

# **THE ROLE OF MIGRATION AND DOMESTIC TRANSMISSION IN THE SPREAD OF HIV-1 NON-B SUBTYPES IN SWITZERLAND**

Viktor von Wyl\* 1,2; Roger D. Kouyos\* 3; Sabine Yerly 4; Jürg Böni 5; Cyril Shah 5; Philippe Bürgisser 6; Thomas Klimkait 7; Rainer Weber 1; Bernard Hirschel 8; Matthias Cavassini 9; Cornelia Staehelin 10; Manuel Battegay 11; Pietro L. Vernazza 12; Enos Bernasconi 13; Bruno Ledergerber 1; Sebastian Bonhoeffer 3; Huldrych F. Günthard 1, and the Swiss HIV Cohort Study.

\* contributed equally.

1 Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, University of Zurich, Zurich, Switzerland;

2 current affiliation: Department of Primary Care and Population Sciences, University College London, London, UK;

3 Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

4 Laboratory of Virology, University Hospital Geneva and University of Geneva Medical School, Geneva, Switzerland;

5 Swiss National Center for Retroviruses, University of Zurich, Zurich, Switzerland;

6 Division of Immunology, University Hospital Lausanne, Lausanne, Switzerland;

7 Institute for Medical Microbiology, University of Basel, Basel, Switzerland

8 Division of Infectious Diseases, Geneva University Hospital, Geneva, Switzerland;

9 University Hospital Lausanne, Lausanne, Switzerland;

10 Clinic for Infectious Diseases, Bern University Hospital and University of Bern, Switzerland;

11 Division of Infectious Diseases and Hospital Epidemiology, University Hospital Basel, Basel, Switzerland;

12 Division of Infectious Diseases, Cantonal Hospital St. Gallen, St. Gallen, Switzerland;

13 Division of Infectious Diseases, Regional Hospital Lugano, Lugano, Switzerland;

Reprints or correspondence: Dr. Viktor von Wyl, University Hospital Zurich, Division of Infectious Diseases and Hospital Epidemiology, Raemistrasse 100, 8091 Zurich, Switzerland ([vovv@usz.uzh.ch](mailto:vovv@usz.uzh.ch));

alternate corresponding author: Dr. Huldrych Günthard, University Hospital Zurich, Division of Infectious Diseases and Hospital Epidemiology, Raemistrasse 100, 8091 Zurich, Switzerland ([huldrych.guenthard@usz.ch](mailto:huldrych.guenthard@usz.ch))

**Word count**

Abstract: 200

Main text: 3406

**Short title:** Non-B HIV-1 epidemics in Switzerland

**Key words:** HIV-1; Molecular epidemiology; HIV-1 subtypes; Phylogeny; Switzerland

**Funding sources:** This study has been financed in the framework of the Swiss HIV Cohort Study, supported by the Swiss National Science Foundation (SNF grant #33CS30-134277). Further support was provided by the SNF grants #3247B0-112594 (to HFG, SY, BL, SB), SNF Grant for Prospective Researchers PBEZP3-125726 (to VvW), and 324730\_130865 (to HFG), SHCS project 470, 528 and 569, the SHCS Research Foundation, by a further research grant of the Union Bank of Switzerland in the name of a donor to HFG and by the European Community's Seventh Framework Programme (grant FP7/2007–2013), under the Collaborative HIV and Anti-HIV Drug Resistance Network (CHAIN; grant 223131). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Potential conflicts of interest:** Potential conflicts of interest: H.F.G. has been an adviser and/or consultant for the following companies: GlaxoSmithKline, Abbott, Novartis, Boehringer Ingelheim, Roche, Tibotec and Bristol-Myers Squibb, and has received unrestricted research and educational grants from Roche, Abbott, Bristol-Myers Squibb,

GlaxoSmithKline, Tibotec and Merck Sharp & Dohme. S.Y. has participated in advisory boards of BMS and Tibotec, and has received travel grants from GSK and MSD. TK served as advisor for BMS and Pfizer and has received travel grants from Abbott and Pfizer. All other named authors have stated no competing interests.

## **Abstract**

**Background:** By analyzing HIV-1 pol sequences from the Swiss HIV Cohort Study (SHCS) we explored whether the prevalence of non-B subtypes reflects domestic transmission or migration patterns.

**Methods:** Swiss non-B sequences and sequences collected abroad were pooled to construct maximum likelihood trees, which were analyzed for Swiss-specific subepidemics, (subtrees including  $\geq 80\%$  Swiss sequences, bootstrap  $>70\%$ ; macroscale analysis) or evidence for domestic transmission (sequence pairs with genetic distance  $<1.5\%$ , bootstrap  $\geq 98\%$ ; microscale analysis).

**Results:** Of 8287 SHCS participants, 1732 (21%) were infected with non-B subtypes, of which A (n=328), C (n=272), CRF01\_AE (n=258), and CRF02\_AG (n=285) were studied further. The macroscale analysis revealed that 21% (A), 16% (C), 24% (CRF01\_AE) and 28% (CRF02\_AG) belonged to Swiss-specific subepidemics. The microscale analysis identified 26 possible transmission pairs: 3 (12%) including only homosexual Swiss men of white ethnicity; 3 (12%) including homosexual white men from Switzerland and partners from foreign countries; and 10 (38%) involving heterosexual white Swiss men and females of different nationality and predominantly non-white ethnicity.

**Conclusions:** Of all non-B infections diagnosed in Switzerland,  $<25\%$  could be prevented by domestic interventions. Awareness should be raised among immigrants and Swiss individuals with partners from high prevalence countries to contain the spread of non-B subtypes.

## **Introduction**

Over the last decade, the prevalence of HIV-1 non-B subtypes has increased in Western Europe. For example, in a time series extending from 1996 to 2005, Yerly et al. noted a steady increase in non-B subtypes among 822 recently infected individuals in Switzerland [1]. Such increases are commonly attributed to migration, as non-B subtypes are mostly found among people from non European origin.

