Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis

*Joris Hemelaar, Ramyiadarsini Elangovan, Jason Yun, Leslie Dickson-Tetteh, Isabella Fleminger, Shona Kirtley, Brian Williams, Eleanor Gouws-Williams, Peter D Ghys, on behalf of the WHO–UNAIDS Network for HIV Isolation and Characterisation**

Summary

Background Global genetic diversity of HIV-1 is a major challenge to the development of HIV vaccines. We aimed to estimate the regional and global distribution of HIV-1 subtypes and recombinants during 1990–2015.

Methods We searched PubMed, EMBASE (Ovid), CINAHL (Ebscohost), and Global Health (Ovid) for HIV-1 subtyping studies published between Jan 1, 1990, and Dec 31, 2015. We collected additional unpublished HIV-1 subtyping data through a global survey. We included prevalence studies with HIV-1 subtyping data collected during 1990–2015. We grouped countries into 14 regions and analysed data for four time periods (1990–99, 2000–04, 2005–09, and 2010–15). The distribution of HIV-1 subtypes, circulating recombinant forms (CRFs), and unique recombinant forms (URFs) in individual countries was weighted according to the UNAIDS estimates of the number of people living with HIV (PLHIV) in each country to generate regional and global estimates of HIV-1 diversity in each time period. The primary outcome was the number of samples designated as HIV-1 subtypes A, B, C, D, F, G, H, J, K, CRFs, and URFs. The systematic review is registered with PROSPERO, number CRD42017067164.

Findings This systematic review and global survey yielded 2203 datasets with 383 519 samples from 116 countries in 1990–2015. Globally, subtype C accounted for 46·6% (16 280 897/34 921 639 of PLHIV) of all HIV-1 infections in 2010–15. Subtype B was responsible for 12·1% (4 235 299/34 921 639) of infections, followed by subtype A (10·3%; 3 587 003/34 921 639), CRF02_AG (7·7%; 2 705 110/34 921 639), CRF01_AE (5·3%; 1 840 982/34 921 639), subtype G (4·6%; 1 591 276/34 921 639), and subtype D (2·7%; 926 255/34 921 639). Subtypes F, H, J, and K combined accounted for 0·9% (311 332/34 921 639) of infections. Other CRFs accounted for 3·7% (1 309 082/34 921 639), bringing the proportion of all CRFs to 16·7% (5 844 113/34 921 639). URFs constituted 6·1% (2 134 405/34 921 639), resulting in recombinants accounting for 22·8% (7 978 517/34 921 639) of all global HIV-1 infections. The distribution of HIV-1 subtypes and recombinants changed over time in countries, regions, and globally. At a global level during 2005–15, subtype B increased, subtypes A and D were stable, and subtypes C and G and CRF02_AG decreased. CRF01_AE, other CRFs, and URFs increased, leading to a consistent increase in the global proportion of recombinants over time.

Interpretation Global and regional HIV diversity is complex and evolving, and is a major challenge to HIV vaccine development. Surveillance of the global molecular epidemiology of HIV-1 remains crucial for the design, testing, and implementation of HIV vaccines.

Funding None.

Introduction

The HIV pandemic remains a major global public health problem, with 36·9 million people living with HIV in 2017.[1] Despite the expansion of antiretroviral treatment programmes, in 2017, 940 000 people died from AIDS-related illnesses and 1·8 million people became newly infected with HIV.[1] An HIV vaccine will probably be necessary to end the HIV pandemic.[2]

One of the key characteristics of the HIV pandemic is its extraordinary global genetic diversity.[3] After zoonotic transmission from non-human primates to humans at the beginning of the 20th century, HIV-1 group M diversified into distinct subtypes, designated by the letters A, B, C, D, F, G, H, J, and K. Recombinants between subtypes, designated as circulating recombinant forms (CRFs) or unique recombinant forms (URFs), also formed.[4] CRFs are strains propagated in the population and are named consecutively according to internationally agreed guidelines (96 distinct CRFs have been identified to date). URFs refer to unique recombinant sequences without evidence of onward transmission.[4] The spread and evolution of HIV has caused differential distributions of HIV-1 subtypes, CRFs, and URFs worldwide.[3]

Global HIV-1 genetic diversity is a major obstacle to the development of a vaccine against HIV, because a globally effective HIV vaccine will need to protect against divergent HIV subtypes and recombinants.[5] The design, testing, and implementation of HIV

Lancet Infect Dis 2019; 19: 143–55

Published **Online** November 30, 2018 http://dx.doi.org/10.1016/S1473-3099(18)30647-9

This online publication has been corrected. The corrected version first appeared at thelancet.com/infection on December 23, 2019

See **Comment** page 114

*Members listed at the end of the Article

Nuffield Department of Women's & Reproductive Health, Women's Centre, John Radcliffe Hospital, University of Oxford, Oxford, UK (J Hemelaar DPhil, R Elangovan BMBCh, J Yun BA, L Dickson-Tetteh BA, I Fleminger MSci)**; Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, University of Oxford, Oxford, UK** (S Kirtley MSc)**; South African Centre for Epidemiological Modelling and Analysis, Stellenbosch University, Stellenbosch, South Africa** (B Williams PhD)**; and Fast-Track Implementation Department,** (E Gouws-Williams PhD) **and Strategic Information Department** (P D Ghys PhD)**, UNAIDS, Geneva, Switzerland**

Correspondence to: Dr Joris Hemelaar, Nuffield Department of Women's & Reproductive Health, Women's Centre, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK **joris.hemelaar@wrh.ox.ac.uk**

For more on **CRFs** see https://www.hiv.lanl.gov

Research in context

Evidence before this study

Global HIV diversity is a major challenge to the development of vaccines against HIV. Designing, testing, and implementing HIV vaccines requires up-to-date and accurate knowledge of the distribution of HIV-1 subtypes and recombinants in countries, regions, and worldwide. Surveillance of the global molecular epidemiology of HIV-1 is therefore essential. We searched PubMed, Embase (Ovid), CINAHL (Ebscohost), and Global Health (Ovid) for studies reporting HIV-1 subtyping data published between Jan 1, 1990, and Dec 31, 2015. Search terms were Mesh headings, Emtree terms, and free text words and synonyms, including "HIV", "subtype", "recombinant", and "epidemiology". The appendix (pp 2–5) contains a full list of all search terms. No methodological or language filters were used. We found that data and information on trends of global molecular epidemiology of HIV-1 are lacking. HIV sequence databases contain sequences with known dates and countries of sample collection, but samples are not representative of populations. Estimates based only on published data have restricted coverage, especially for recent years, and are prone to publication bias. Estimates based on both published and unpublished data and adjusted for the number of people living with HIV (PLHIV) in each country are only available for 2000–07.

Added value of this study

To our knowledge, this is the largest study to date reporting on the regional and global distribution of HIV-1 subtypes and recombinants, by combining data from a systematic literature review and a global survey of experts in molecular epidemiology of HIV-1 who were members of the WHO-UNAIDS Network for HIV Isolation and Characterisation. The time period (1990–2015) covers the whole period for which reliable estimates of national prevalences of HIV are available. We achieved high coverage both globally and regionally. We collected 2203 datasets with 383 519 samples from 116 countries for 1990–2015. Distinct distributions of HIV-1 subtypes and recombinants were seen in different countries and regions. Globally, subtype C accounted for about half of all HIV-1 infections worldwide in 2010–15. Subtypes B and A were responsible for 12·1% and 10·3% of infections, respectively, followed by circulating recombinant forms (CRFs), CRF02_AG and CRF01_AE, subtype G, and subtype D. Subtypes F, H, J, and K combined accounted for 0·9% of infections globally. Other CRFs accounted for 3·7%, bringing the proportion of all CRFs to 16·7%. Unique recombinant froms (URFs) constituted 6·1%, resulting in recombinants accounting for 22·8% of all global HIV-1 infections. Changes in the distribution of HIV-1 subtypes and recombinants were seen over time in countries, regions, and globally. In 2005–15 at a global level, subtype B increased, subtypes A and D were stable, and subtypes C and G and CRF02_AG decreased. CRF01_AE, other CRFs, and URFs increased, leading to a consistent increase in the global proportion of recombinants over time.

Implications of all the available evidence

Global and regional HIV diversity is complex and evolving and is a major challenge to the development of HIV vaccines. Surveillance of the global molecular epidemiology of HIV-1 remains crucial for the design, testing, and implementation of HIV vaccines. Global HIV diversity also affects HIV diagnostic assays, viral load measurements, development of drug resistance, and response to antiretroviral treatment.

vaccines requires up-to-date and accurate knowledge of the distribution of HIV subtypes and recombinants in different parts of the world, because vaccine immunogen sequences need to match, as closely as possible, the viral sequences circulating in the target population.[6] Furthermore, HIV diversity also affects HIV diagnostic assays, viral load measurements,[7] development of drug resistance, and response to antiretroviral treatment.[8,9]

Surveillance of the global molecular epidemiology of HIV-1 is therefore essential, but recent data and information on trends are insufficient.[10,11] HIV sequence databases contain sequences with known dates and countries of collection, but samples are not representative of populations. Estimates based only on published data have restricted coverage, especially for recent years (due to the delay between sample collection and publication), and are prone to publication bias.[12] In this study, we aimed to estimate the regional and global distribution of HIV-1 subtypes and recombinants between 1990 and 2015, by combining country-specific HIV-1 subtyping data from a systematic review and a global survey with UNAIDS estimates of the number of people living with HIV (PLHIV) in each country for 1990–2015.

