campaigns, as well as reporting on social media, television, radio, and in newspapers have been powerful ways to raise awareness about HIV, treatments, and prevention, and to facilitate public health campaigns aiming to change attitudes and behaviours. However, the way the media frames HIV and reports study outcomes can affect both the long-term and short-term success of any campaigns and can generate unintentional consequences, including a lack of trust in health-care services.<sup>69,72,73</sup> Any ambiguous or misleading reporting of phylogenetic studies might reduce frequency of HIV testing, increase scepticism about participating in studies, and make risk groups less likely to access health care. Therefore, education of the media, local health-care personnel, and the community about these studies is essential.

Care must be taken when reporting findings relevant to specific population subgroups, including identifiable geographic areas or population groups that might be stigmatised or targeted by government, police, others in the community, or subject to criminal charges. Researchers will need to consider the potential social harms and political impact of findings before deciding exactly what information should be publicly shared or published.

**Panel 5: Eight considerations for ethically responsible implementation of phylogenetic analyses****1. Careful risk-benefit assessment**

Done before designing, conducting, and reporting phylogenetic analysis. Risk assessment should address risks to individuals and to groups that might be identified in the research.

**2. Protection of the rights and interests of study participants**

Individuals who participate in studies and the social and geographic groups that might be identified in phylogenetic networks need to be protected. Clinically relevant results should be returned to the patient or care provider.

**3. Social and legal context**

An awareness of the social environment, legal environment, human rights violations, and other potential negative consequences is essential. This understanding includes knowledge both of when and how data are subject to subpoena, and of precedent criminal cases. These challenges are specific to the context and the legal, political, and social environments are subject to change. Furthermore, considerations of gender-specific risks and concerns faced by women and key populations should be evaluated, owing to coercive social and legal environments.

**4. Risk mitigation strategies**

Study designs should address risks to individuals and groups and take into account the potential for anonymisation and masking of individual and group identifiers as protective strategies, as well as accounting for scientific needs of the project and its value in informing public health strategies. The technical nature of sequence data collected needs to be considered in terms of potential for relinkage or other harms, and data can be preprocessed to reduce this risk in line with the needs of the study. Research staff should be trained on the risks as well as the importance of anonymity, confidentiality, informed consent, and protection of research participants and communities. Monitoring and redress mechanisms should be established to accompany and respond to misuse of phylogenetic data.

**5. Informed consent and other safeguards**

Study participants and patients whose samples are being used for phylogenetic analysis should have consented to such use. In the absence of such consent, waivers of consent must have been obtained from the appropriate ethics committees. Researchers must ensure that specific populations are protected against non-stigmatisation and non-discrimination.

**6. Community engagement**

The engagement process should be started during the research design process, thereby ensuring that the research is relevant to participating communities, and local perspectives are included in the design and overall conduct of the research studies, including risk assessment, risk mitigation, informed consent, and communication.

**7. Communication**

Expertise is needed to do and to interpret phylogenetic results. The results are usually ambiguous and therefore the uncertainty associated with these methods must be communicated appropriately during dissemination to the wider scientific community, government bodies, media, and participating communities. Specific efforts are needed to sensitise public health officials, the police, and communities on the use of phylogenetic analysis in the context of public health, including its benefits and limitations.

**8. Equitable data sharing**

Accountability of phylogenetic and sequencing data should be ensured and we recommend that a governance plan is created to address confidentiality, safety, and potential unauthorised use of phylogenetic data. Information and protocols for data sharing, including controlled access, must be addressed in the governance plan. Only certain information should be routinely published with each sequence, and care is needed to ensure human DNA sequences are not inadvertently released with next-generation sequencing data.

**Equitable data sharing**

Largely as a result of funders' requirements, many anonymised HIV sequences are being made publicly available on GenBank and LosAlamos. This availability is advantageous for some research studies, such as vaccine development. However, the lack of awareness that every sequence is associated with a patient or study participant is a real risk. Care must be taken to ensure human DNA sequences are not inadvertently released with NGS data. Routinely publishing only certain information (eg, the year of sampling and the country where the samples are collected) with each sequence would help to minimise risk. Any other anonymised information should be provided by use of a controlled access protocol, which ensures that the research proposed is scientifically valid, does not pose any risks to study participants, and is in line with the informed consent obtained. This protocol would require the development of a clear governance plan.