While Swiss subtype B epidemics have already been studied extensively [2], the possible role of domestic infections with non-B subtypes is not well investigated. A phylogenetic study from the UK identified two clusters of epidemiologically linked A1 subtype sequences involving men who have sex with men of British nationality [3]. Significant clustering of non-B subtypes was also observed for HIV-1 infected individuals who have acquired the disease through heterosexual contacts. In the UK, at least 14% (subtype A) and 6% (subtype C) were found to be included in UK-specific clusters, therefore suggesting ongoing transmissions within the UK [4]. Local infections with non-B subtypes in B-dominated regions have also been identified by phylogenetic linkage of individuals diagnosed with primary HIV-1 infection [5]. For example, by screening recently infected patients Yerly et al identified a subtype CRF11 subepidemic among intravenous drug users in Western Switzerland [6].

The present analysis aimed to study three questions: How abundant are non-B subtypes in Switzerland, do Swiss-specific non-B subepidemics exist (macroscale analysis), and finally, can we find evidence for ongoing non-B transmissions in Switzerland (i.e. domestic transmission) by identifying possible transmission pairs (microscale analysis)?

In order to address these questions, we applied a molecular phylogeny approach to non-B sequences collected from participants of the Swiss HIV Cohort Study (SHCS).

## **Materials and Methods**

### *Sequences & Patients*

The SHCS is a nationwide, prospective, clinic-based cohort study with continuous enrolment and semi-annual study visits [7, 8]. The SHCS has been approved by ethical committees of all participating institutions, and written informed consent has been obtained from participants. The SHCS drug resistance database contains all HIV resistance tests performed by the four laboratories engaged in HIV resistance testing in Switzerland, stored in a central database developed and hosted by SmartGene (Zug, Switzerland, IDNS version 3.5.0) [9]. Resistance data stem from routine clinical testing (60% of tests) and tests performed retrospectively from frozen repository plasma samples (40% of tests). The retrospective sequencing was systematically performed using for each patient the earliest available plasma sample. All laboratories are performing population-based sequencing of the full protease gene and in minimum codons 28-225 of the reverse transcriptase gene using commercial assays (Viroseq Vs.1 PE Biosystems, Rotkreuz, Switzerland; Virsoeq Vs. 2, Abbott AG, Baar, Switzerland; vircoTYPE HIV-1 Assay, Virco Lab, Mechelen, Belgium) and in-house methods [10]. Subtyping was performed on the contig of the protease and the reverse transcriptase sequence using the REGA 2 System [11]. If this method returned inconclusive results, the analysis was repeated with the Star analyzer (<http://www.vgb.ucl.ac.uk/starn.shtml>)[12].

The full SHCS drug resistance database contains 13750 sequences from 9593 patients with protease and reverse transcriptase sequences available. To study the epidemiology of non-B subtypes in Switzerland we selected all viral sequences from individuals who were seen at least once in the SHCS between January 1, 1996 (earliest availability of plasma samples for retrospective sequencing) and June 30, 2010 (database closure for this analysis; see Acknowledgements for Genbank accession numbers). Time trends in distribution of B and non-B viruses were analyzed for a subset of these patients who received their HIV diagnosis between January 1, 1996 and December 31, 2009 (to account for time lags in data reporting) through a positive HIV test.

The phylogenetic analysis presented here focuses on non-B subtypes A and C as well as CRF AE and AG, because only these occur in large enough numbers in Switzerland to perform meaningful analyses (see results). In order to identify Swiss-specific subepidemics, we pooled the earliest available sequence per SHCS participant with all non-Swiss sequences that were available for the same subtype/CRF from the Los Alamos Sequence database (<http://www.hiv.lanl.gov/>, date of accession: September 2010). The following numbers of sequences were available for the different subtypes/CRF: 2018 (subtype A), 2954 (subtype C), 1455 (CRF01\_AE), 1643 (CRF02\_AG). To avoid a distortion of our analysis by convergent evolution driven by antiretroviral therapy, we removed all major amino acid positions that are associated with antiretroviral drug resistance according to the International AIDS Society - USA (IAS-USA) guidelines [13] (positions 30,32,33,46,47,48,50,54,76,82,84,88, and 90 in the PR and 41,62,65,67,69,70,74,75,77,100,103,106,108,115,116,151,181,184,188,190,210,215,219, 225, and 236 in the RT).

Demographic (transmission groups, geographical origin within Switzerland, nationality) and administrative data (sampling year for genotypic test, year of patient's enrollment in the cohort, year of first positive HIV test) were obtained from the SHCS database. Transmission groups were categorized into heterosexual transmission, men having sex with men, and intravenous drug users. Other transmission modes (perinatal, transfusion risks) were not considered. Information on patient ethnicity is self-reported according to one of the five categories: white, black, asian, hispanic, or other. Patients' nationalities were categorized as follows for this analysis: Western Europe, sub-Saharan Africa, and South-East Asia. Because other regions only occurred in small numbers they were subsumed in one category.

### *Phylogenetic methods*

The phylogenetic methods are described in detail in [2]. For each of the considered subtypes/CRF (subtypes A and C, CRF01\_AE and CRF02\_AG), we inferred maximum-likelihood trees (+100 bootstrap runs) using the GTR model with  $\Gamma$ -distributed rate heterogeneity implemented in RAxML [14]. These trees were then searched for patterns that indicate Swiss-specific subepidemics, which were defined as subtrees including  $\geq 80\%$  Swiss sequences and separated by bootstrap support  $> 70\%$  from the main tree (macroscale analysis). In addition, trees were analyzed for evidence of domestic transmission, which was defined as possible transmission chains with non-B virus that involved at least one individual of white ethnicity and Swiss nationality (microscale



analysis). Possible transmission chains were identified on the basis of sequence pairs with a genetic distance of <1.5% and bootstrap support  $\geq 98\%$ .

### Statistical analyses

Categorical data were analyzed with the Chi2 test. Time trends were investigated with the Cochran-Armitage test for trend. Statistical calculations were done with Stata 11.1 (Stata Corp., College Station, TX). All p-values are two-sided.

