

## **A Phylogenetic Analysis of HIV-1 Sequences in Kiev: Findings among Key Populations**

Esther Fearnhill<sup>1</sup>, Annabelle Gourlay<sup>1</sup>, Ruslan Malyuta<sup>2</sup>, Ruth Simmons<sup>1</sup>, R. Bridget Ferns<sup>1</sup>, Paul Grant<sup>3</sup>, Eleni Nastouli<sup>1,2</sup>, Iryna Karnets<sup>4</sup>, Gary Murphy<sup>5</sup>, Antonia Medoeva<sup>4</sup>, Yuri Kruglov<sup>6</sup>, Alexander Yurchenko<sup>4</sup>, Kholoud Porter<sup>1</sup> on behalf of CASCADE Collaboration in EuroCoord\*

1. University College London, London,
2. Perinatal Prevention of AIDS Initiative, Odessa, Ukraine
3. University College London Hospital NHS Foundation Trust, London
4. Kyiv City AIDS Centre, Kiev, Ukraine
5. Public Health England, London, UK
6. Institute of Epidemiology, Kiev, Ukraine

\* Appendix

**Key words:** HIV, Ukraine, MSM, PWID, Phylogenetic

**Running Title:** HIV Transmission Dynamics in Kiev

**Correspondence to:** Esther Fearnhill, Department of Infection and Population Health, Faculty of Population Health Sciences, University College London, Mortimer Market Centre, off Capper Street, London WC1E 6JB. Telephone: [+44 \(0\)20 7679 7611](tel:+442076797611). Email: [e.fearnhill@ucl.ac.uk](mailto:e.fearnhill@ucl.ac.uk)

**Alternate corresponding author:** Dr Annabelle Gourlay, Department of Infection and Population Health, Faculty of Population Health Sciences, University College London, Mortimer Market Centre, off Capper Street, London WC1E 6JB. Telephone: [+44 \(0\)20 7679 7611](tel:+442076797611). Email: [a.gourlay@ucl.ac.uk](mailto:a.gourlay@ucl.ac.uk)

**Summary:** Transmission dynamics of HIV-1 in Kiev were investigated using phylogenetic analysis of pol sequences from recently diagnosed individuals. This revealed bridging between the three key populations, and evidence that the sexually transmitted epidemic in Kiev is becoming self-sustaining.

## Abstract

### BACKGROUND

The HIV epidemic in Ukraine has been driven by a rapid rise among people who inject drugs, but recent studies have shown an increase through sexual transmission.

### METHODS

Protease and RT sequences from 876 new HIV diagnoses (April 2013 – March 2015) in Kiev were linked to demographic data. We constructed phylogenetic trees for 794 subtype A1 and 64 subtype B sequences and identified factors associated with transmission clustering. Clusters were defined as  $\geq 2$  sequences,  $\geq 80\%$  local branch support and maximum genetic distance of all sequence pairs in the cluster  $\leq 2.5\%$ . Recent infection was determined through the LAg avidity EIA assay. Sequences were analysed for transmitted drug resistance (TDR) mutations.

### RESULTS

30% of subtype A1 and 66% of subtype B sequences clustered. Large clusters (maximum 11 sequences) contained mixed risk groups. In univariate analysis, clustering was significantly associated with subtype B compared to A1 (OR 4.38 [95% CI 2.56-7.50]), risk group (OR 5.65 [3.27-9.75]) for men who have sex with men compared to heterosexual males, recent, compared to long-standing, infection (OR 2.72 [1.64-4.52]), reported sex work contact (OR 1.93 [1.07-3.47]) and younger age groups compared to age  $\geq 36$  (OR 1.83 [1.10-3.05] for age  $\leq 25$ ). Females were associated with lower odds of clustering than heterosexual males (OR 0.49 [0.31-0.77]). In multivariate analysis, risk group, subtype and age group were independently associated with clustering ( $p < 0.001$ ,  $p = 0.007$  and  $p = 0.033$ ). 18 sequences (2.1%) indicated evidence of TDR.