See Online for appendix

For more on the **HIV drug resistance database** see https://hivdb.stanford.edu

Methods

Systematic literature review

We searched PubMed, EMBASE (Ovid), CINAHL (Ebscohost), and Global Health (Ovid) for studies reporting HIV-1 subtyping data published between Jan 1, 1990, and Dec 31, 2015. Search terms were Mesh headings, Emtree terms, and free text words and synonyms, including "HIV", "subtype", "recombinant", and "epidemiology". The appendix (pp 2–5) contains a full list of all search terms. No methodological or language filters were used. We combined all search results to form a central database of citations in Endnote reference manager (Endnote X7; Clarivate Analytics, Philadelphia, PA, USA). Duplicate references were removed. Reviewers (RE, JY, LD-T, and JH) screened titles and abstracts of references, retrieved full text articles of potentially eligible studies, and assessed articles against the eligibility criteria (figure 1).

Additional published data sources

We obtained further published data from four sources (appendix pp 6–8): the WHO HIV Drug Resistance Report 2012,[13] references in published reports and reviews on HIV diversity, tables of contents of four specialist journals (issues of *AIDS*, *Journal of AIDS*, *Journal of Virology*, and *AIDS Research and Human Retroviruses* published between January, 1990, and February, 2016), and publications indexed on the Scopus citation database that referenced previous publications on global HIV-1 molecular epidemiology.

Global survey

We collected unpublished original HIV-1 subtyping data through a survey among experts in the field who were members of the WHO–UNAIDS Network for HIV Isolation and Characterisation. We contacted researchers who were known to be working on HIV-1 molecular epidemiology based on previous publications, conference abstracts, or informal networking by email or fax and asked them to contribute unpublished HIV-1 subtyping data that had been collected as part of independent studies by completing a pre-formulated data collection template. The eligibility criteria for unpublished data sources were the same as those applied to published sources.

Eligibility criteria

We included prevalence studies of PLHIV with known country and year of sample collection (1990–2015), original HIV-1 subtyping data, and a minimum of 20 samples in each study. All study types with prevalence data for HIV-1 subtyping were eligible, including randomised controlled trials, cohort studies, and cross-sectional studies. We excluded studies containing only incident infections or new diagnoses. No restrictions were applied in relation to age, sex, ethnicity, nationality, duration of infection, CD4-positive T cell count, viral load, treatment with antiretrovirals or other medications, or co-infections. The country designation of a dataset was determined by the country where the blood samples were taken and not by the country of origin of the participants. Different data sets from each country were permitted to have been collected from different sites and populations.

HIV-1 subtyping methods included sequencing (genotyping), heteroduplex mobility assay, and serotyping (appendix p 10). Any genome segment could be used for subtyping (including *gag*, *pol*, *env*, and other genome segments) or the full-length genome (appendix pp 11, 12). No minimum sequence length was specified and all online subtyping tools were accepted. The subtyping data as provided in each manuscript or submitted dataset were taken as correct. We excluded untyped samples.

Data extraction

Reviewers (RE, JY, LD-T, and JH) extracted the following data from published and unpublished sources: country, city or region, year(s) when samples were collected, study type, population, HIV-1 subtyping method(s), genome segment(s) analysed, and HIV-1 subtyping data (minimum of 20 samples per dataset). The primary outcome was the number of samples designated as HIV-1 subtypes A, B, C, D, F, G, H, J, K, CRFs, and URFs in each dataset. No formal study quality assessments were done. No patient identifiable information was retrieved at any stage and consent was presumed to have been obtained by the researchers who submitted or published each dataset. Any ambiguities and discrepancies were resolved by the senior reviewer (JH).

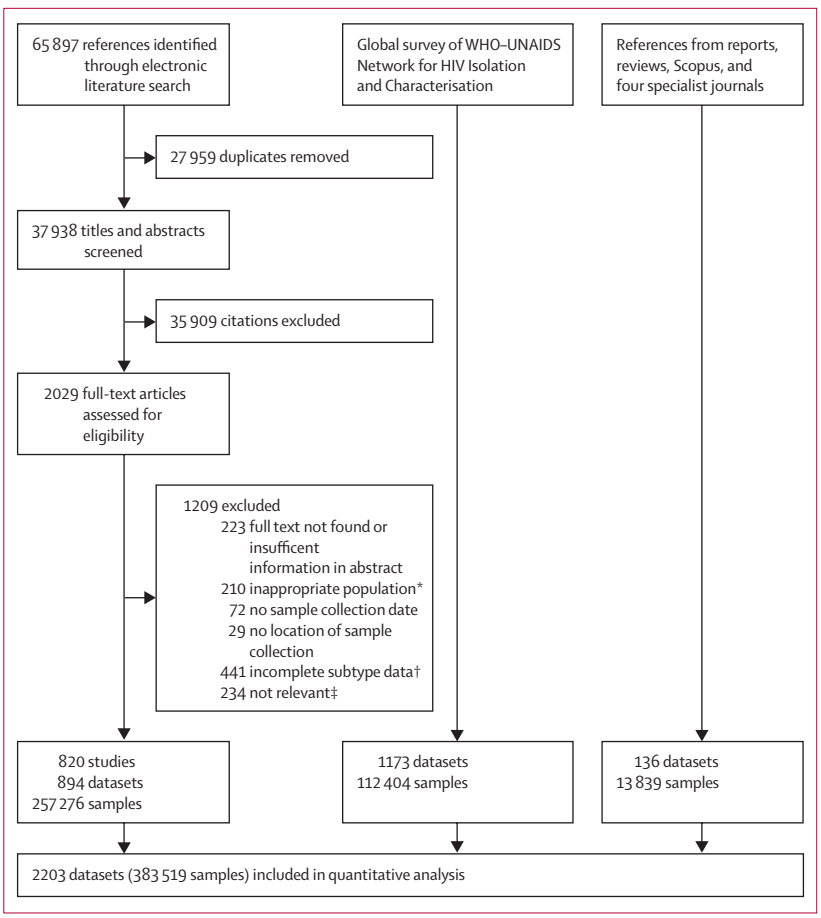

Figure 1: **Study profile**
*For example, HIV-positive immigrants only. †For example, data given for subtype B and non-B samples only. ‡For example, subtypes referred to disease states, not HIV subtypes.

HIV epidemiology data and country groupings

Country-specific estimates of the number of PLHIV in each year for the period 1990–2015 were obtained from UNAIDS.[14] We grouped countries into regions according to the UNAIDS classification, with a few modifications based on subtype distributions in countries and regions (appendix p 9). India and Ethiopia were separated from their respective regions because of their distinct subtype distribution compared with the other countries in their respective regions.

	Caribbean	Latin America	Western and central Europe, and North America	Eastern Europe and central Asia	Oceania	South Asia (India)	Southeast Asia	East Asia	Middle East and north Africa	West Africa	East Africa	Ethiopia	Central Africa	Southern Africa	Global
Number of countries with data available*															
1990–99	1	9	24	8	1	1	8	3	3	5	3	1	5	5	77
2000–04	3	10	28	8	3	1	7	3	4	6	6	1	6	7	93
2005–09	4	9	28	6	2	1	10	4	3	7	7	1	6	9	97
2010–15	3	9	24	6	1	1	7	2	4	8	6	1	5	7	84
1990–2015	4	12	31	11	4	1	10	4	7	8	8	1	6	9	116
Number of datasets															
1990–99	3	63	175	25	4	7	53	25	3	54	42	5	19	24	502
2000–04	6	136	216	30	16	14	31	40	6	60	58	7	16	48	684
2005–09	13	83	241	16	19	18	52	87	9	93	97	4	31	131	894
2010–15	6	57	253	9	11	6	18	42	8	59	28	1	9	45	552
1990–2015	21	263	736	70	43	36	130	150	21	238	186	15	68	226	2203
Number of samples															
1990–99	143	4043	23709	1348	157	382	6743	1323	75	4272	6836	248	1591	1451	52319
2000–04	595	14907	62406	2334	1544	1057	4763	5313	405	4020	6011	237	1351	2571	107513
2005–09	998	10868	84086	839	5686	1342	5637	14235	781	5001	5896	352	1708	9301	146728
2010–15	670	4467	41180	645	4748	272	3134	7899	706	4268	1976	235	853	5907	76959
1990–2015	2406	34285	211381	5166	12135	3052	20277	28769	1967	17561	20719	1072	5503	19229	383519
Coverage†															
1990–99	2·3%	95·5%	99·1%	98·7%	56·6%	100%	97·8%	99·7%	32·4%	88·7%	91·3%	100%	91·5%	90·6%	91·1%
2000–04	65·7%	94·3%	99·5%	96·5%	99·5%	100%	87·6%	99·7%	65·0%	94·1%	95·3%	100%	84·3%	96·4%	95·1%
2005–09	95·0%	84·6%	99·3%	87·4%	34·8%	100%	96·2%	99·7%	61·2%	96·1%	95·4%	100%	100%	98·1%	96·6%
2010–15	39·9%	83·8%	98·6%	90·1%	36·9%	100%	92·0%	97·1%	87·7%	84·8%	97·8%	100%	84·9%	88·6%	90·4%
1990–2015	97·8%	96·8%	92·9%	98·3%	100%	100%	97·2%	99·7%	91·1%	98·8%	99·0%	100%	100%	100%	98·9%
Depth of sampling‡															
1990–99	0·053%	0·520%	1·955%	0·727%	0·627%	0·042%	0·871%	0·636%	0·312%	0·143%	0·174%	0·034%	0·211%	0·026%	0·285%
2000–04	0·141%	1·347%	3·840%	0·415%	3·228%	0·048%	0·358%	0·981%	0·513%	0·082%	0·136%	0·023%	0·121%	0·023%	0·353%
2005–09	0·272%	0·818%	4·421%	0·100%	10·107%	0·060%	0·350%	2·062%	0·620%	0·097%	0·138%	0·044%	0·146%	0·080%	0·455%
2010–15	0·226%	0·282%	1·841%	0·051%	7·265%	0·013%	0·165%	0·975%	0·411%	0·079%	0·042%	0·033%	0·075%	0·047%	0·220%
Mean (SD) number of people living with HIV															
1990–99	270647 (43132)	777108 (58917)	1213002 (71695)	185325 (9309)	25114 (3452)	917464 (228350)	773999 (105211)	208047 (27295)	24066 (6628)	2991282 (376909)	3933298 (183000)	720140 (61262)	755554 (157416)	5550537 (248749)	18345584 (1406452)
2000–04	422471 (48047)	1106688 (88151)	1625178 (103491)	562173 (25378)	47847 (4010)	2185896 (259046)	1329216 (111512)	541451 (45324)	79018 (10343)	4920236 (475926)	4412631 (160012)	1038564 (68766)	1119699 (230893)	11081253 (401209)	30472321 (1780806)
2005–09	366821 (39849)	1327862 (109341)	1901896 (113871)	841816 (39269)	56252 (4248)	2226780 (237499)	1609187 (121453)	690389 (52115)	125841 (13567)	5140275 (462921)	4269141 (148161)	794908 (70770)	1173716 (246435)	11690044 (405665)	32214928 (1801513)
2010–15	295748 (28985)	1586605 (132112)	2236833 (118123)	1256028 (51347)	65354 (4861)	2136545 (227036)	1893512 (140831)	810004 (60250)	171944 (25870)	5419010 (486112)	4704986 (169339)	707751 (59595)	1143531 (258071)	12493788 (449695)	34921639 (1942011)