## Results

### *Epidemiology and time trends on non-B subtypes in Switzerland*

Of 10540 individuals seen at least once in the Swiss HIV Cohort Study between 1996 and 2010, subtype information was available for 8287 individuals (78.6%). Of these 8287 persons, 6555 (79.1%) were infected with subtype B. There were four non-B subtypes, which occurred at a prevalence of 3% or greater: A (n= 328; 4.0 %), C (n= 272; 3.3 %), CRF01\_AE (n=258; 3.1%), and CRF02\_AG (n=285; 3.4%). All remaining subtypes contributed approximately 1% or less to the sequence database; these were D (n=55), F (n=74), G (n=99), H (n=8), J (n=4), K (n=2), CRF03\_AB (n=2), CRF06\_CPX (n=26), CRF10\_CD (n=1), CRF11\_CPX (n=38), CRF12\_BF (n=8), and CRF13\_CPX (n=8). In addition, 267 viral samples (3.2%) were classified as unspecified recombinants of different viral subtypes by the subtyping algorithms used [11, 12].

Time trends in occurrence of B and non-B subtypes were studied for 4767 individuals whose HIV diagnoses (positive HIV tests) fell between 1996 and 2009. Over this time

period the proportion of non-B subtype viruses increased from 22% (86 non-B subtypes / 385 HIV diagnoses in 1996) to 33% (78 non-B subtypes / 237 HIV diagnoses in 2009) in our study population (figure 1a;  $p$  trend = 0.004). When analyzing the data for each mode of HIV acquisition separately, there were also appreciable, although in some cases non-linear, time trends for increases in non-B infections among the risk groups of heterosexuals (figure 1b; total  $n$  = 2196; 1996: 79 non-B subtypes / 174 diagnosed (45%); 2009: 63/95 (66%);  $p < 0.001$ ), intravenous drug users (total  $n$  = 514; 1996: 5/85 (6%); 2009: 4/7 (50%);  $p < 0.001$ , not shown). In contrast, the proportion of non-B subtypes remained comparably low among 2056 homosexual men (1996: 2/126 (2%); 2009: 11/135 (8%);  $p$  trend  $< 0.001$ , not shown).

*Macro-level phylogenetic analyses identify Swiss-specific non-B subepidemics*

In total, we identified 1143 individuals infected with one of the four subtypes A, C, CRF01\_AE and CRF02\_AG. Characteristics of these patients are displayed in table 1. The three subtypes (A, C, CRF02\_AG) are endemic in sub-Saharan Africa, which is also reflected by the high proportions of individuals with black ethnicity and African origin in these groups (48% in A, 60% in C, 61% in CRF02\_AG). In contrast, individuals of Asian ethnicity and origin only contributed 32% to the Southeast-Asia endemic subtype CRF01\_AE. Each of the four subtype groups also consisted of 24% (C) to 57% (CRF01\_AE) of individuals of Western European nationality and white ethnicity (other subtypes: 28% in CRF02\_AG, 38% in A).

In order to identify non-B subepidemics specific for Switzerland, the phylogenetic maximum likelihood trees constructed on data from the 1143 SHCS participants were

searched for subtrees, which were separated from the full tree with a bootstrap support >70% and contained at least 80% of Swiss sequences (figure 2, macroscale analysis). We empirically found that these criteria were well suited to detect Swiss subepidemics, because of the low frequency of Swiss sequences relative to the number of added non-B sequences collected outside of Switzerland, which made the risk for falsely detecting Swiss clusters small. Overall, 108 Swiss-specific subepidemics involving 254 out of 1143 (22%) non-B infected SHCS participants were detected. Ninety-two of these subepidemics predominantly consisted of persons with a heterosexual orientation (including 204 of the 254 individuals, of whom 7 with a history of intravenous drug use), whereas MSM-dominated clusters were observed in 16 instances, (50 individuals). Most of these sub-epidemics tended to be small, with only 17% of the Swiss specific subtrees consisting of more than two individuals (11% [n=10] and 5% [n=4] with cluster size 3 and 4, respectively, and one cluster [1%] of size 5). Although not reaching statistical significance, there was a trend for MSM dominated subepidemics to form larger transmission chains as illustrated by a somewhat higher fraction of subepidemics reaching a size >2 (25%) compared with heterosexually dominated subepidemics (18%) and the greater size of the largest cluster (14 vs. 5 in heterosexual subepidemics).

As illustrated by table 2, different patterns of clustering were observed across subtypes. The proportion of non-B infected SHCS participants included in Swiss-specific subepidemics varied significantly (Test for heterogeneity in proportions of clustering across subtypes  $p = 0.008$ ). The highest degree of clustering was observed for the CRF02\_AG group, where 80 of 285 (28%) individuals were contained within 34 Swiss

specific subepidemics, followed by CRF01\_AE with 61 of 258 SHCS participants (24%, 23 subepidemics). Corresponding proportions of clustering sequences (number of clusters) for subtype A and C were 21% (69 of 328 individuals, 31 clusters) and 16% (69 of 328 individuals, 20 clusters), respectively. Of note, we found no indication for a time trend in these proportions, irrespective of viral subtype (all tests for trend  $p > 0.14$ ; data not shown). In particular, there was no indication for a higher degree of clustering of viral samples collected in later years, which – if found - could be suggestive for emerging self-sustained non-B epidemics.

In order to distinguish between possible domestic transmission and immigration, we analyzed clustering patterns according to groups defined by ethnicity and nationality. In particular, we considered sequences contained within Swiss-specific clusters to have likely resulted out of domestic transmission events (i.e. to have occurred within Switzerland), and the remaining infections as possibly having been introduced from abroad through migration. Among the African subtypes A, C, and CRF02\_AG, the overall proportion of individuals with black ethnicity and Sub-Saharan African nationality included in Swiss-specific sub-epidemics was 16.4% (81/449; subtype A: 31/157 (19.8%), C: 15/164 (9.2%), CRF02\_AG: 35/173 (20.2%); test for across subtypes of proportions  $p = 0.0184$ ), whereas viral sequences from white persons from Western Europe were linked to Swiss subepidemics in 30.2% of cases (81/268; subtype A: 29/123 (23.6%); C: 19/66 (28.8%); CRF02\_AG: 33/79 (41.8%); test for heterogeneity of proportions across subtypes  $p = 0.114$ ). For the Southeast Asia-endemic subtype CRF01\_AE the proportion of sequences from Asian individuals contained within Swiss-

specific clusters was 22.9% (19/83), and the proportion of clustering sequences among white, Western European individuals was 24.7% (36/146).