Finally, different participant information sheets and consent forms can allow for different amounts of data sharing, and laws can differ as to how data can be reused. Any phylogenetic researcher must abide by the amount of sharing outlined in the forms, even if this impacts on the quality of the research conducted.

**Conclusions and recommendations**

Phylogenetic analysis, either alone or in combination with linked epidemiological data, is a powerful method with the potential to help reduce the spread of the HIV epidemic. However, an effective and sustainable model of good ethical practice in phylogenetic research is required to help minimise the risks to individuals or groups participating in studies while optimising the scientific benefits. Although a one-model approach to address any ethical issues is impractical, given the vast variations in studies and contexts, this Review highlights

For **GenBank** see <https://www.ncbi.nlm.nih.gov/genbank/>

For **LosAlamos HIV sequence database** see <https://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html>



### Search strategy and selection criteria

An outline document was created before the workshop, which included the PubMed search terms “HIV transmission”, “HIV and viral linkage”, “HIV and genetic linkage”, “HIV & phylogenetic”, “sequencing technologies”, “Stigma and HIV”, “Stigma and MSM”, “HIV criminalisation”, “MSM, HIV and criminalisation”, “HIV transmission laws/criminalisation”, and “HIV and female sex workers”. Additional information was obtained by reviewing resources available from UNAIDS and from human rights associations such as the International Lesbian Gay Bisexual Trans and Intersex Association. The outline document was used to draft an agenda, identify a list of experts to be invited to the workshop and to debrief all workshop attendees. Workshop attendees included social scientists, bioethicists, phylogeneticists, clinicians, virologists, lawyers, human rights advocates, HIV activists, and community engagement members. All speakers were asked to provide an abstract and give a presentation on the key issues that, in line with their field of expertise, needed to be considered for HIV phylogenetic analysis. The issues were subsequently debated and a consensus was formed. This Review was drafted on the basis of meeting minutes, abstracts submitted before the meeting, and follow-on email discussions. Reference suggestions from experts were included in this Review.

several themes we believe are essential to consider to do phylogenetic studies in an ethically responsible manner.

We have clustered the key issues into eight domains, which provide a framework through which to consider them: risk and benefit assessments; protection of the rights and interests of study participants; local social and legal context, including human rights violations; risk mitigation strategies to protect individual and group identities; valid informed consent and other safeguards; community engagement; communication; and equitable data sharing. We have also provided recommendations for each domain, based on our review of the published literature and input from experts (panel 5).

Viral sequencing has effectively been restricted to large, highly regulated laboratories. Emerging DNA sequencing technologies are more powerful in terms of their applications to epidemiology,<sup>21</sup> more portable, more robust to field conditions, and have lower capital costs.<sup>74</sup> Therefore, HIV phylogenetics (and increasingly for other pathogens) might become more widely distributed across geographic areas and laboratory types, with widely differing regulatory frameworks. Furthermore, because phylogenetics is inherently relational, data are likely to be shared among wider and looser networks of investigators. These developments have the potential to promote rapid scientific advances, and they also pose new challenges for governance, enhancing the use of disseminating clear ethical frameworks, and promoting positive social norms.

Any researcher doing phylogenetic analysis should be aware of the risks such analyses pose and take steps to mitigate these risks. These steps are particularly pertinent in LMICs, which often have weak governance structures and few laws to protect vulnerable populations. These issues are likely to become more problematic as sequence costs decrease and data become more routinely available. Looking forward, real-time phylogenetics could be used more frequently to direct public health responses and increasingly form the basis for surveillance programmes. Whatever the scenario, the fundamental principle of protecting participating individuals and groups must be central to the design and implementation of any study and the reporting of results.

### Contributors

CEMC drafted the original manuscript. AH, MP, LD, GG, GHe, GHa, MKG, CF, MSC, and DP made substantial contributions to the manuscript. OL produced the content for figure 2. JDT and team produced the infographics. The content of the manuscript was based on discussions at the Ethics in HIV Phylogenetics meeting and excerpts of meeting abstracts (JJA, MS, PE, VN, and A-MV) have been incorporated into the manuscript. Members of the Ethics in HIV Phylogenetics Working Group (CEMC, AH, GG, OL, JDT, JS, GD, PG-F, MSC, and DP) organised the meeting. AH led the working group with support from MSC and DP, under the auspices of the PANGEA-HIV Consortium (principal investigator DP). All authors and working group members reviewed and approved the final version of the manuscript.