*Country-level HIV-1 subtyping data are given in the appendix (pp 17–31). †The proportion of people living with HIV in a region who lived in the countries for which HIV subtyping data were available in that region. If any subtyping data were available for a country (independent of the number of samples collected), the whole HIV-infected population in that country was deemed to be represented in this analysis. ‡The number of HIV samples subtyped in a region as a proportion of the number of people living with HIV in that region.

Table 1: **HIV-1 subtyping data collection, 1990–2015**

Data analysis

We split the data into four time periods (1990–99, 2000–04, 2005–09, and 2010–15) on the basis of the spread of datasets and samples across the years. The earliest (1990–99) and latest (2010–15) time periods encompass more years to account for the relatively fewer data available in these years.

For each time period, we calculated the mean number of PLHIV in each country. Coverage was calculated as the proportion of PLHIV in a region who lived in the countries for which HIV-1 subtyping data was available in that region. Depth of sampling was calculated as the number of HIV samples subtyped in a region and is given as a proportion of the mean number of PLHIV in that region.

For each country, HIV-1 subtyping data were split into the four time periods and we determined the subtype distribution (ie, proportions) in each time period on the basis of the number of samples for each subtype, CRF, and URF. For country-specific datasets with sampling years that crossed multiple time periods (eg, 2003–06), we divided the total number of samples for each subtype, CRF, and URF by the number of sampling years, which gave the number of samples for each subtype, CRF, and URF per year. These sample numbers were then assigned to the relevant time periods.

To estimate regional HIV-1 subtype distributions in each time period, we multiplied the proportion of each subtype, CRF, and URF in each country by the estimated mean number of PLHIV in the country. We summed, per region, the resultant (absolute) numbers of each subtype, CRF, and URF in each country and used these values to calculate the proportions of each subtype, CRF, and URF in each region. Countries that lacked HIV-1 subtyping data were not included in these regional calculations.

To estimate global distributions of HIV-1 subtypes, we multiplied the regional proportions of each subtype, CRF, and URF by the regional estimates of the number of PLHIV. We summed the resultant regional (absolute) numbers of each subtype, CRF, and URF to generate a total number of each subtype, CRF, and URF globally, which provided global proportions of each subtype, CRF, and URF. Monte Carlo simulations were used to generate repeated random samples from the cumulative distribution defined by the observed set of subtypes, from which we assessed the level of uncertainty and derived 95% CIs on the proportions of subtypes for the different time periods (appendix pp 13, 14).

We used Windows Excel for all calculations. This systematic review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis guidelines, as applicable. This review is registered online with PROSPERO, number CRD42017067164.

Role of the funding source

This study received no funding. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

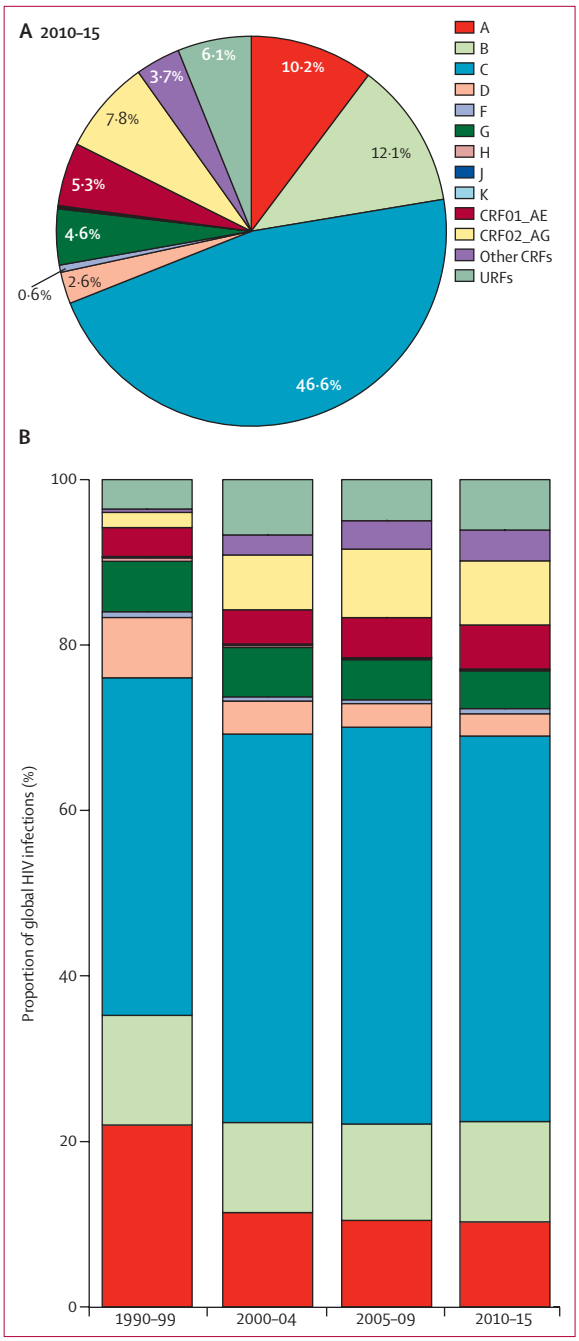

Figure 2: **Global distribution of HIV-1 subtypes, CRFs, and URFs, 1990–2015**
(A) Global proportions of HIV-1 subtypes, CRFs, and URFs in 2010–15. Pie charts for 1990–99, 2000–04, and 2005–09 are shown in the appendix (p 15). (B) Proportions and 95% CIs are shown in table 2 and the appendix (pp 13, 14). CRF=circulating recombinant form. URFs=unique recombinant forms.

Results

Data collection

Our systematic literature search for 1990–2015 yielded 894 datasets with 257 276 samples (figure 1). This search was supplemented with 1173 datasets (112 404 samples) from the global survey of the WHO–UNAIDS Network