Next, we studied sociodemographic factors associated with inclusion of sequences in Swiss-specific subepidemics using logistic regression. Due to expected differences in socio-demographic factors between the three African subtypes A, C, and CRF02\_AG and the South-East Asian subtype CRF01\_AE we chose to analyze these two subtype groups separately. Univariable logistic regression models restricted to the African subtypes detected associations of male sex (odds ratio 1.8 [95% confidence interval 1.3-2.4]), white ethnicity (2.3 [1.7-3.3] compared to blacks), Western European nationality (2.1 [1.5-3.0] compared to Sub-Saharan Africa) and homosexual mode of HIV acquisition (2.0 [1.0-2.8] compared to heterosexual infection) with a higher probability for inclusion in Swiss-specific subepidemics. In contrast, the only positively associated factor for the Asian subtype CRF01\_AE was HIV acquisition through homosexual contacts (4.6 [2.4-8.8]). Due to the strong cross-correlation of these demographic characteristics, for example between nationality and ethnicity or sex and mode of HIV acquisition, multivariable analyses were not performed.

Instead, characteristics of Swiss-specific subepidemics were further analyzed descriptively and on the basis of demographic subgroups defined according to ethnicity and mode of HIV acquisition. Figure 3, panel A highlights differences in these demographic subgroups between the African subtypes and the Asian subtype CRF01\_AE. The major differences found were i) lower proportions of men who have sex with men in

subepidemics of African subtype when compared to subtype CRF01\_AE and ii) the comparably small numbers of white females (n=4) and heterosexual asian males (n=1) were observed in CRF01\_AE subepidemics (both present at <5% and therefore not plotted). Conversely, the proportion of white and black men and women appeared to be in balance in clusters of African subtypes.

When performing sensitivity analyses we found that our conclusions were robust to methodological variations of our approach. For instance, the number of Swiss subepidemics decreases by less than 20% on average if the analysis was restricted to subtrees with bootstrap support above 95% (instead of > 70%). Likewise, the number of subepidemics changed by <10% if the fraction of Swiss sequences defining a Swiss subepidemic was increased from 80% to 100%.

#### *Identification of possible transmission pairs in micro-scale analysis*

In addition to the macro-scale analysis, which identified Swiss-specific non-B subepidemics, we further aimed to detect possible transmission pairs by more conservative criteria (micro-scale analysis), defined as sequence pairs clustering with a bootstrap support  $\geq 98\%$  and a genetic distance  $< 0.015$  on the trees calculated for the macro-scale analysis. [15, 16]

A total of 26 possible transmission pairs were detected (figure 3, panel B). Among the 17 pairs infected with African subtypes, the most frequently observed types of transmission events were among pairs of heterosexual men and women both of black ethnicity (n total= 6; A: n=0; CRF02\_AG: n=2; C: n=4), and heterosexual transmission between

individuals of mixed ethnicities (n total=6; A: n=1; C: n=1, CRF02\_AG: n= 4). Noteworthy, in all of the latter 6 transmission pairs the male partner was of white ethnicity. Possible heterosexual transmission between white individuals was observed 2 times (A: n=1; CRF02\_AG: n=1) and possible homosexual transmission between individuals of mixed ethnicity was seen in 3 instances (A: n=1, C: n=1, CRF02\_AG: n=1). Among the patients infected with the Asian subtype CRF01\_AE 5 possible transmission events between Asian females and white males were identified, as well as 3 events among white homosexual men. In addition, 1 possible heterosexual transmission event included two individuals of white ethnicity.

Overall, when also taking information on nationality into account, local transmission was likely in 16 (62%) of the 26 possible transmission events, because at least one partner was of white ethnicity and Swiss nationality. Of 6 possible events involving homosexual transmission there were three pairs of white men from Switzerland and 3 pairs, which included white men from Switzerland and partners from other world regions. Ten possible transmission pairs consisted of heterosexual white Swiss men and females from other world regions, 8 of those were of non-white ethnicity. In addition, there were three potential heterosexual transmission pairs involving at least one white individual from Western European countries other than Switzerland, thus making infection in Switzerland seem less certain.

## **Discussion**

By analyzing >8000 HIV-1 pol sequences from Switzerland we detected significant increases in the proportion of non-B subtypes over a time period of 12 years, which were most apparent among individuals who acquired HIV through heterosexual contacts. Among the most frequent non-B subtypes with a prevalence >3% were A, C, CRF01\_AE, and CRF02AG. A macro scale phylogenetic analysis revealed that only around 20% of newly detected non-B infections among individuals of Asian or African nationality may have originated from domestic transmission, whereas the remainder of infections most likely have occurred outside of Switzerland. Although this proportion was higher among white, Western European individuals and reaching up to 42% for subtype CRF02\_AG, substantial numbers of sequences were not associated with Swiss-specific sub-epidemics, possibly reflecting infections acquired during travel and stays abroad [17]. Along these lines, the macro-scale analysis further indicated that, consistent with findings from a previous analysis of the Swiss subtype B epidemic among heterosexuals [2], the non-B epidemic can only be maintained through frequent reintroduction from abroad. The presence of self-sustained non-B epidemics would lead to large transmission chains in the phylogenetic tree (similar to those observed among subtype B infected MSM in Switzerland [2]), and possibly also to an increase in the ratio of clustering (i.e. Swiss-specific) and non-clustering sequences over time. Neither was seen in our macro-scale analysis.

Our observations from the macroscale analysis are in line with results from a modeling study, which also found a much greater contribution of migration to newly diagnosed non-B infections among Africans than domestic transmission [18]. The Dutch study predicted a proportion of 60% of newly diagnosed infections among African immigrants



to have occurred outside the Netherlands, whereas in our study 80% and more of all non-B infections among Africans may have originated outside of Switzerland, given that only around 20% of all sequences from this group were contained within Swiss-specific clusters.