### Declaration of interests

This work was funded by the Bill & Melinda Gates Foundation and done with support from the HIV Prevention Trials Network and the PANGEA-HIV project (grant code OPP1084362). CEMC is funded by the Wellcome Trust (grant code 106551/Z/14/A). OL received additional support from the Division of Intramural Research National Institute of Allergy and Infectious Diseases. A-MV received funding from Fonds voor Wetenschappelijk Onderzoek—Flanders (grant G.0692.14), and the VIROGENESIS project receives funding from the European Union's Horizon 2020 Research and Innovation Programme (under grant agreement 634650). GD is employed by the Bill & Melinda Gates Foundation. LD is employed by the US National Institutes of Health. DP is funded by the Wellcome Trust. No other funding sources or representatives thereof had any role in the writing of the manuscript or the decision to submit it for publication. None of the authors have been paid to write this Review by a pharmaceutical company or other agency. The views expressed in this Review do not represent any policy or position of the US Department of Health and Human Services or any of its components. All other authors declare no competing interests.

### Acknowledgments

We would like to thank Shufang Wei for producing the infographics for this paper, Habiba Cooper Diallo and Cait Breeden for their assistance in organising the Ethics in HIV Phylogenetics meeting, Anne Johnson for her helpful discussions and commenting on the manuscript, and the PANGEA-HIV Steering Committee for their support.

### References

- UNAIDS. AIDS by numbers 2015. Geneva: Joint United Nations Programme on HIV/AIDS, 2015.
- Dennis AM, Herbeck JT, Leigh Brown A, et al. Phylogenetic studies of transmission dynamics in generalized HIV epidemics: an essential tool where the burden is greatest? *J Acquir Immune Defic Syndr* 2014; **67**: 181–95.
- Grabowski MK, Lessler J, Redd AD, et al. The role of viral introductions in sustaining community-based HIV epidemics in rural Uganda: evidence from spatial clustering, phylogenetics, and egocentric transmission models. *PLoS Med* 2014; **11**: e1001610.
- de Oliveira T, Kharsany ABM, Gräf T, et al. Transmission networks and risk of HIV infection in KwaZulu-Natal, South Africa: a community-wide phylogenetic study. *Lancet HIV* 2017; **4**: e41–50.