	HIV-1 subtypes									CRFs			URFs	Total CRFs*	Total recombinants†
	A	B	C	D	F	G	H	J	K	CRF01_AE	CRF02_AG	Other			
Global															
1990–99	22·0%	13·2%	40·8%	7·3%	0·7%	6·1%	0·4%	0·2%	0·0%	3·5%	1·9%	0·4%	3·6%	5·7%	9·3%
2000–04	11·4%	10·9%	47·0%	3·9%	0·5%	6·0%	0·2%	0·1%	0·0%	4·2%	6·6%	2·5%	6·7%	13·1%	19·7%
2005–09	10·5%	11·6%	48·0%	2·8%	0·5%	4·8%	0·2%	0·1%	0·0%	4·8%	8·3%	3·4%	5·0%	15·9%	20·8%
2010–15	10·3%	12·1%	46·6%	2·7%	0·6%	4·6%	0·1%	0·1%	0·1%	5·3%	7·7%	3·7%	6·1%	16·7%	22·8%
Caribbean															
1990–99	0·0%	47·1%	3·1%	25·8%	0·9%	6·5%	0·9%	0·9%	0·0%	0·0%	0·0%	1·8%	12·9%	1·8%	14·8%
2000–04	0·0%	99·3%	0·1%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·5%	0·1%	0·5%	0·6%
2005–09	0·1%	98·8%	0·1%	0·2%	0·0%	0·1%	0·0%	0·0%	0·0%	0·0%	0·0%	0·6%	0·1%	0·6%	0·7%
2010–15	0·5%	91·0%	0·6%	0·3%	0·0%	0·7%	0·1%	0·0%	0·0%	0·0%	0·1%	3·7%	2·9%	3·9%	6·7%
Latin America															
1990–99	0·0%	87·0%	2·6%	0·1%	7·8%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·9%	1·5%	0·9%	3·2%
2000–04	0·0%	81·0%	6·5%	0·1%	5·2%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	1·6%	5·6%	1·6%	7·1%
2005–09	0·1%	74·7%	8·5%	0·0%	3·9%	0·0%	0·0%	0·0%	0·0%	0·0%	0·1%	2·1%	10·6%	2·2%	12·7%
2010–15	0·1%	76·0%	7·3%	0·0%	4·4%	0·0%	0·0%	0·0%	0·0%	0·0%	0·1%	2·4%	9·6%	2·5%	12·2%
Western and central Europe, and North America															
1990–99	3·6%	88·5%	2·1%	0·4%	0·9%	0·9%	0·1%	0·0%	0·0%	1·4%	0·8%	0·4%	0·7%	2·6%	3·3%
2000–04	2·1%	86·5%	4·0%	0·7%	0·9%	1·0%	0·0%	0·0%	0·0%	0·8%	2·2%	0·8%	1·1%	3·8%	4·8%
2005–09	1·9%	85·0%	3·9%	0·4%	1·1%	1·1%	0·0%	0·0%	0·0%	0·8%	2·9%	1·1%	1·7%	4·7%	6·4%
2010–15	1·9%	83·3%	3·9%	0·3%	1·7%	1·1%	0·0%	0·0%	0·0%	0·9%	3·0%	1·9%	1·9%	5·7%	7·6%
Eastern Europe and central Asia															
1990–99	63·2%	21·3%	3·1%	1·6%	0·0%	4·6%	0·9%	0·0%	0·0%	0·1%	0·0%	4·3%	0·9%	4·4%	5·3%
2000–04	82·1%	13·0%	1·0%	0·4%	0·2%	0·1%	0·0%	0·0%	0·0%	0·1%	0·7%	1·8%	0·5%	2·6%	3·1%
2005–09	91·3%	6·3%	0·0%	0·0%	0·2%	0·0%	0·0%	0·0%	0·0%	0·0%	0·8%	1·5%	0·0%	2·3%	2·3%
2010–15	52·8%	17·4%	6·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	1·4%	21·1%	1·2%	22·6%	23·7%
South Asia (India)															
1990–99	1·9%	0·9%	96·9%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·3%	0·0%	0·3%
2000–04	1·1%	3·0%	95·2%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·1%	0·0%	0·0%	0·5%	0·2%	0·7%
2005–09	1·1%	1·0%	97·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·1%	0·0%	0·0%	0·8%	0·1%	0·9%
2010–15	1·0%	1·2%	94·8%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	3·1%	0·0%	3·1%
Southeast Asia															
1990–99	0·0%	22·7%	1·1%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	73·1%	0·0%	0·0%	3·1%	72·5%	75·6%
2000–04	0·0%	9·9%	2·6%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	80·4%	0·0%	1·2%	5·8%	78·0%	83·7%
2005–09	1·7%	4·4%	2·0%	0·1%	0·0%	0·0%	0·0%	0·0%	0·0%	81·7%	0·4%	1·6%	8·2%	81·9%	90·1%
2010–15	0·0%	18·0%	1·3%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	72·8%	0·2%	1·8%	5·9%	74·2%	80·1%
East Asia															
1990–99	0·9%	68·4%	2·3%	0·0%	0·2%	0·9%	0·0%	0·0%	0·0%	17·7%	0·0%	8·2%	1·4%	25·9%	27·3%
2000–04	0·6%	36·7%	0·8%	0·0%	0·0%	0·1%	0·0%	0·0%	0·0%	22·4%	0·3%	32·4%	6·7%	55·1%	61·8%
2005–09	0·1%	38·7%	1·7%	0·0%	0·0%	0·1%	0·0%	0·0%	0·0%	27·2%	0·2%	29·5%	2·5%	56·7%	59·1%
2010–15	0·0%	17·8%	1·6%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	47·2%	0·1%	28·2%	5·0%	75·5%	80·5%
Oceania															
1990–99	2·2%	80·7%	8·2%	0·0%	0·0%	0·9%	0·0%	0·0%	0·0%	8·0%	0·0%	0·0%	0·0%	8·0%	8·0%
2000–04	0·4%	37·3%	60·0%	0·1%	0·1%	0·0%	0·0%	0·0%	0·0%	1·4%	0·5%	0·1%	0·1%	2·0%	2·1%
2005–09	1·3%	84·9%	5·5%	0·1%	0·0%	0·2%	0·0%	0·0%	0·0%	5·9%	0·8%	0·2%	1·0%	6·9%	7·9%
2010–15	1·7%	75·9%	6·9%	0·3%	0·3%	0·5%	0·0%	0·0%	0·0%	10·7%	1·6%	0·4%	1·7%	12·7%	14·4%
Middle East and north Africa															
1990–99	10·2%	9·2%	26·8%	44·4%	0·0%	0·4%	0·0%	0·0%	0·0%	0·0%	0·1%	0·0%	8·8%	0·1%	8·9%
2000–04	48·9%	41·6%	0·8%	0·4%	0·3%	0·4%	0·0%	0·0%	0·0%	0·2%	5·2%	1·1%	1·1%	6·5%	7·7%
2005–09	9·7%	37·5%	1·2%	0·0%	0·3%	0·6%	0·0%	0·0%	0·0%	1·0%	5·9%	39·5%	4·4%	46·3%	50·7%
2010–15	1·5%	28·9%	1·1%	0·0%	0·2%	0·5%	0·0%	0·0%	0·0%	0·9%	6·7%	59·8%	0·4%	67·4%	67·7%

(Table 2 continues on next page)

	HIV-1 subtypes									CRFs			URFs	Total CRFs*	Total recombinants†
	A	B	C	D	F	G	H	J	K	CRF01_AE	CRF02_AG	Other			
(Continued from previous page)															
West Africa															
1990–99	51·1%	0·1%	0·6%	0·8%	0·9%	32·3%	0·0%	0·2%	0·0%	0·1%	10·1%	0·3%	3·4%	10·5%	14·0%
2000–04	4·3%	0·1%	1·4%	1·0%	0·7%	33·1%	0·1%	0·1%	0·0%	0·8%	37·8%	7·1%	13·5%	45·7%	59·2%
2005–09	3·2%	0·4%	0·9%	1·3%	0·6%	27·8%	0·1%	0·0%	0·0%	0·1%	48·0%	10·4%	7·3%	55·6%	62·7%
2010–15	2·3%	0·0%	1·0%	0·9%	0·5%	26·8%	0·0%	0·0%	0·0%	0·1%	46·2%	6·6%	15·5%	52·9%	68·4%
East Africa															
1990–99	50·3%	0·0%	10·4%	28·5%	0·1%	0·7%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	10·0%	0·0%	10·0%
2000–04	39·4%	0·0%	23·4%	15·7%	0·1%	0·7%	0·1%	0·0%	0·0%	0·0%	0·3%	0·2%	20·0%	0·5%	20·6%
2005–09	50·0%	0·3%	17·4%	17·2%	0·0%	0·4%	0·0%	0·0%	0·0%	0·1%	0·4%	1·4%	12·9%	1·7%	14·4%
2010–15	53·4%	0·1%	14·8%	16·8%	0·1%	0·4%	0·0%	0·4%	0·0%	0·7%	0·0%	0·7%	12·6%	1·4%	13·9%
Ethiopia															
1990–99	0·3%	0·4%	98·4%	0·8%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%
2000–04	1·3%	0·0%	98·3%	0·4%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%
2005–09	0·2%	0·0%	98·7%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·6%	0·0%	0·6%	0·6%	1·1%
2010–15	0·6%	0·0%	99·4%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%
Central Africa															
1990–99	39·7%	0·1%	8·1%	10·0%	3·4%	9·9%	9·1%	2·2%	1·0%	1·3%	3·7%	2·6%	8·8%	7·7%	16·5%
2000–04	30·4%	0·3%	5·5%	11·1%	3·6%	10·5%	5·0%	2·4%	0·9%	2·3%	7·9%	4·7%	15·5%	14·8%	30·3%
2005–09	13·6%	0·3%	13·2%	6·8%	4·1%	6·3%	4·6%	1·4%	0·2%	1·6%	9·2%	13·4%	25·5%	22·0%	47·5%
2010–15	14·7%	0·8%	11·9%	6·5%	6·1%	6·9%	3·2%	1·4%	1·7%	1·4%	7·7%	16·4%	21·3%	25·5%	46·8%
Southern Africa															
1990–99	0·7%	2·8%	95·8%	0·4%	0·0%	0·2%	0·0%	0·0%	0·0%	0·0%	0·0%	0·0%	0·1%	0·0%	0·1%
2000–04	5·5%	0·6%	88·8%	2·9%	0·0%	0·3%	0·0%	0·0%	0·0%	0·0%	0·0%	0·9%	0·9%	1·0%	1·9%
2005–09	0·4%	2·0%	97·0%	0·2%	0·0%	0·1%	0·0%	0·0%	0·0%	0·0%	0·1%	0·1%	0·2%	0·1%	0·4%
2010–15	0·4%	0·4%	98·8%	0·0%	0·0%	0·1%	0·0%	0·0%	0·0%	0·0%	0·1%	0·1%	0·2%	0·1%	0·3%

Distribution of HIV-1 subtypes, CRFs, and URFs as a proportion of global and regional HIV infections for four time periods. 95% CIs of proportions are given in the appendix (pp 13, 14). CRF=circulating recombinant form. URFs=unique recombinant forms. *Total CRFs is the sum of CRF01_AE, CRF02_AG, and other CRFs. †Total recombinants is the sum of total CRFs and URFs.

Table 2: **Global and regional distribution of HIV-1 subtypes, CRFs, and URFs, 1990–2015**

for HIV Isolation and Characterisation and 136 datasets (13 839 samples) from references from reports, reviews, Scopus, and specialist journals, yielding a total of 2203 datasets and 383 519 samples from 116 countries over 1990–2015 (figure 1, table 1).

For the analysis, the dataset was split into four time periods (1990–99, 2000–04, 2005–09, and 2010–15; table 1). The countries for which HIV-1 subtype distribution data were available had greater than 90% of global HIV-infected individuals in each time period, and greater than 95% in 2000–04 and 2005–09 (table 1). Ten of 14 regions had greater than 90% coverage in 1990–09, compared with seven regions in 2010–15 (table 1). Low coverage was seen in some time periods for the Caribbean, Oceania, and Middle East and north Africa. Sampling depth (defined as the number of samples out of the number of PLHIV) was highest in western and central Europe, North America and Oceania, and was lowest in south Asia (India), Ethiopia, and southern Africa (table 1). The latter three regions have large HIV-infected populations and very little subtype diversity.