Nevertheless, our micro-scale analysis yielded substantial evidence for possible domestic non-B transmissions, because, even with a very conservative criterion, we detected 16 instances of genetically linked viral samples including at least one individual of white ethnicity and Swiss nationality. A surprisingly large proportion (11/16, 69%) of these possible local transmission events involved a constellation of white Western European males and non-white females, which is suggestive for an important role of interracial partnerships in transmission of non-B subtypes within Switzerland. These patterns are also consistent with a contribution of sex tourism or prostitution to non-B infections in Switzerland, as occasionally reported by newly HIV-diagnosed individuals as a possible source of infection [unpublished observation, Günthard].

Several study limitations should be noted. Our analysis is subject to the usual and unavoidable caveats associated with molecular epidemiological analyses such as sampling bias (see [2] for a detailed discussion), but these problems are substantially alleviated by the high representativeness of the SHCS [8]. Furthermore, the criteria used to identify possible transmission pairs in the micro-scale analysis were strict and have successfully been used previously, confirmed by clonal sequences of a second gene, to establish linkages between HIV-infected individuals with primary HIV infection [16]. However, linkage between individuals can never be established with absolute certainty

due to possible intermediary links, which may not have been sampled. Moreover, we cannot fully exclude that white and non-white individuals may have been infected abroad and later migrated to Switzerland together. However, the finding that 4 of the 26 possible transmission pairs were embedded in Swiss-specific subepidemics of size 3 (n=3) or greater (size 14: n=1; figure 2) provides strong evidence that domestic infections with the four most prevalent non-B subtypes indeed do occur in Switzerland.

Overall, our results stand in contrast to findings of the subtype B epidemics among IDUs and MSMs in Switzerland, which were driven by domestic transmission [2], but also to the common notion that HIV infections with non-B subtypes are mainly a problem of migration. Instead, the epidemiology of non-B HIV-1 in Europe should be understood as the combined result of both migration and domestic transmission. Moreover, these findings have significant public health implications, because they demonstrate that less than one fourth of all non-B infections diagnosed in Switzerland could possibly be prevented by local interventions, and they also suggest that awareness should be raised among immigrants and Swiss individuals with partners from countries with high HIV prevalence to contain the spread of HIV in Switzerland.

### **The Swiss HIV Cohort Study Members:**

Barth J, Battegay M, Bernasconi E, Böni J, Bucher HC, Bürgisser P, Burton-Jeangros C, Calmy A, Cavassini M, Dubs R, Egger M, Elzi L, Fehr J, Fischer M, Flepp M, Francioli P (President of the SHCS), Furrer H (Chairman of the Clinical and Laboratory Committee), Fux CA, Gorgievski M, Günthard H (Chairman of the Scientific Board), Hasse B, Hirsch HH, Hirschel B, Hösli I, Kahlert C, Kaiser L, Keiser O, Kind C, Klimkait T, Kovari H, Ledergerber B, Martinetti G, Martinez de Tejada B, Müller N, Nadal D, Pantaleo G, Rauch A, Regenass S, Rickenbach M (Head of Data Center), Rudin C (Chairman of the Mother & Child Substudy), Schmid P, Schultze D, Schöni-Affolter F, Schüpbach J, Speck R, Taffé P, Telenti A, Trkola A, Vernazza P, von Wyl V, Weber R, Yerly S.

### **Acknowledgements:**

A sample of 10% of the sequences used in this analysis were submitted to Genbank (Genbank Accession Numbers EF449787, EF449788, EF449796, EF449798, EF449828, EF449829, EF449838, EF449844, EF449852, EF449853, EF449854, EF449860, EF449880, EF449883, EF449889, EF449895, EF449901, EF449904, EF449905, EF449917, EF449921, EF449928, EF449930, EF449943, EF449950, EF449960, EF449971, EF449980, EF449987, EF450004, EF450005, EF450011, EF450024, EF450026, GQ848113, GQ848120, GQ848140, GQ848145, GQ848149, JF769777-JF769851].

We thank the patients for participation in the SHCS and the clinical trials, the physicians and study nurses for excellent patient care, the laboratory technicians of the Swiss resistance laboratories for the quality of the data, SmartGene, Zug, Switzerland for providing excellent technical support, and Brigitte Remy, Martin Rickenbach and Yannick Vallet from the SHCS data center in Lausanne for the data management.

## REFERENCES

1. Yerly S, von Wyl V, Ledergerber B, et al. Transmission of HIV-1 drug resistance in Switzerland: a 10-year molecular epidemiology survey. *AIDS* **2007**;21:2223-9.
2. Kouyos RD, von Wyl V, Yerly S, et al. Molecular Epidemiology Reveals Long-Term Changes in HIV Type 1 Subtype B Transmission in Switzerland. *Journal of Infectious Diseases* **2010**;201:1488-1497.
3. Gifford RJ, Oliveira T, Rambaut A, et al. Phylogenetic surveillance of viral genetic diversity and the evolving molecular epidemiology of human immunodeficiency virus type 1. *Journal of Virology* **2007**;81:13050-13056.
4. Hughes GJ, Fearnhill E, Dunn D, Lycett SJ, Rambaut A and Leigh Brown AJ. Molecular phylodynamics of the heterosexual HIV epidemic in the United Kingdom. *PLoS Pathog* **2009**;5:e1000590.
5. Brenner BG, Roger M, Routy JP, et al. High rates of forward transmission events after acute/early HIV-1 infection. *Journal of Infectious Diseases* **2007**;195:951-9.
6. Yerly S, Junier T, Gayet-Ageron A, et al. The impact of transmission clusters on primary drug resistance in newly diagnosed HIV-1 infection. *AIDS* **2009**;23:1415-23.
7. Ledergerber B, Egger M, Opravil M, et al. Clinical progression and virological failure on highly active antiretroviral therapy in HIV-1 patients: a prospective cohort study. *Lancet* **1999**;353:863-868.
8. Cohort profile: the Swiss HIV Cohort study. *Int J Epidemiol*;39:1179-89.
9. Von Wyl V, Yerly S, Boni J, et al. Emergence of HIV-1 drug resistance in previously untreated patients initiating combination Antiretroviral treatment - A comparison of different regimen types. *Archives of Internal Medicine* **2007**;167:1782-1790.
10. Yerly S, Vora S, Rizzardì P, et al. Acute HIV infection: impact on the spread of HIV and transmission of drug resistance. *Aids* **2001**;15:2287-2292.
11. de Oliveira T, Deforche K, Cassol S, et al. An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics* **2005**;21:3797-800.
12. Myers RE, Gale CV, Harrison A, Takeuchi Y and Kellam P. A statistical model for HIV-1 sequence classification using the subtype analyser (STAR). *Bioinformatics* **2005**;21:3535-40.
13. Hirsch MS, Gunthard HF, Schapiro JM, et al. Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society-USA panel. *Clinical Infectious Diseases* **2008**;47:266-285.
14. Stamatakis A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **2006**;22:2688-2690.
15. Hue S, Clewley JP, Cane PA and Pillay D. HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. *AIDS* **2004**;18:719-28.
16. Rieder P, Joos B, von Wyl V, et al. HIV-1 transmission after cessation of early antiretroviral therapy among men having sex with men. *AIDS*;24:1177-83.
17. Perrin L, Kaiser L and Yerly S. Travel and the spread of HIV-1 genetic variants. *Lancet Infect Dis* **2003**;3:22-7.