- 5 Pillay D, Herbeck JT, Cohen MS, et al. PANGEA-HIV: phylogenetics for generalised epidemics in Africa. *Lancet Infect Dis* 2015; **15**: 259–61.
- 6 World Medical Association. WMA Declaration of Helsinki—ethical principles for medical research involving human subjects. 2013. <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/> (accessed March 28, 2018).
- 7 Council for International Organizations of Medical Sciences, WHO. International Ethical Guidelines for health-related research involving humans. 2016. <https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf> (accessed March 28, 2018).
- 8 Geller G, Dvoskin R, Thio CL, et al. Genomics and infectious disease: a call to identify the ethical, legal and social implications for public health and clinical practice. *Genome Med* 2014; **6**: 106.
- 9 Kaye J, Meslin EM, Knoppers BM, et al. Research priorities. ELSI 2.0 for genomics and society. *Science* 2012; **336**: 673–74.
- 10 Parker M, Kwiatkowski DP. The ethics of sustainable genomic research in Africa. *Genome Biol* 2016; **17**: 44.
- 11 De Vries J, Munung SN, Matimba A, et al. Regulation of genomic and biobanking research in Africa: a content analysis of ethics guidelines, policies and procedures from 22 African countries. *BMC Med Ethics* 2017; **18**: 8.
- 12 The Wellcome Trust. Ethical sharing of health research data in low- and middle-income countries. 2014. <https://wellcome.ac.uk/sites/default/files/ethical-sharing-of-health-research-data-in-low-and-middle-income-countries-phrfd-2014.pdf> (accessed March 28, 2018).
- 13 H3Africa. High-level principles on ethics, governance and resource sharing. 2013. <https://h3africa.org/about/ethics-and-governance> (accessed March 28, 2018).
- 14 De Vries J, Tindana P, Littler K, et al. The H3Africa policy framework: negotiating fairness in genomics. *Trends Genet* 2015; **31**: 117–19.
- 15 Cohen J. Pinpointing HIV spread in Africa poses risks. *Science* 2017; **356**: 568–69.
- 16 Batorsky R, Sergeev RA, Rouzine IM. The route of HIV escape from immune response targeting multiple sites is determined by the cost-benefit trade off of escape mutations. *PLoS Comput Biol* 2014; **10**: e1003878.
- 17 Grabowski MK, Redd AD. Molecular tools for studying HIV transmission in sexual networks. *Curr Opin HIV AIDS* 2014; **9**: 126–33.
- 18 Ratmann O, van Sighem A, Bezemer D, et al. Sources of HIV infection among men having sex with men and implications for prevention. *Sci Transl Med* 2016; **8**: 320ra2.
- 19 Volz EM, Frost SD. Inferring the source of transmission with phylogenetic data. *PLoS Comput Biol* 2013; **9**: e1003397.
- 20 Romero-Severson EO, Bulla I, Leitner T. Phylogenetically resolving epidemiologic linkage. *Proc Natl Acad Sci USA* 2016; **113**: 2690–95.
- 21 Wymant C, Hall M, Ratmann O, et al. PHYLOSCANNER: inferring transmission from within- and between-host pathogen genetic diversity. *Mol Biol Evol* 2017; **35**: 719–33.
- 22 Noguera-Julian M, Edgil D, Harrigan PR, Sandstrom P, Godfrey C, Paredes R. Next-generation human immunodeficiency virus sequencing for patient management and drug resistance surveillance. *J Infect Dis* 2017; **16** (suppl 9): S829–33.
- 23 Moscona R, Ram D, Wax M, et al. Comparison between next-generation and Sanger-based sequencing for the detection of transmitted drug-resistance mutations among recently infected HIV-1 patients in Israel, 2000–2014. *J Int AIDS Soc* 2017; **20**: 21846.
- 24 Hue S, Clewley JP, Cane PA, Pillay D. Investigation of HIV-1 transmission events by phylogenetic methods: requirement for scientific rigour. *AIDS* 2005; **19**: 449–50.
- 25 Brown AE, Gifford RJ, Clewley JP, et al. Phylogenetic reconstruction of transmission events from individuals with acute HIV infection: toward more-rigorous epidemiological definitions. *J Infect Dis* 2009; **199**: 427–31.
- 26 Vandamme A-M. Basic concepts of molecular evolution. In: Lemey P, Salemi M, Vandamme A-M, eds. *The Phylogenetic Handbook*. Cambridge: Cambridge University Press, 2009: 3–30.
- 27 Efron B, Halloran E, Holmes S. Bootstrap confidence levels for phylogenetic trees. *Proc Natl Acad Sci USA* 1996; **93**: 7085–90.
- 28 Ratmann O, Hodcroft EB, Pickles M, et al. Phylogenetic tools for generalized HIV-1 epidemics: findings from the PANGEA-HIV methods comparison. *Mol Biol Evol* 2017; **34**: 185–203.