Sequencing has been the mainstay of HIV-1 subtyping since 2000 (>97% of samples in 2000–15, with 99·8% in 2010–15; appendix, p 10), mostly in *pol* (>77% of samples in 2000–15, with 94·2% in 2010–15; appendix p 11). In addition, the proportion of samples assessed in only one genome segment increased from 77·0–83·4% in 1990–2004 to 87·2–90·1% in 2005–15, and the proportion of full length sequences decreased from 1·4–1·9% to 0·3–0·4% in the same time frame (appendix p 12).

Global distribution of HIV-1 subtypes, CRFs, and URFs

In 2010–15, subtype C accounted for 46·6% (16 280 897/34 921 639 of PLHIV) of all HIV-1 infections worldwide (figure 2A, table 2, appendix pp 13–15). Subtype B was responsible for 12·1% (4 235 299/34 921 639) of infections and subtype A for 10·3% (3 587 003/34 921 639) of infections, followed by CRF02_AG (7·7%; 2 705 110/34 921 639), CRF01_AE (5·3%; 1 840 982/34 921 639), sub type G (4·6%; 1 591 276/34 921 639), and subtype D (2·7%; 926 255/34 921 639). Subtypes F, H, J, and K combined

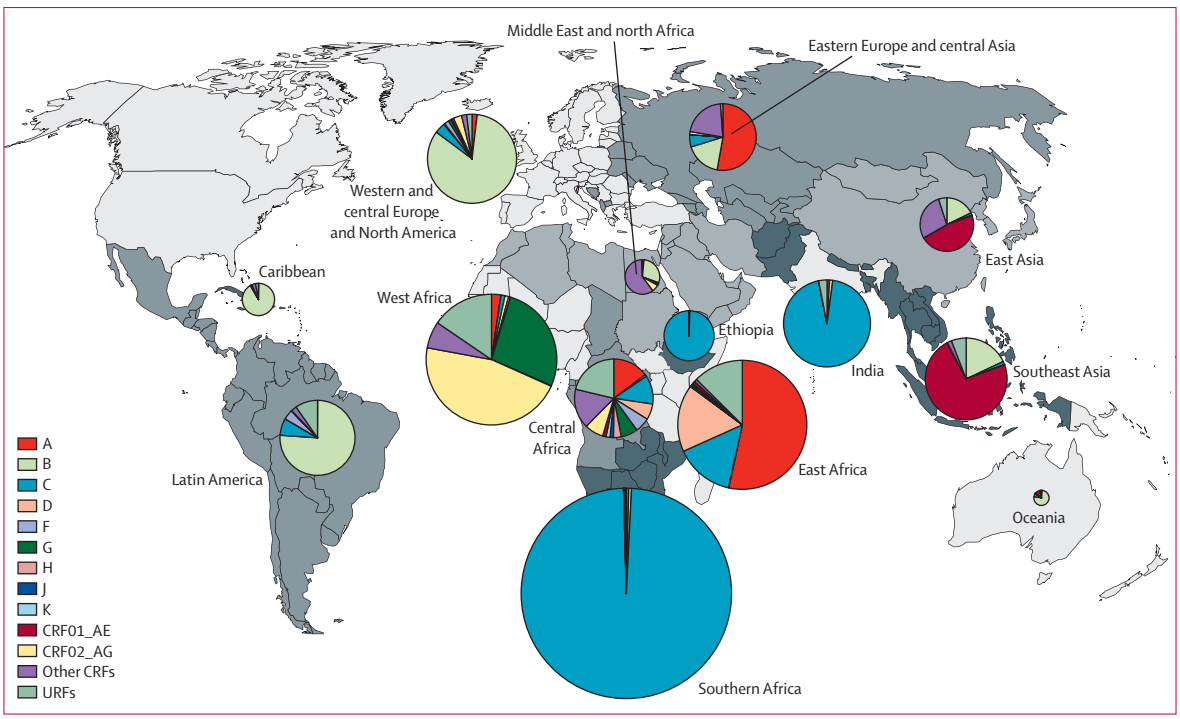

Figure 3: **Regional distributions of HIV-1 subtypes, CRFs, and URFs, 2010–15**
We grouped countries into 14 regions (appendix p 9). Countries forming a region are shaded in the same colour on the world map. The surface area of each pie-chart corresponds to the number of people living with HIV in each region. Equivalent maps for 1990–99, 2000–04, and 2005–09 are given in the appendix (p 16). CRF=circulating recombinant form. URFs=unique recombinant forms.

accounted for 0·9% (311 332/34 921 639) of infections globally. Other CRFs accounted for 3·7% (1 309 082/34 921 639) of infections, bringing the total proportion of CRFs to 16·7% (5 844 113/34 921 639). URFs accounted for 6·1% (2 134 405/34 921 639) of infections, which brought the proportion of all global HIV-1 infections attributable to recombinants to 22·8% (7 978 517/34 921 639).

Changes in the global distribution were seen over time (figure 2B, table 2, appendix pp 13–15). Subtype C increased its proportion in 1990–2009, followed by a decrease in 2010–15. After a decrease between 1990 and 2004, the proportion of subtype B consistently increased throughout 2000–15. The proportions of subtypes A, D, and G decreased throughout 1990–2015, although subtypes A and D stabilised in 2010–15. Subtypes F, H, J, and K contribute a small proportion of infections, with a combined global total of around 1% throughout 1990–2015. Amongst the recombinants, CRF02_AG increased in 1990–2009, followed by a decrease in 2010–15. CRF01_AE and other CRFs had consistently increased proportions in 1990–2015, contributing to a consistent increase in the proportion of all CRFs over all time periods. URFs made a substantial contribution throughout 1990–2015. Overall, a consistent increase in the total proportion of recombinants was seen in 1990–2015 (figure 2B, table 2, appendix pp 13–15).

Regional distribution of HIV-1 subtypes, CRFs, and URFs

The proportions of HIV-1 subtypes, CRFs, and URFs differed among regions and changed over time (figure 3, table 2, appendix pp 13, 14, 16–31).

Diversity was greatest in central Africa, where all HIV-1 subtypes and many CRFs and URFs were found, throughout all time periods (figure 4, table 2, appendix pp 13, 14, 30). Over time, decreases in the proportions of subtypes A, D, G, and H were accompanied by increases in proportions of subtype C, other CRFs, and URFs. In 2010–15, central Africa had the highest proportion of URFs (21·3%; 243 041/1 143 531) of any region, contributing to a total proportion of recombinants of 46·8% (534 997/1 143 531).

West Africa had the highest proportion of CRF02_AG (46·2% [2 504 438/5 419 010] in 2010–2015) and subtype G (26·8%; 1 454 444/5 419 010) of any region in each time period. With many URFs (15·5%; 838 476/5 419 010), the total proportion of recombinants was 68·4% (3 706 246/5 419 010; figure 3, table 2, appendix pp 13, 14, 16–31).

In east Africa, 53·4% (2 510 665/4 704 986) of all infections were caused by subtype A in 2010–15, which remained broadly stable over time, with notable contributions by subtype C (14·8%; 696 163/4 704 986), subtype D (16·8%; 791 501/4 704 986), and URFs (12·6%; 591 140/4 704 986; table 2).

Subtype C dominated in southern Africa, Ethiopia, and south Asia (India), where it contributed at least 89% of infections throughout 1990–2015 and few other subtypes or recombinants were found (table 2).

Subtype B was the main contributor in western and central Europe and North America, Caribbean, Latin America, and Oceania, where it accounted for at least 75% of infections in 2010–2015. The proportion of subtype B generally decreased over time in western and central Europe and North America, and Latin America, with concomitant increases in CRFs and URFs.

In eastern Europe and central Asia, more than 50% of infections were caused by subtype A (across all time periods), with notable contributions by subtype B and other CRFs (table 2).

In the Middle East and north Africa, infections were caused by a decreasing proportion of subtype B and an increasing proportion of other CRFs (mainly CRF35_AD [data not shown]), which accounted for 59·8% (102 853/171 944) of infections in the region in 2010–15—the highest proportion of other CRFs in any region at any time (table 2).

Infections in southeast Asia and east Asia were dominated by CRF01_AE, with an overall proportion of recombinants of about 80% in each region in 2010–15. Southeast Asia consistently had the highest proportion of CRF01_AE infections of any region at any time. In east Asia, CRF01_AE increased consistently over time to 47·2% (381 996/810 004) in 2010–2015, with further contributions by other CRFs (28·2% [228 720/810 004] in 2010–15; mainly CRF07_BC and CRF08_BC [data not shown]), resulting in the highest proportion of total CRFs (75·5%; 611 520/810 004) and total recombinants (80·5%; 652 181/810 004) of any region (figure 3, table 2, appendix pp 13, 14, 16–31).