### CONCLUSIONS

Our findings suggest high levels of transmission and bridging between risk groups.

## INTRODUCTION

With an estimated 238,000 people living with HIV/AIDS, and AIDS-related annual mortality of around 17,000, Ukraine is experiencing one of the most severe HIV epidemics in Europe (1).

The initial epidemic followed a rapid rise of illegal drug use after the breakup of the Soviet Union. By 1996 there were 12,000 HIV infections diagnosed a year, mainly in people who inject drugs (PWID), but from late 1990s sexually-transmitted infections, and infections in children, increased (2, 3). Since 2006, new diagnoses among PWID fell, and from 2008 the majority have been through sexual contacts, most believed to be linked to PWID (1, 4, 5). Diagnoses among men who have sex with men (MSM) are low in official statistics but other studies have shown these figures to be under-estimated (6-8).

Previous studies recommend interventions targeting bridge populations (sexual partners of PWID, clients of sex workers, and female partners of MSM), responsible for linking high risk groups with the general population (1, 9-11). The prevention programs delivered through the International HIV/AIDS Alliance in Ukraine since 2006 have increased opioid substitution therapy, HIV testing and counselling, access to health services and antiretroviral therapy (ART) coverage, but continued efforts are required to reach WHO targets (1). Currently there is no government support for prevention programs and financial resources are limited due to ongoing armed conflict in parts of the country.

We examined HIV transmission dynamics in Ukraine using viral sequences. Previous studies have used phylogenetic analysis of HIV-1 sequences from Ukraine and other former Soviet Union (FSU) countries to describe the molecular epidemiology of HIV-1 in Ukraine from earlier years (2, 12, 13-16). Ours is the first using phylogenetics to examine new diagnoses and recently-acquired infection, and is unique in having a large number of HIV-1 sequences with detailed demographic data.

Sequences obtained from samples collected in the Kiev region between April 2013 and March 2015 were analysed to identify clusters of closely-related infections, and to

characterise factors associated with high levels of ongoing transmission. We examined bridge groups by describing the association and separation of risk groups within transmission clusters.

## **METHODS**

### **Study Population and Data Collection**

This analysis used data collected in a larger study within the CASCADE Collaboration in EuroCoord ([www.EuroCoord.net](http://www.EuroCoord.net)), which estimated HIV incidence in Kiev. The study population and data collection methods were described in detail in earlier publications (17, 18). Briefly, data were collected from all persons over the age of 16 who presented or were referred for an HIV test at any of the four infectious disease clinics in the Kiev province (known collectively as Kiev City AIDS Centre), between 1<sup>st</sup> April 2013 and 29<sup>th</sup> March 2015. Tests from routine screening programmes, for example antenatal screening and blood donation, were excluded due to different incidence estimation methods in these programs. Tests from general screening, for example, surgical interventions, army recruitment, prisoners and visa applications were included.

Data collected from an anonymous questionnaire included year of birth, area of residence, sex, reason for test, risk factors for HIV including likely route of infection, and testing history.

A number of individuals with long-standing HIV infections, sampled to assess the false recent rate of the recent testing algorithm (RITA) in the main study, were included for phylogenetic robustness but did not complete the survey and were not included in the analysis of factors associated with clustering.

### **Laboratory Methods**

HIV positive residual blood samples were tested for recent infection, using the limiting antigen (LAg) avidity EIA assay at Public Health England. For those classified as recent (infections likely to have occurred within the last six months), HIV viral load (VL) was quantified, and those with low VL (<1000 copies/mL) were reclassified as longstanding. Methods are described in detail by Simmons et al (17).