- 29 Volz EM, Ionides E, Romero-Severson EO, Brandt MG, Mokotoff E, Koopman JS. HIV-1 transmission during early infection in men who have sex with men: a phylodynamic analysis. *PLoS Med* 2013; **10**: e1001568.
- 30 Stadler T, Kuhner D, Bonhoeffer S, Drummond AJ. Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proc Natl Acad Sci USA* 2013; **110**: 228–33.
- 31 Faria NR, Azevedo Rdo S, Kraemer MU, et al. Zika virus in the Americas: early epidemiological and genetic findings. *Science* 2016; **352**: 345–49.
- 32 Pozniak A, Gazzard B, Anderson J, et al. British HIV Association (BHIVA) guidelines for the treatment of HIV-infected adults with antiretroviral therapy. *HIV Med* 2003; **4** (suppl 1): 1–41.
- 33 Dunn D, Pillay D. UK HIV drug resistance database: background and recent outputs. *J HIV Ther* 2007; **12**: 97–98.
- 34 Schoeni-Affolter F, Ledergerber B, Rickenbach M, et al. Cohort profile: the Swiss HIV Cohort study. *Int J Epidemiol* 2010; **39**: 1179–89.
- 35 Hughes GJ, Fearnhill E, Dunn D, Lycett SJ, Rambaut A, Leigh Brown AJ. Molecular phylogenetics of the heterosexual HIV epidemic in the United Kingdom. *PLoS Pathog* 2009; **5**: e1000590.
- 36 Poon AFY, Gustafson R, Daly P, et al. Near real-time monitoring of HIV transmission hotspots from routine HIV genotyping: an implementation case study. *Lancet HIV* 2016; **3**: e231–38.
- 37 Wertheim JO, Kosakovsky Pond SL, Forgiione LA, et al. Social and genetic networks of HIV-1 transmission in New York City. *PLoS Pathog* 2017; **13**: e1006000.
- 38 WHO, United States Centers for Disease Control and Prevention, The Global Fund to Fight AIDS, Tuberculosis, and Malaria. HIV drug resistance report 2017. 2017. <http://www.who.int/hiv/pub/drugresistance/hivdr-report-2017/en/> (accessed March 28, 2018).
- 39 Ou CY, Ciesielski CA, Myers G, et al. Molecular epidemiology of HIV transmission in a dental practice. *Science* 1992; **256**: 1165–71.
- 40 Lemey P, Van Dooren S, Van Laethem K, et al. Molecular testing of multiple HIV-1 transmissions in a criminal case. *AIDS* 2005; **19**: 1649–58.
- 41 Bernard EJ, Azad Y, Vandamme AM, Weait M, Geretti AM. HIV forensics: pitfalls and acceptable standards in the use of phylogenetic analysis as evidence in criminal investigations of HIV transmission. *HIV Med* 2007; **8**: 382–87.
- 42 Abecasis AB, Pingarilho M, Vandamme AM. Phylogenetic analysis as a forensic tool in HIV transmission investigations. *AIDS* 2018; **32**: 543–54.
- 43 Carlson JM, Du VY, Pfeifer N, et al. Impact of pre-adapted HIV transmission. *Nat Med* 2016; **22**: 606–13.
- 44 McGrath N, Eaton JW, Newell M-L, Hosegood V. Migration, sexual behaviour, and HIV risk: a general population cohort in rural South Africa. *Lancet HIV* 2015; **2**: e252–59.
- 45 Dobra A, Barnighausen T, Vandormael A, Tanser F. Space-time migration patterns and risk of HIV acquisition in rural South Africa. *AIDS* 2017; **31**: 137–45.
- 46 SASPEN. Migration and health care: the case of HIV and AIDS in Botswana. 2015. [http://www.saspen.org/brief/SASPEN-brief\\_2015-6\\_Refilwe-Sinkamba\\_Migration-and-Healthcare-HIV-Botswana.pdf](http://www.saspen.org/brief/SASPEN-brief_2015-6_Refilwe-Sinkamba_Migration-and-Healthcare-HIV-Botswana.pdf) (accessed March 28, 2018).
- 47 Singh R, Dwivedi VK. Population and housing census 2011 analytical report. Gaborone, Botswana: Statistics Botswana, 2014. <https://www.statsbots.org.bw/sites/default/files/publications/Population%20%26%20Housing%20Census%20Dissemination%20analytical%20report%20.pdf> (accessed March 28, 2018).
- 48 Jacques G, Mmatli TO. Addressing ethical non-sequiturs in Botswana's HIV and AIDS policies: harmonising the halo effect. *Ethics Social Welf* 2013; **7**: 342–58.
- 49 Amon JJ, Kasambala T. Structural barriers and human rights related to HIV prevention and treatment in Zimbabwe. *Glob Public Health* 2009; **4**: 528–45.
- 50 Cohen MS, Chen YQ, McCauley M, et al. Antiretroviral therapy for the prevention of HIV-1 transmission. *N Engl J Med* 2016; **375**: 830–39.
- 51 Cohen MS, McCauley M, Sugarman J. The ethical odyssey in testing HIV treatment as prevention. *Clin Trials* 2012; **9**: 340–47.
- 52 Munthe C. The goals of public health: an integrated, multidimensional model. *Public Health Ethics* 2008; **1**: 39–52.