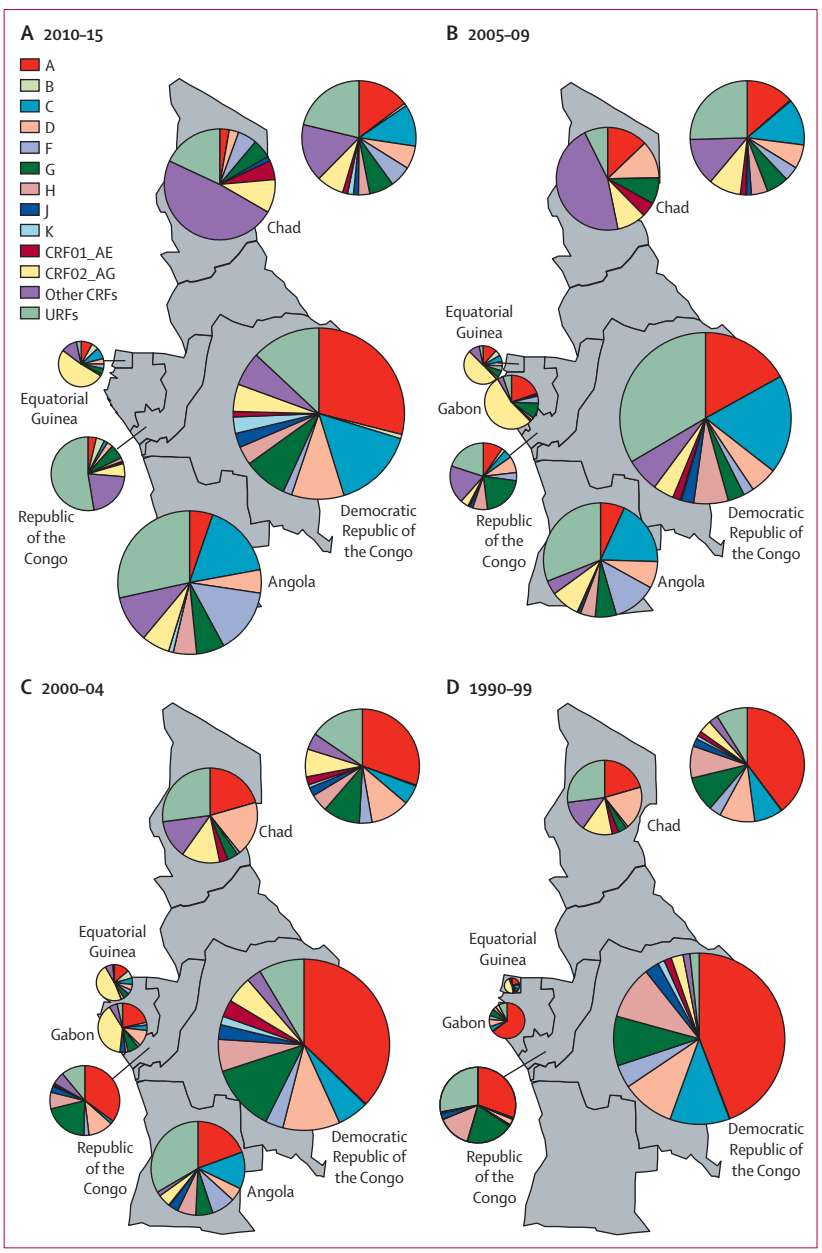

Figure 4: **Country and regional distribution of HIV-1 subtypes, CRFs, and URFs in central Africa, 1990–2015**
The surface area of each pie chart corresponds to the number of people living with HIV in each country per time period. The regional distribution of HIV-1 subtypes, CRFs, and URFs in each time period is given in the top-right corner of each panel. CRF=circulating recombinant form. URFs=unique recombinant forms.

Discussion

This systematic review and global survey is the largest study to date reporting on the global and regional distribution and trends of HIV-1 subtypes and recombinants, with a sample size of 383 519, and spanning 1990–2015, covering the whole period for which reliable estimates of national HIV prevalence are available for most countries. With our approach, we achieved a high coverage both globally and regionally throughout the time periods.

Distinct distributions of HIV-1 subtypes and recombinants and dynamic changes were seen in different countries and regions. Over the most recent decade (2005–15) at a global level, the proportion of subtype B increased, subtypes A and D were stable, and subtypes C and G and CRF02_AG decreased. CRF01_AE, other CRFs, and URFs increased, leading to a consistent increase in the global proportion of recombinants over time. These global and regional trends are the result of changes in HIV-1 subtype distributions at country level and changes in the number of PLHIV in countries over time. Overall, trends reflect the relative differences in the numbers of new infections and deaths associated with each subtype and recombinant. The factors that lead to these patterns are complex and multifactorial, and include possible biological differences between subtypes leading to differences in transmission and disease progression.[15,16] However, there are also many geographical and socioeconomic factors, including transportation links, founder effects, migration, urbanisation, transmission networks, and population growth.[17,18] Furthermore, the wider

availability of antiretroviral treatment globally has reduced HIV-related deaths and new infections, leading to a slower renewal of PLHIV. However, disparities in treatment and prevention coverage and effectiveness between geographical regions and risk groups might lead to differential control of the HIV epidemic in different regions or risk groups with different HIV variants, thereby shaping the global and regional distributions of HIV-1 subtypes and recombinants.

A consistent increase in the proportion of CRFs, and recombinants overall, throughout the time periods, was seen globally and in most regions. Over time, the number of different CRFs identified increased—with 96 identified thus far—and the formation of new recombinants is ongoing.[3] The recognition of CRF02_AG led to a reclassification of some strains that would previously have been assigned as subtype A on the basis of the characterisation of *env* only.[19] This change led to a decrease in subtype A and an increase in CRF02_AG between the periods 1990–99 and 2000–04 reported for west Africa and, to a lesser extent, central Africa, which translated into global changes that do not represent genuine epidemiological changes and should be interpreted with caution. However, the increase over time in the reported regional and global proportions of CRFs are probably not due to better characterisation and assignment of CRFs alone. The methods and number and type of genome segments used to subtype samples have narrowed over time, and HIV-1 subtyping is now dominated by *pol* sequencing—often as part of testing for resistance. Indeed, the prevalence of recombinants is probably underestimated in our study because of the restricted proportion of samples that are subtyped in more than one genome segment.

Our study had some further limitations. Although coverage was generally high, some sample sizes were small and few data were from nationally representative surveys. Low coverage in some time periods for the Caribbean, Oceania, and Middle East and north Africa—with different countries and populations sampled in different time periods—led to regional variations in the distribution of subtypes over time, which do not represent genuine epidemiological changes and should be interpreted with caution (appendix pp 17, 25, 26). However, these three regions contain the smallest absolute numbers of PLHIV of all regions and hence these changes have little effect on global proportions. Furthermore, dividing datasets that straddled different analytic time periods assumed that subtype data were collected evenly over the sampling years or that subtype distribution was stable during the sampling years (or both), which might not have been the case. Publication bias might have occurred. Finally, heterogeneity among datasets in study types, populations sampled, subtyping methods used, and genome segments analysed could have introduced biases. In particular, different inclusion criteria for different studies and different risk groups sampled would lead to characterisation of subtype distributions in specific populations, or risk groups, which might not be fully representative of the distribution of subtypes in a country. Use of certain subtyping methods (eg, heteroduplex mobility assay) only detects a restricted number of prespecified subtypes or CRFs and examination of different genome segments in different studies could lead to differential classification of recombinants that correspond to different subtypes in different genome segments. These biases could affect comparative results between countries or regions, but also within countries and regions over time.

The choice of a vaccine immunogen sequence is crucial to the efficacy of a vaccine against HIV. Of the candidate HIV vaccines in development, the two most clinically advanced concepts have been reported in phase 1 and 2a trials in humans,[20,21] demonstrating adequate safety and immunogenicity to justify progression to human phase 2b and 3 efficacy trials. One of these human HIV vaccine efficacy trials (HVTN702; NCT02968849), which is ongoing in South Africa, investigates a vaccine based on the vaccine used in the successful RV144 trial in Thailand,[22] with the important adaptation that subtype B/CRF01_AE immunogens have been replaced with subtype C isolate immunogens, to match the HIV subtype that dominates in South Africa. If this subtype C vaccine is effective against subtype C infection, one of the next steps would be to test whether the vaccine offers cross-protection against other HIV-1 subtypes and recombinants, by conducting trials in regions with other circulating subtypes or recombinants, such as east Africa, where subtypes A, C, and D co-circulate and hence relative protection against diverse subtypes can be evaluated. The other vaccine that entered into an efficacy trial (HVTN705; NCT03060629) in southern Africa is a polyvalent mosaic vaccine designed to cover all global HIV-1 group M viruses.

Matching immunogen sequences to circulating strains is challenging. Primary isolates belonging to the same HIV subtype typically differ by 8–17% (maximum 30%) at the amino acid level, whereas this difference is 17–35% (maximum 42%) between isolates from different subtypes, depending on the subtypes and genome segments examined.[23] Recombination adds to this complexity. However, the amino acid difference between immunogen sequence and circulating viral strains of the same subtype might be halved by using artificial centralised sequences, such as consensus, ancestral, or centre-of-tree sequences.[6,24] These approaches can be expanded to include all global subtypes (eg, centralised sequences of HIV-1 group M).[6] Other approaches to address HIV diversity include mosaic and polyvalent vaccines and focusing on evolutionarily conserved regions of the HIV genome.[5] These strategies are not mutually exclusive and can be used in combination. Other vaccine approaches involving different vectors and immunogens might be more effective when tested in human trials.[5]

In addition to aiding vaccine design and development, knowledge of the regional and global prevalence of HIV-1 subtypes and recombinants is also instrumental in the estimation of the global need and market size for therapeutic and prophylactic HIV vaccines based on different subtypes, the reason being that this allows prioritisation of vaccines on the basis of subtypes with the greatest potential global benefit.[25] Our analysis shows that subtype C causes the greatest number of global infections and that subtype C is the near-exclusive strain of HIV in southern Africa. Given that, in addition, the prevalence of HIV is greatest in southern Africa, it is logical to focus on a subtype C-specific HIV vaccine and test it in South Africa, the country with the greatest number of PLHIV.[1,20] In many countries and regions, however, the HIV epidemic is not solely composed of a few dominant subtypes, but instead a substantial and increasing proportion of infections is caused by various minority subtypes and recombinants. It remains to be seen whether a subtype-specific vaccine will protect against these diverse subtypes and recombinants, or whether complex cocktails of subtype-specific vaccines or a global group M-specific vaccine will be required.[5,21] Moreover, since the distribution of HIV subtypes is evolving and the number and proportion of recombinants are increasing, a future vaccine against HIV might need to be changed periodically, like influenza vaccines.[23]

HIV-1 genetic diversity also impacts the efficiency of HIV diagnostic assays and viral load assays, which affects blood donor screening, surveillance, clinical diagnosis, and management.[3] A study[7] compared a range of commercially available fourth generation antigen-antibody, p24 antigen, and viral load assays against a standardised panel of HIV-1 subtypes and recombinants. Differences in the efficiency of detection of different subtypes and recombinants by diagnostic assays were seen but, notably, the different assays tested had the same rank order of detection of diverse isolates, with similar levels of detection across subtypes. Viral load assays did well across a range of genotypes and viral loads.[7] However, ongoing recombination is expected to lead to the generation of new URFs and CRFs for which the assays have not been designed, and vigilance is therefore required. Most major resistance mutations found in subtype B are also found in non-B subtypes, although the prevalence of some mutations is higher in certain subtypes (eg, K65R/N mutation in the reverse transcriptase gene with tenofovir treatment) than in others.[8] Several novel mutations occur in non-B subtypes.[9,26] However, drug resistance is not well studied in the subtypes, CRFs, and URFs that are less common. Continued surveillance for transmitted drug resistance and treatment failure is therefore essential in all HIV subtypes and recombinants.