18. Xiridou M, van Veen M, Coutinho R and Prins M. Can migrants from high-endemic countries cause new HIV outbreaks among heterosexuals in low-endemic countries? *AIDS*;24:2081-8.

## Tables

**Table 1:** Characteristics of SHCS study participants infected with the four most frequently observed non-B subtypes

	CRF01_AE	CRF02_AG	A	C	p-value
N	258 (100)	285 (100)	328 (100)	272 (100)	
Female sex	92 (35.7)	168 (58.9)	213 (64.9)	158 (58.1)	<0.001
<u>Ethnicity</u>					
White	158 (61.2)	93 (32.6)	146 (44.5)	77 (28.3)	<0.001
Black	2 (0.8)	186 (65.3)	170 (51.8)	180 (66.2)	
Hispanic	2 (0.8)	2 (0.7)	1 (0.3)	8 (2.9)	
Asian	96 (37.2)	4 (1.4)	11 (3.4)	7 (2.6)	
<u>Nationality</u>					
Western Europe	157 (60.9)	88 (30.9)	133 (40.5)	82 (30.1)	<0.001
Sub-Saharan Africa	2 (0.8)	174 (61.1)	159 (48.5)	168 (61.8)	
South- and East Asia	84 (32.6)	3 (1.1)	10 (3.0)	4 (1.5)	
Other	15 (5.8)	20 (7.0)	26 (7.9)	18 (6.6)	
Study recruitment in German-speaking Switzerland (vs. French-speaking) <sup>b</sup>	232 (89.9)	156 (54.7)	235 (71.6)	154 (56.6)	<0.001
<u>Mode of HIV acquisition</u>					
Heterosexual contacts	189 (73.3)	269 (94.4)	302 (92.1)	247 (90.8)	<0.001
Intravenous drug use	9 (3.5)	2 (0.7)	12 (3.7)	9 (3.3)	
Homosexual contacts	60 (23.3)	14 (4.9)	14 (4.3)	16 (5.9)	
Year of HIV diagnosis (median [interquartile range])	2002 [1998 to 2005]	2003 [2000 to 2005]	1999 [1995 to 2003]	2000 [1997 to 2004]	<0.001
Documented seroconversion or HIV diagnosis during primary HIV infection <sup>a</sup>	36 (14.0)	42 (14.7)	33 (10.1)	32 (11.8)	0.294
Viral sequence obtained while antiretroviral therapy naive	215 (83.3)	231 (81.1)	206 (62.8)	176 (64.7)	<0.001
First measured CD4 count (median [interquartile range])	319 [94 to 488]	284 [168 to 460]	283 [154 to 490]	283 [152 to 510]	0.882

All numbers are N (%) unless otherwise stated. Tests of significance for differences across subtypes were done with the Chi2-test for categorical variables and the Mann-Whitney U test for continuous variables.

<sup>a</sup> negative and positive HIV test < 1 year apart or clinical diagnosis of acute retroviral syndrome

<sup>b</sup> Individuals from Italian-speaking Switzerland (<4%) were assigned to the German-speaking group due to closer geographical proximity

**Table 2:** Characterization of sequences included in Swiss-specific subepidemics

	CRF01_AE			CRF02_AG			A			C		
	Cluster	No cluster	p-value	Cluster	No cluster	p-value	Cluster	No cluster	p-value	Cluster	No cluster	p-value
N	61 (100)	197 (100)		80 (100)	205 (100)		69 (100)	259 (100)		44 (100)	228 (100)	
Female sex	16 (26.2)	76 (38.6)	0.078	41 (51.3)	127 (62)	0.099	39 (56.5)	174 (67.2)	0.099	17 (38.6)	141 (61.8)	0.004
<u>Ethnicity</u>			0.044			0.001			0.040			<0.001
White	40 (65.6)	118 (59.9)		39 (48.8)	54 (26.3)		36 (52.2)	110 (42.5)		24 (54.5)	53 (23.2)	
Black	0 (0)	2 (1)		38 (47.5)	148 (72.2)		32 (46.4)	138 (53.3)		19 (43.2)	161 (70.6)	
Hispanic	2 (3.3)	0 (0)		1 (1.3)	1 (0.5)		1 (1.4)	0 (0)		1 (2.3)	7 (3.1)	
Asian	19 (31.1)	77 (39.1)		2 (2.5)	2 (1)		0 (0)	11 (4.2)		0 (0)	7 (3.1)	
<u>Nationality</u>			0.694			0.003			0.207			<0.001
Western Europe	37 (60.7)	120 (60.9)		36 (45)	52 (25.4)		30 (43.5)	103 (39.8)		22 (50)	60 (26.3)	
Sub-Saharan Africa	0 (0)	2 (1)		36 (45)	138 (67.3)		31 (44.9)	128 (49.4)		15 (34.1)	153 (67.1)	
South- and East Asia	19 (31.1)	65 (33)		2 (2.5)	1 (0.5)		0 (0)	10 (3.9)		0 (0)	4 (1.8)	
Other	5 (8.2)	10 (5.1)		6 (7.5)	14 (6.8)		8 (11.6)	18 (6.9)		7 (15.9)	11 (4.8)	
<u>Mode of HIV acquisition</u>			<0.001			0.311			0.439			0.078
Heterosexual contacts	30 (49.2)	159 (80.7)		74 (92.5)	195 (95.1)		64 (92.8)	238 (91.9)		36 (81.8)	211 (92.5)	
Intravenous drug use	3 (4.9)	6 (3)		0 (0)	2 (1)		1 (1.4)	11 (4.2)		3 (6.8)	6 (2.6)	
Homosexual contacts	28 (45.9)	32 (16.2)		6 (7.5)	8 (3.9)		4 (5.8)	10 (3.9)		5 (11.4)	11 (4.8)	
Study recruitment in German-speaking Switzerland (vs. French-speaking) <sup>b</sup>	56 (91.8)	176 (89.3)	0.577	42 (52.5)	114 (55.6)	0.636	50 (72.5)	185 (71.4)	0.865	22 (50)	132 (57.9)	0.333
Documented seroconversion or HIV diagnosis during primary HIV infection <sup>a</sup>	14 (23)	22 (11.2)	0.020	19 (23.8)	23 (11.2)	0.007	8 (11.6)	25 (9.7)	0.634	13 (29.5)	19 (8.3)	<0.001
Viral sequence obtained while antiretroviral therapy naive	52 (85.2)	163 (82.7)	0.646	69 (86.3)	162 (79)	0.162	43 (62.3)	163 (62.9)	0.925	32 (72.7)	144 (63.2)	0.224
Year of HIV diagnosis	2003 [2001-2006]	2001 [1997-2005]	0.010	2003.5 [1999.5-2006]	2003 [2000-2005]	0.563	2000 [1996-2003]	1999 [1995-2003]	0.591	2002 [1997-2004.5]	2000 [1997-2003.5]	0.164