- 53 Sox HC. Resolving the tension between population health and individual health care. *JAMA* 2013; **310**: 1933–34.
- 54 Gostin LO, Bayer R, Fairchild AL. Ethical and legal challenges posed by severe acute respiratory syndrome: implications for the control of severe infectious disease threats. *JAMA* 2003; **290**: 3229–37.
- 55 Earley CL, Strong LC. Certificates of confidentiality: a valuable tool for protecting genetic data. *Am J Hum Genet* 1995; **57**: 727–31.
- 56 WHO. Human rights and health. 2017. <http://www.who.int/mediacentre/factsheets/fs323/en/> (accessed March 28, 2018).
- 57 Shannon K, Strathdee SA, Goldenberg SM, et al. Global epidemiology of HIV among female sex workers: influence of structural determinants. *Lancet* 2015; **385**: 55–71.
- 58 de Oliveira T, Pybus OG, Rambaut A, et al. Molecular epidemiology: HIV-1 and HCV sequences from Libyan outbreak. *Nature* 2006; **444**: 836–37.
- 59 Human Rights Watch. “Tell me where I can be safe.” The impact of Nigeria’s same sex marriage (prohibition) act. 2016. [https://www.hrw.org/sites/default/files/report\\_pdf/nigeria1016\\_web.pdf](https://www.hrw.org/sites/default/files/report_pdf/nigeria1016_web.pdf) (accessed March 28, 2018).
- 60 Bernard EJ, Cameron S. Advancing HIV justice 2. Building momentum in global advocacy against HIV criminalisation. Brighton and Amsterdam: HIV Justice Network and Global Network of People Living with HIV. 2016. [http://www.hivjustice.net/wp-content/uploads/2016/05/AHJ2.final2\\_10May2016.pdf](http://www.hivjustice.net/wp-content/uploads/2016/05/AHJ2.final2_10May2016.pdf) (accessed July 31, 2018).
- 61 O’Byrne P, Willmore J, Bryan A, et al. Nondisclosure prosecutions and population health outcomes: examining HIV testing, HIV diagnoses, and the attitudes of men who have sex with men following nondisclosure prosecution media releases in Ottawa, Canada. *BMC Public Health* 2013; **13**: 94.
- 62 Abecasis AB, Geretti AM, Albert J, Power L, Weait M, Vandamme A-M. Science in court: the myth of HIV fingerprinting. *Lancet Infect Dis* 2011; **11**: 78–79.
- 63 Leitner T. Guidelines for HIV in court cases. *Nature* 2011; **473**: 284.
- 64 Galletly CL, Pinkerton SD. Conflicting messages: how criminal HIV disclosure laws undermine public health efforts to control the spread of HIV. *AIDS Behav* 2006; **10**: 451–61.
- 65 O’Byrne P, Bryan A, Woodyatt C. Nondisclosure prosecutions and HIV prevention: results from an Ottawa-based gay men’s sex survey. *JANAC* 2013; **24**: 81–87.
- 66 Department of Health, Education, and Welfare; National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report. Ethical principles and guidelines for the protection of human subjects of research. *J Am Coll Dent* 2014; **81**: 4–13.
- 67 UNAIDS. Good participatory practice: guidelines for HIV biomedical prevention trials. Geneva: Joint United Nations Programme on HIV/AIDS, 2011. [http://www.unaids.org/sites/default/files/media\\_asset/JC1853\\_GPP\\_Guidelines\\_2011\\_en\\_0.pdf](http://www.unaids.org/sites/default/files/media_asset/JC1853_GPP_Guidelines_2011_en_0.pdf) (accessed March 28, 2018).
- 68 WHO. Ethical considerations in biomedical HIV prevention trials: UNAIDS-WHO guidance document. Geneva: Joint United Nations Programme on HIV/AIDS and World Health Organization, 2012. [http://www.unaids.org/en/resources/documents/2012/20120701\\_jc1399\\_ethical\\_considerations](http://www.unaids.org/en/resources/documents/2012/20120701_jc1399_ethical_considerations) (accessed March 28, 2018).
- 69 Singh JA, Mills EJ. The abandoned trials of pre-exposure prophylaxis for HIV: what went wrong? *PLoS Med* 2005; **2**: e234.
- 70 Vermund SH, Blevins M, Moon TD, et al. Poor clinical outcomes for HIV infected children on antiretroviral therapy in rural Mozambique: need for program quality improvement and community engagement. *PLoS One* 2014; **9**: e110116.
- 71 London School of Hygiene & Tropical Medicine. PopART. 2017. <https://www.lshtm.ac.uk/popart> (accessed March 28, 2018).
- 72 Boyson AR, Zimmerman RS, Shoemaker S. Exemplification of HAART and HIV/AIDS: a news experiment. *Health Commun* 2015; **30**: 901–10.
- 73 Zimmerman RS, Kirschbaum AL. News of biomedical advances in HIV: relationship to treatment optimism and expected risk behavior in US MSM. *AIDS Behav* 2017; **22**: 367–78.
- 74 Quick J, Loman NJ, Duraffour S, et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature* 2016; **530**: 228–32.