In summary, the current study, which included 383 519 subtyped samples, shows great global HIV-1 diversity and dynamic trends over time. Continued effective surveillance of the global and regional molecular epidemiology of HIV-1 is crucial for the design, testing, and implementation of vaccines against HIV.[27]

Contributors

JH conceived, designed, and coordinated the study, wrote the systematic review protocol, assisted with the literature search, assessed eligibility of manuscripts, collected additional published data, did the global survey, extracted the data, designed and did data analysis, designed figures and tables, interpreted the data, and wrote the manuscript. RE, JY, and LD-T screened the electronic literature search results for relevant manuscripts, assessed their eligibility, extracted data, and collected additional published data. JY and LD-T analysed and interpreted data and made figures. IF analysed data. SK designed and did the electronic literature search. BW assisted with the statistical analyses. EG-W provided data on the number of people living with HIV in each country, assisted with the statistical analysis, and interpreted the data. PDG assisted with the statistical analysis and interpreted the data. All authors read and approved the final version of the manuscript.

WHO-UNAIDS Network for HIV Isolation and Characterisation

Alash'le G Abimiku, Simon Agwale, Chris Archibald, Boaz Avidor, María Gabriela Barbás, Francoise Barre-Sinoussi, Banson Barugahare, El Hadj Belabbes, Silvia Bertagnolio, Deborah Birx, Aleksei F Bobkov, James Brandful, Helba Bredell, Catherine A Brennan, James Brooks, Marie Bruckova, Luigi Buonaguro, Franco Buonaguro, Stefano Buttò, Anne Buve, Mary Campbell, Jean Carr, Alex Carrera, Manuel Gómez Carrillo, Connie Celum, Beth Chaplin, Macarthur Charles, Dimitrios Chatzidimitriou, Zhiwei Chen, Katsumi Chijiwa, David Cooper, Philip Cunningham, Anoumou Dagnra, Cillian F de Gascun, Julia Del Amo, Elena Delgado, Ursula Dietrich, Dominic Dwyer, Dennis Ellenberger, Barbara Ensoli, Max Essex, Herve Fleury, Peter N Fonjungo, Vincent Foulongne, Deepak A Gadkari, Feng Gao, Federico García, Roger Garsia, Guy Michel Gershy-Damet, Judith R Glynn, Ruth Goodall, Zehava Grossman, Monick Lindenmeyer Guimarães, Beatrice Hahn, Raph L Hamers, Osamah Hamouda, Ray Handema, Xiang He, Joshua Herbeck, David D Ho, Africa Holguin, Mina Hosseinipour, Gillian Hunt, Masahiko Ito, Mohamed Ali Bel Hadj Kacem, Erin Kahle, Pontiano Kaleebu, Marcia Kalish, Adeeba Kamarulzaman, Chun Kang, Phyllis Kanki, Edward Karamov, Jean-Claude Karasi, Kayitesi Kayitenkore, Tony Kelleher, Dwip Kitayaporn, Leondios G Kostrikis, Claudia Kucherer, Claudia Lara, Thomas Leitner, Kirsi Liitsola, Jai Lingappa, Marek Linka, Ivette Lorenzana de Rivera, Vladimir Lukashov, Shlomo Maayan, Luzia Mayr, Francine McCutchan, Nicolas Meda, Elisabeth Menu, Fred Mhalu, Doreen Mloka, John L Mokili, Brigitte Montes, Orna Mor, Mariza Morgado, Fausta Mosha, Awatef Moussi, James Mullins, Rafael Najera, Mejda Nasr, Nicaise Ndembi, Joel R Neilson, Vivek R Nerurkar, Florian Neuhann, Claudine Nolte, Vlad Novitsky, Philippe Nyambi, Marianna Ofner, Fem J Paladin, Anna Papa, Jean Pape, Neil Parkin, Chris Parry, Martine Peeters, Alexandra Pelletier, Lucía Pérez-Álvarez, Deenan Pillay, Angie Pinto, Trinh Duy Quang, Cecilia Rademeyer, Filimone Raikanikoda, Mark A. Rayfield, Jean-Marc Reynes, Tobias Rinke de Wit, Kenneth E Robbins, Morgane Rolland, Christine Rousseau, Jesus Salazar-Gonzales, Hanan Salem, Mika Salminen, Horacio Salomon, Paul Sandstrom, Mario L Santiago, Abdoulaye D Sarr, Bryan Schroeder, Michel Segondy, Philippe Selhorst, Sylvester Sempala, Jean Servais, Ansari Shaik, Yiming Shao, Amine Slim, Marcelo A Soares, Elijah Songok, Debbie Stewart, Julie Stokes, Shambavi Subbarao, Ruengpung Sutthent, Jun Takehisa, Amilcar Tanuri, Kok Keng Tee, Kiran Thapa, Michael Thomson, Tyna Tran, Willy Urassa, Hiroshi Ushijima, Philippe van de Perre, Guido van der Groen, Kristel van Laethem, Joep van Oosterhout, Ard van Sighem, Eric van Wijngaerden, Anne-Mieke Vandamme, Jurgen Vercauteren, Nicole Vidal, Lesley Wallace, Carolyn Williamson, Dawit Wolday, Jianqing Xu, Chunfu Yang, Linqi Zhang, Rong Zhang

Affiliations

Institute of Human Virology, University of Maryland, Baltimore, MD, USA (A G Abimiku, J Carr); Gede Foundation, Abuja, Nigeria (S Agwale); Public Health Agency of Canada, Ottawa, ON, Canada (C Archibald, J Brooks, M Ofner, P Sandstrom, J Stokes); Tel-Aviv Sourasky Medical