In HIV pol gene sequencing, nucleic acid extraction followed by nested PCR (outer forward and reverse primers AATGATGACAGCATGYCAGGGAGT and

AGTCTTTCCCATATTACTATGCTTTC, inner forward and reverse primers  
GGAAAAGGGCTGTTGGAAATGTG and GGCTCTTGATAAATTTGATATGTCCAT)  
amplified a pol gene product which was then sequenced on Sanger or Illumina Miseq  
platforms.

### **Phylogenetic Methods**

Sequences were aligned using MPAlign (19) and positions 1-1320 retained. Rega3  
subtyping was performed (20). Sequences were examined for transmitted drug resistance  
(TDR), according to the World Health Organization (WHO) 2009 surveillance list (21).

Maximum-likelihood trees were created in Fasttree 2.1 (22) for subtypes A1 and B with the  
General Time Reversal method of nucleotide substitution (GTR), which excluded nucleotide  
ambiguity codes. 1130 subtype A, and 230 subtype B sequences from FSU countries  
between 2010 and 2014, were included from the Los Alamos Database to increase  
phylogenetic robustness.

Analysis of clusters was performed using programs written in R. Fasttree SH local branch  
support values of  $\geq 80\%$ , and a maximum genetic distance of 2.5% between all sequence  
pairs were used to identify clusters. Nucleotide ambiguity codes were included only if they  
generated a definite change (e.g. A to S). The 2.5% threshold was calculated using a  
modified version of an approach described by Aldous et al (23) as follows: we created 1000  
randomly generated clusters, sizes 2-50, from the combined FSU and Kiev sequences, and  
calculated the maximum pairwise genetic distance in each cluster. The distance of 2.5% fell  
below the 0.01 percentile of the distribution of these values, making it highly unlikely that  
clusters with a maximum genetic distance of this value or below would occur by chance in  
phylogenetically unrelated sequences.

### **Data Preparation/Management**

Survey data were linked to laboratory results. Age at diagnosis was grouped in four bands:  
 $\leq 25$ , 26-30, 31-35 and  $\geq 36$  years. Risk group was based on survey questions (have you ever

injected drugs, had sex with a person of the opposite sex, had sex with someone of the same sex, paid for sex, been paid for sex) and categorised in the following hierarchical order (if more than one was reported): PWID, MSM or heterosexual contact. 'Sex worker/contact with sex worker' was based on responses to 'ever paid for' and/or 'ever been paid for sex'. Reason for test was categorised as either clinical indication (symptoms), high risk group (has injected drugs, has multiple partners, had a sexually transmitted infection (STI), had a needle stick injury, had contact with an HIV-positive person, suspects they are HIV-positive), or general screening (described above). A hierarchy was applied as follows if multiple reasons were recorded: screening, clinical indication, high risk group. Repeat testing was defined as having at least one prior HIV test result more than one month before the current test.

### **Statistical Analysis of factors associated with clustering**

All sequences were categorised as clustering or not clustering. To assess associations with clustering, we performed univariate and multivariate logistic regression analysis, controlling for the following variables: sex, age group, subtype, assigned risk group, reported sex work/contact with sex worker, repeated testing history, recent or longstanding infection, and reason for test. All variables which were significantly associated with the outcome ( $p \leq 0.05$ ) in univariate analyses were assessed in the multivariate models. We used a forwards stepwise approach so that variables were retained in the model if they significantly improved the fit of the model ( $p \leq 0.05$ , based on likelihood ratio tests). To avoid collinearity, sex and the assigned risk group variables are reported under one variable, categorised as MSM, Heterosexual Contact: Male, Heterosexual Contact: Female, PWID: Male, PWID: Female.