© 2018 Elsevier Ltd. All rights reserved.

# THE LANCET HIV

## Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Coltart CEM, Hoppe A, Parker M, et al, on behalf of the Ethics in HIV Phylogenetics Working Group. Ethical considerations in global HIV phylogenetic research. *Lancet HIV* 2018; published online August 30. [http://dx.doi.org/10.1016/S2352-3018\(18\)30134-6](http://dx.doi.org/10.1016/S2352-3018(18)30134-6).



## **SUPPLEMENTARY MATERIALS**

### **Ethics of HIV Phylogenetics Working Group Members:**

Africa Health Research Institute, KwaZulu-Natal, ZA (Guy Harling, ScD; Prof Deenan Pillay, PhD; Prof Janet Seeley, PhD);

Bill & Melinda Gates Foundation, Washington, DC, US (Gina Dallabetta, MD);

Centers for Disease Control and Prevention, Atlanta, GA, US (Irene Hall, PhD);

European AIDS Treatment Group, Brussels, BE (Nikos Dedes)

Harvard T.H. Chan School of Public Health, Boston, MA, US (Guy Harling, ScD; Vladimir Novitsky, PhD);

HIV Justice Network, Brighton, UK (Edwin J Bernard, BA (Hons));

Johns Hopkins University, Baltimore, MD, US (Prof Gail Geller, ScD; Danielle German, PhD; M. Kate Grabrowski, PhD);

Joint United Nations Programme on HIV/AIDS, Geneva, CH (Patrick M. Eba, PhD; Prof Peter Godfrey-Faussett, FRPC);

London School of Hygiene & Tropical Medicine, London, UK (Prof Peter Godfrey-Faussett, FRCP; Prof Janet Seeley, PhD);

National Institutes of Health, Bethesda, MD (Prof David Burns, MD; Liza Dawson, PhD; Oliver Laeyendecker, PhD);

Princeton University, Princeton, NJ, US (Dr Joseph J. Amon, PhD)

Public Health England, London, UK (Valerie Delpech, FPHM)

Treatment Advocacy and Literacy Campaign, Lusaka, ZM (Felix Mwanza, BA);

Universidade Nova de Lisboa, Lisbon, PT (Prof Anne-Mieke Vandamme, PhD);

University California, San Diego, La Jolla, CA, US (Joel O. Wertheim, PhD);

University College London, London, UK (Cordelia C. Coltart, PhD; Guy Harling, ScD; Anne Hoppe, PhD;

Zisis Kozlakidis, PhD; Prof Deenan Pillay, PhD)

University of KwaZulu-Natal, Pietermaritzburg, ZA (Patrick M. Eba, PhD);

University of Leuven, Leuven, BE (Prof Anne-Mieke Vandamme, PhD);

University of Missouri, St Louis, MO, US (Prof Rick Zimmerman, PhD);

University of North Carolina, Chapel Hill, NC, US (Prof Myron S. Cohen, MD; Prof Gail Henderson;  
PhD; Joseph D. Tucker, PhD);

University of Oxford, Oxford, UK (Prof Christophe Fraser, PhD; Prof Michael Parker, PhD)

World Health Organization, Geneva, CH (Rachel Baggaley, MBBS; Andreas Reis, MD);

Zambart, Lusaka, ZM (Musonda Simwinga, PhD)