Center, Tel-Aviv, Israel (B Avidor); Ministerio de Salud, Córdoba, Argentina (M G Barbás); Institut Pasteur, Paris, France (F Barre-Sinoussi, E Menu); Ministry of Health, Entebbe, Uganda (B Barugahare); National Reference Laboratory on HIV/AIDS, Institut Pasteur d'Algérie, Algiers, Algeria (E Belabbes); World Health Organization, Geneva, Switzerland (S Bertagnolio); Office of the Global AIDS Coordinator, Washington, DC, USA (D Birx); The D I Ivanovsky Institute of Virology, Moscow, Russia(A F Bobkov); Noguchi Memorial Institute for Medical Research, University of Ghana, Accra, Ghana (J Brandful); University of Cape Town, Cape Town, South Africa (H Bredell, C Rademeyer, P Selhorst, D Stewart, C Williamson); Abbott Laboratories, Chicago, IL, USA (C A Brennan); National Institute of Public Health, Prague, Czech Republic (M Bruckova, M Linka); AIDS Reference Center, National Cancer Institute "Fond. G. Pascale", Naples, Italy (L Buonaguro, F Buonaguro); National AIDS Center, Istituto Superiore di Sanità, Rome, Italy (S Buttò, B Ensoli); Institute of Tropical Medicine, Antwerp, Belgium (A Buve, G van der Groen); University of Washington School of Medicine, Seattle, WA, USA (M Campbell, C Celum, J Herbeck, E Kahle, J Lingappa, J Mullins, M Rolland, C Rousseau); St Vincent's Hospital, Sydney, Australia (A Carrera, P Cunningham); University of Buenos Aires, Buenos Aires, Argentina (M Carrillo, H Salomon); Harvard T H Chan School of Public Health, Boston, MA, USA (B Chaplin, P Kanki); Gheskio Center, Port-au-Prince, Haiti (M Charles, C Nolte, J Pape); Aristotle University of Thessaloniki, Thessaloniki, Greece (D Chatzidimitriou, A Papa); Chinese Academy of Medical Sciences, Peking Union Medical School, Beijing, China (Z Chen, L Zhang, R Zhang); Fukuoka Institute of Health and Environmental Sciences, Kyushu University Hospital, Dazaifu, Japan (K Chijiwa); The Kirby Institute, Sydney, NSW, Australia (D Cooper, T Kelleher, A Pinto, A Shaik); Faculté des Sciences de la Santé, Université de Lomé, Lomé, Togo (A Dagnra); University College Dublin, Dublin, Ireland (C de Gascun); Instituto de Salud Carlos III, Madrid, Spain (J Del Amo, E Delgado, R Najera, L Pérez-Álvarez, M Thomson); Chemotherapeutisches Forschungsinstitut, Georg-Speyer-Haus, Frankfurt, Germany (U Dietrich); Pathology West, Westmead Hospital, Westmead, NSW, Australia (D Dwyer, K Thapa, T Tran); Centers for Disease Control and Prevention, Atlanta, GA, USA (D Ellenberger, P N Fonjungo, M A Rayfield, K E Robbins, S Subbarao, C Yang); Harvard School of Public Health, Boston, MA, USA (M Essex, V Novitsky, A D Sarr); Duke University Medical Center, Durham, NC, USA (F Gao); University of Bordeaux, Bordeaux, France (H Fleury); Montpellier University Hospital, Montpellier, France (V Foulongne, P van de Perre); National AIDS Research Institute, Pune, India (D A Gadkari); Complejo Hospitalario Universitario de Granada, Granada, Spain (F García); Royal Prince Alfred Hospital, Sydney, Australia (R Garsia, H Salem); HIV Laboratory Programme on AIDS/AFRO, World Health Organisation, Ouagadougou, Burkina Faso (G M Gershy-Damet); London School of Hygiene and Tropical Medicine, London, UK (J R Glynn); University College London, London, UK (R Goodall); National HIV Reference Laboratory, Ministry of Health, Tel Aviv, Israel (Z Grossman, O Mor); Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil (M L Guimarães, M Morgado); University of Pennsylvania, Philadelphia, PN, USA (B Hahn); Amsterdam Institute for Global Health and Development, Amsterdam, Netherlands (R L Hamers, T Rinke de Wit); Robert Koch Institute, Berlin, Germany (O Hamouda, C Kucherer); Yamanashi Medical University, Yamanashi, Japan (R Handema, M Ito); National Center for AIDS/STD Control and Prevention, China CDC, Beijing, China (X He, Y Shao, J Xu); Aaron Diamond AIDS Research Center, The Rockefeller University, New York, NY, USA (D D Ho, L G Kostrikis); Ramón y Cajal Research Institute, Hospital Universitario Ramón y Cajal de Madrid, Madrid, Spain (A Holguin); University of North Carolina, Chapel Hill, NC, USA (M Hosseinipour); National Institute for Communicable Diseases, Johannesburg, South Africa (G Hunt); Charles Nicolle Hospital, Tunis, Tunisia (M Kacem, A Moussi, M Nasr, A Slim); Medical Research Council, Entebbe, Uganda (P Kaleebu, C Parry); Vanderbilt Institute for Global Health, Vanderbilt University School of Medicine, Nashville, TN, USA (M Kalish); University of Malaya, Kuala Lumpur, Malaysia (A Kamarulzaman, K-K Tee); Institute for Molecular Biology and Genetics and Medical College, Seoul National University, Seoul, Korea (C Kang); Gamaleya Center for Epidemiology and Microbiology, Moscow, Russian (E Karamov); National Reference Laboratory, Kigali, Rwanda (J-C Karasi); Emory University School of Medicine, Atlanta, GA, USA (K Kayitenkore); HIV/AIDS Collaboration, Nonthaburi, Thailand (D Kitayaporn); Karolinska Institute, Huddinge University Hospital, Stockholm, Sweden (C Lara); Los Alamos National Laboratory, Los Alamos, NM, USA (T Leitner); National Institute for Health and Wellfare, Helsinki, Finland (K Liitsola, M Salminen); National Autonomous University of Honduras, Tegucigalpa, Honduras (I Lorenzana de Rivera); Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands (V Lukashov); Hadassah U Hospital, Jerusalem, Israel (S Maayan); New York University School of Medicine, New York, NY, USA (L Mayr, P Nyambi); Henry M Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD, USA (F McCutchan); Centre Muraz, Bobo-Dioulasso, Burkina Faso (N Meda); Muhimbili University of Health Sciences, Dar-es-salaam, Tanzania (F Mhalu, D Mloka, F Mosha, W Urassa); University of Edinburgh, Edinburgh, UK (J L Mokili); Montpellier University Hospital, Montpellier, France (B Montes, M Segondy); Institute of Human Virology, Abuja, Nigeria (N Ndembi); University of Washington, Seattle, WA, USA (J R Neilson); University of Hawaii, Honolulu, HI, USA (V R Nerurkar); University Clinic Heidelberg, Heidelberg, Germany & Lighthouse Trust, Lilongwe, Malawi (F Neuhann); Research Institute for Tropical Medicine, Muntinlupa City, Manila, Philippines (F J Paladin, M L Santiago); Data First Consulting, Inc, Belmont, CA, USA (N Parkin); University of Montpellier, Montpellier, France (M Peeters, N Vidal); Centre de Recherche Public-Santé, Luxembourg, Luxembourg (A Pelletier, J Servais); Africa Health Research Institute, Durban, KwaZulu-Natal, South Africa & Division of Infection and Immunity, University College London, London, UK (D Pillay); Institute of International Health, University of Tokyo, Tokyo, Japan (T D Quang); University of Sydney, Sydney, NSW, Australia (F Raikanikoda); Institut Pasteur du Cambodge, Phnom Penh, Cambodia (J-M Reynes); University of Alabama at Birmingham, Birmingham, AL, USA (J Salazar-Gonzales); Auckland City Hospital, Auckland, New Zealand (B Schroeder); Uganda Virus Research Institute, Entebbe, Uganda (S Sempala); Instituto Nacional de Câncer, Rio de Janeiro, Brazil (M A Soares); Kenya Medical Research Institute, Nairobi, Kenya (E Songok); National HIV Repository and Bioinformatic Center, Siriraj Hospital, Mahidol University, Thailand (R Sutthent); Laboratory of Viral Pathogenesis, Kyoto University, Kyoto, Japan (J Takehisa); Federal University of Rio de Janeiro, Rio de Janeiro, Brazil (A Tanuri); Aino Health Science Center and Aino University, Tokyo, Japan (H Ushijima); Rega Institute for Medical Research, KU Leuven, Belgium (K van Laethem, E van Wijngaerden, A-M Vandamme, J Vercauteren); Department of Medicine, Blantyre, Malawi (J van Oosterhout); Stichting HIV Monitoring, Amsterdam, Netherlands (A van Sighem); Health Protection Scotland, Glasgow, UK (L Wallace); and Ethiopian Health & Nutrition Research Institute, Addis Ababa, Ethiopia (D Wolday).

Declaration of interests

We declare no competing interests.

Acknowledgments

JH is supported by the Oxford University Clinical Academic Graduate School (Oxford, UK) and Linacre College, Oxford University (Oxford, UK).

References

1 UNAIDS. Global AIDS update 2018. Geneva: UNAIDS, 2018.
2 Fauci AS. An HIV vaccine is essential for ending the HIV/AIDS pandemic. *JAMA* 2017; **318:** 1535–36.
3 Hemelaar J. The origin and diversity of the HIV-1 pandemic. *Trends Mol Med* 2012; **18:** 182–92.
4 Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. *Science* 2000; **288:** 55–56.
5 Barouch DH, Korber B. HIV-1 vaccine development after STEP. *Annu Rev Med* 2010; **61:** 153–67.
6 Gaschen B, Taylor J, Yusim K, et al. Diversity considerations in HIV-1 vaccine selection. *Science* 2002; **296:** 2354–60.
7 Stone M, Bainbridge J, Sanchez AM, et al. Comparison of detection limits of fourth- and fifth-generation combination HIV antigen-antibody, p24 antigen, and viral load assays on diverse HIV isolates. *J Clin Microbiol* 2018; **56:** JCM.02045–17.
8 TenoRes Study Group. Global epidemiology of drug resistance after failure of WHO recommended first-line regimens for adult HIV-1 infection: a multicentre retrospective cohort study. *Lancet Infect Dis* 2016; **16:** 565–75.

9 Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM. The challenge of HIV-1 subtype diversity. *N Engl J Med* 2008; **358:** 1590–602.

10 Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS* 2011; **25:** 679–89.

11 Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 2006; **20:** W13–23.

12 Arien KK, Vanham G, Arts EJ. Is HIV-1 evolving to a less virulent form in humans? *Nat Rev Microbiol* 2007; **5:** 141–51.

13 WHO. WHO HIV Drug Resistance Report 2012. Geneva: World Health Organization, 2012.

14 UNAIDS. Global AIDS update 2016. Geneva: UNAIDS, 2016.

15 Kiwanuka N, Laeyendecker O, Quinn TC, et al. HIV-1 subtypes and differences in heterosexual HIV transmission among HIV-discordant couples in Rakai, Uganda. *AIDS* 2009; **23:** 2479–84.

16 Venner CM, Nankya I, Kyeyune F, et al. Infecting HIV-1 subtype predicts disease progression in women of sub-Saharan Africa. *EBioMedicine* 2016; **13:** 305–14.

17 Faria NR, Rambaut A, Suchard MA, et al. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science* 2014; **346**: 56–61.

18 Tebit DM, Arts EJ. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. *Lancet Infect Dis* 2011; **11:** 45–56.

19 Carr JK, Salminen MO, Albert J, et al. Full genome sequences of human immunodeficiency virus type 1 subtypes G and A/G intersubtype recombinants. *Virology* 1998; **247:** 22–31.

20 Bekker LG, Moodie Z, Grunenberg N, et al. Subtype C ALVAC-HIV and bivalent subtype C gp120/MF59 HIV-1 vaccine in low-risk, HIV-uninfected, South African adults: a phase 1/2 trial. *Lancet HIV* 2018; **5:** e366–78.

21 Barouch DH, Tomaka FL, Wegmann F, et al. Evaluation of a mosaic HIV-1 vaccine in a multicentre, randomised, double-blind, placebo-controlled, phase 1/2a clinical trial (APPROACH) and in rhesus monkeys (NHP 13-19). *Lancet* 2018; **392:** 232–43.

22 Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med* 2009; **361:** 2209–20.

23 Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, Detours V. Evolutionary and immunological implications of contemporary HIV-1 variation. *Br Med Bull* 2001; **58:** 19–42.

24 Nickle DC, Jensen MA, Gottlieb GS, et al. Consensus and ancestral state HIV vaccines. *Science* 2003; **299:** 1515–18.

25 Marzetta CA, Lee SS, Wrobel SJ, Singh KJ, Russell N, Esparza J. The potential global market size and public health value of an HIV-1 vaccine in a complex global market. *Vaccine* 2010; **28:** 4786–97.

26 Kantor R, Katzenstein DA, Efron B, et al. Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. *PLoS Med* 2005; **2:** e112.

27 Aldrich C, Hemelaar J. Global HIV-1 diversity surveillance. *Trends Mol Med* 2012; **18:** 691–94.