All numbers are N (%) unless otherwise stated. Tests of significance were done with the Chi2-test for categorical variables and the Mann-Whitney U test for continuous variables.

<sup>a</sup> negative and positive HIV test < 1 year apart or clinical diagnosis of acute retroviral syndrome

<sup>b</sup> Individuals from italian-speaking Switzerland (<4%) were assigned to the german-speaking group due to closer geographical proximity



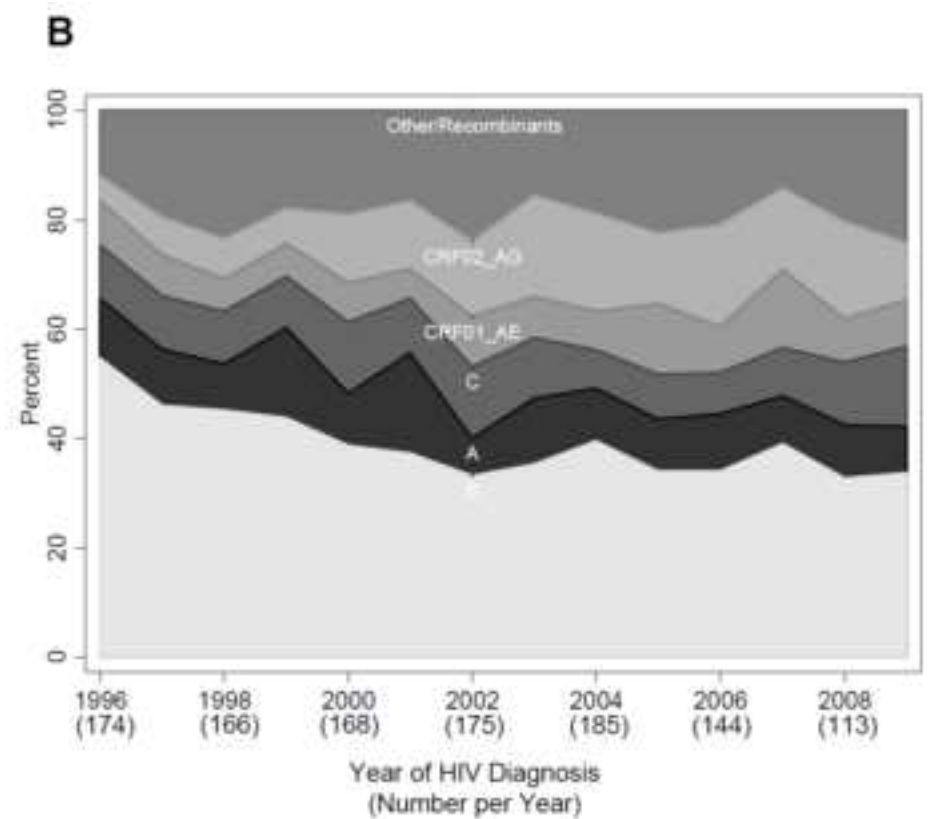
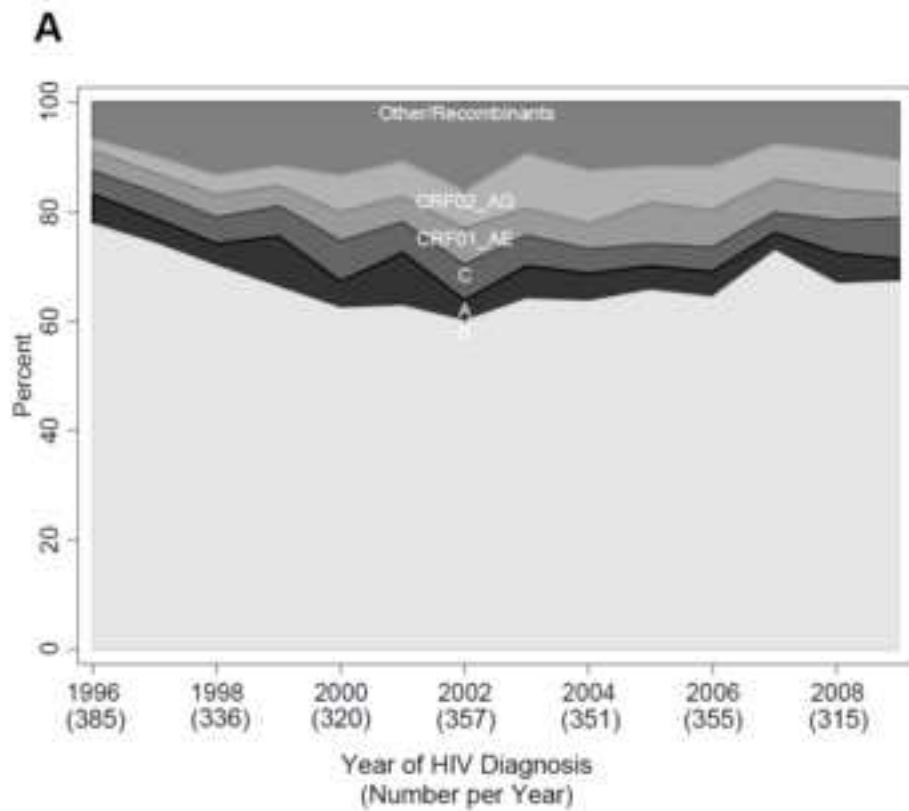
## Figure Legends