**Supplementary Table 1: Summary of key documents, position statements and initiatives relevant to ethical issues of HIV phylogenetics and referred to within the document**

<b>Document/ Position statement/ Initiative</b>	<b>Description</b>
The Declaration of Helsinki (first published in 1964; amended most recently in 2013)(6)	The World Medical Association developed the Helsinki declaration as a statement of ethical principles for medical research involving human subjects, including research on identifiable human material and data. This declaration states that the interest and well-being of the individual takes precedence over the science and well-being of communities and populations. Many of the principles are relevant to performing HIV phylogenetic studies.
The Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines (7)	CIOMS is an international nongovernmental organization in official relationship with WHO, founded in 1949. The guidelines aim to provide internationally vetted ethical principles and detailed commentary on how universal ethical principles should be applied, with particular attention to conducting research in lower-income countries (LIC). There have been four revisions of the guidelines since they were first published (1982) to take into account scientific developments and bring the guidelines into line with current thinking on ethics and human rights.
Wellcome Trust report on Ethical	This report aims to provide evidence to inform the

<p>sharing of health research data in LMIC: views of stakeholders (12)</p>	<p>development, implementation and evaluation of data-sharing models and identify further research priorities. It is based on a multi-site collaborative study of stakeholder experiences and views in LMIC of best practices in sharing individual-level data from clinical and public health research.</p>
<p>The Human Heredity and Health in Africa (H3Africa) Initiative (13)</p>	<p>H3Africa Initiative aims to facilitate a contemporary research approach to the study of genomics and environmental determinants of common diseases with the goal of improving the health of African populations. To accomplish this, the H3Africa Initiative aims to contribute to the development of the necessary expertise among African scientists, and to establish networks of African investigators.</p>
<p>The ELSI (Ethical, Legal, and Social Implications) Program (8, 9)</p>	<p>The ELSI Program is a multi-disciplinary program funded by the National Human Genome Research Institute at NIH. It focuses on exploring ELSI of human genomics, and developing policy options to address these implications, although the scope has broadened over years in response to rapidly evolving genomic technologies, legal and commercial developments, and translation to clinical applications. Many of the issues from human genomics also apply to viral genomics: Biobank governance is a particular focus of the “ELSI 2.0” initiative (9).</p>



**Supplementary Panel 1: Key methodological concepts to be considered when inferring HIV molecular transmission clusters**

- Phylogenetic support methods use bootstrap or posterior probability to identify groups more closely related to each other than to the rest of the population being analysed:
  - Bootstrapping: a statistical resampling method of random sampling of nucleotide sites with replacement. This process is repeated multiple times and the frequency of identical branch reproduction gives a bootstrap value indicating the robustness of the cluster assignment.
  - Posterior probability (PP): combines the prior probability of a tree with the likelihood of the given data to indicate the probability of the cluster assignment to be correct.
- Phylogenetic distance methods identify groups whose mean/median/maximum genetic distance suggests a common ancestor in recent time. Phylogenetic support and distance methods are often combined.
- Molecular clock methods indicate the timing of the most recent common ancestor, which can contribute to understanding the timing of infection.

**Supplementary Panel 2: Important social and legal considerations**

- Are there particular approaches to handling reporting of results that will reduce risk?
- What role can local and external ethics committees play in addressing risks and handling the potential for political issues that arise in the local context?
- What kinds of discussions with policy makers, government officials or other stakeholders might be helpful in planning the research and communicating findings?
- What on-going monitoring will be conducted to ensure respect for study participants and impact on people with HIV or key populations? What resources are available for advocacy and redress if concerns arise?
- Have groups of people with HIV and key populations been meaningfully consulted? Are their views and concerns taken into account?

**Supplementary Panel 3: Five Key questions that need to be addressed as part of responsible and ethical community engagement in phylogenetic research**

1. What is the best community engagement strategy for phylogenetic studies and how sustainable is it?
2. How should be feedback provided to communities? Does informing entire communities add value and/or pose risks?
3. How valid is informed consent when a phylogenetic study is nested within an existing study or healthcare setting? Does it become a question of trust?
4. How can we avoid stigmatization when public health interventions are tailored towards specific communities?
5. How can researchers best share results at community level? Is it disrespectful not to do so?