Figure 1: Time trends in distribution of subtypes in the SHCS. Panel A: All modes of HIV transmission (n=4767); Panel B: Heterosexual contacts only (n=2197). Graphs for intravenous drug users (low numbers) and homosexual men (no discernible time trend) are not shown.

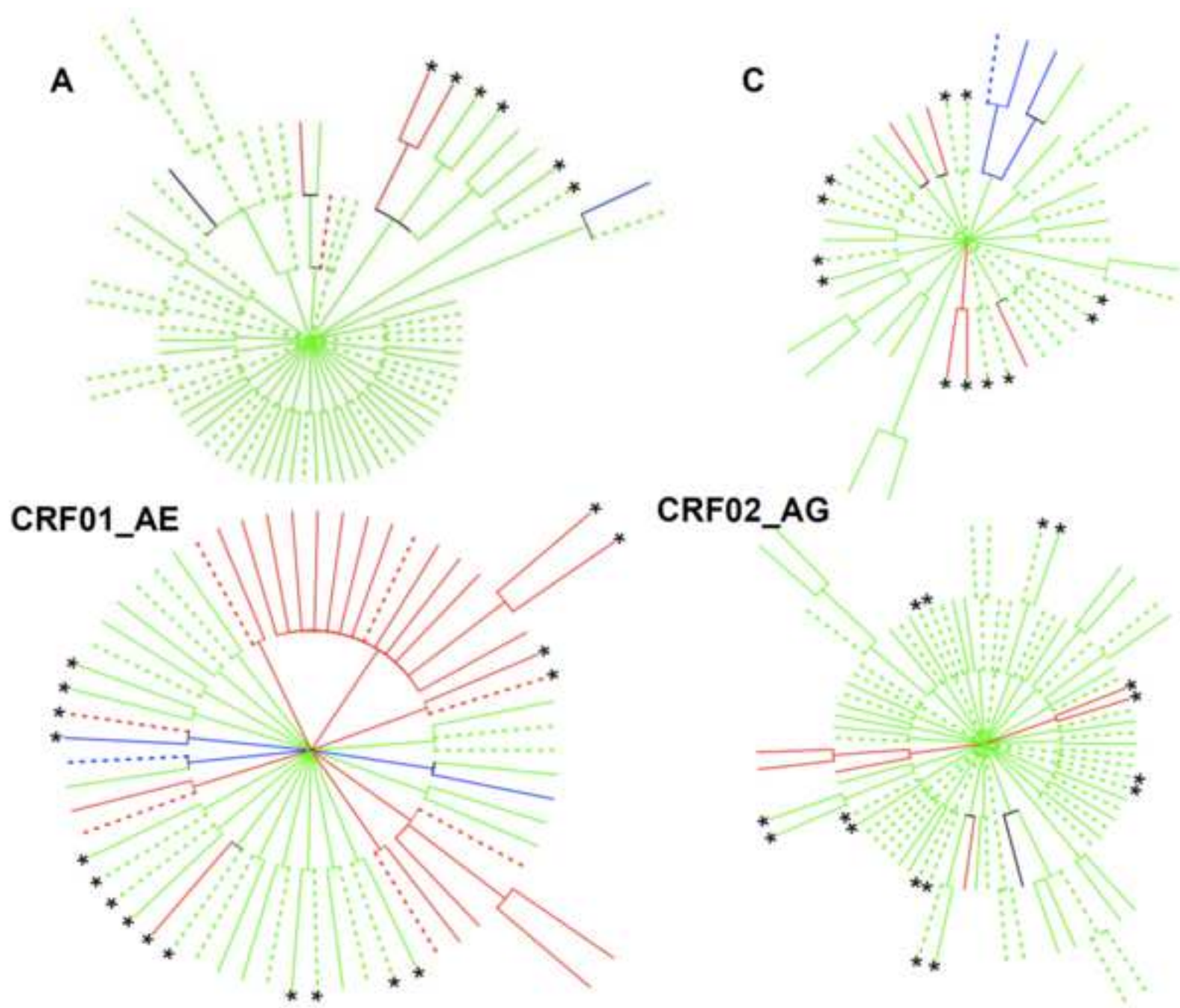
Figure 2: Swiss-specific subepidemics for subtypes A and E and CRFs AE and AG. Only tips belonging to Swiss patients in a Swiss transmission cluster are depicted. Each edge emerging from the center corresponds to one transmission cluster. Colors indicates the transmission group (green: heterosexual, red: MSM, blue: IDU, turquoise: unknown transmission group). Dashed lines indicate patients of non-white ethnicity. Stars indicate possible transmission pairs

Figure 3: Distribution of demographic groups (defined on sex, ethnicity, and mode of HIV acquisition) among Swiss-specific subepidemics (macro-scale analysis, panel A) and within possible transmission pairs (micro-scale analysis, panel B). Note that in panel A only demographic groups with a proportion  $\geq 5\%$  are plotted. Mode of HIV acquisition is heterosexual if not mentioned otherwise. Abbreviations: f, female; m, male; msm; men who have sex with men.

**Figure 1**  
[Click here to download high resolution image](#)



**Figure 2**  
[Click here to download high resolution image](#)



**Figure 3**  
[Click here to download high resolution image](#)