Clusters were assigned to a group, based on reported route of exposure of over 60% of samples in the cluster belonging to a particular risk group or combination of risk groups, in the following hierarchical way; 1) MSM 2) Heterosexual contact: Male and female, 3) PWID: Male and female, 4) Male heterosexual alone or with MSM 5) Heterosexual contact and PWID mixed. Groups with no data on  $\geq 40\%$  of samples were recorded as missing. Using a



method described by Poon et al (24), we also examined the association and separation of different risk groups within clusters: multiple scatterplot diagrams of all clusters were created, where each point represented a cluster, with the area of the marker scaled to cluster size. The x and y axis denoted pairs of risk groups, and the location of the marker showed the proportion of samples in the cluster belonging to each of the risk groups.

## RESULTS

### Study Population Characteristics

During the study period, 1192 persons tested positive for HIV in the Kiev region. Sequences were obtained from 898 of these samples (294 samples were not received by the laboratory, or failed to amplify/contained insufficient sample volume). Sample pairs with identical sequences (after removing resistance positions) were manually checked for matching demographic data, and 22 samples were identified as duplicates and excluded.

Among the remaining 876 samples, most (794, 91%) were subtype A1 and 64 (7%) were subtype B; the remaining 18 samples comprised a mixture of subgroups 02\_AG, G, 03\_AB and other recombinants. Only the 858 subtype A1 and B samples were included in the analysis (Table 1). Of these, 54 samples could not be linked to survey data and had no demographic data. A further 41 samples were missing data on one or more of the following variables; risk group, recent infection, repeat tester, sex work/contact with sex worker and reason for test.

The majority of individuals were male (71%) with assigned source of infection being heterosexual sex (52%). 37% were assigned as PWID, the majority male (83%), and 11% were assigned as MSM. Of 794 samples with risk group data, 48 (6%) reported that they had been paid, or had paid for sex. The median age was 33, IQR: 29-38. We found a significant difference in risk group category between the 876 sequenced and 294 un-sequenced samples. Sequenced data had a higher proportion of MSM (11% compared to 6%) and heterosexual males (29% compared to 20%) and a lower proportion of female PWID (7% compared to 13%),  $p < 0.001$ .

From 795 samples with results, 67 (8%) were identified as recent infections. Individuals assigned as MSM were more likely to have a recent infection compared to individuals assigned to heterosexual transmission (OR 4.81 [95% CI 2.69-8.60]  $p < 0.001$ ) while PWID

were less likely to have a recent infection compared to heterosexuals (OR 0.39 [95% CI 0.18-0.83] p = 0.015) in a univariate logistic regression model.

The majority of individuals reported the reason for test as either high risk behaviours (47%) or clinical indication (47%).

### **Drug Resistance Mutations**

18/858 (2.1%) samples contained drug resistance mutations, indicating TDR, as follows; 3 with non-nucleoside reverse transcriptase inhibitor (NNRTI) resistance, 4 with nucleoside reverse transcriptase inhibitor (NRTI), 5 with NRTI and NNRTI resistance, 1 with NRTI and protease inhibitor (PI) resistance, and 5 with only PI resistance.

### **Factors Associated With Clustering**

A third of the sequences appeared in clusters of size 2 or above. Clustering was higher in subtype B, with 42/64 sequences clustering (66%) compared to 241/794 (30%) in subtype A1. This additional subtype B clustering was mainly among heterosexual males, with 69% clustering compared to 30% in subtype A1.

In univariate analyses (Table 1), clustering was significantly associated with the following: subtype B compared to subtype A1 (OR 4.38 [95% CI 2.56-7.50]), younger age groups compared to over 35s (OR 1.83 [1.10-3.05], for age  $\leq 25$  and OR 2.09 [1.42—3.07] for age 26-30), MSM compared to heterosexual males (OR 5.65 [3.27-9.75]), recent compared to long-standing infection (OR 2.72 [1.64-4.52]), and reported sex work or contact with a sex worker (OR 1.93 [1.07-3.47]). Testing due to clinical indication, compared to high risk behaviours also showed higher levels of clustering although this was not significant at the 5% level. Lower levels of clustering were seen in heterosexual females compared to heterosexual males (OR 0.49 [0.31-0.77]), but not in female compared to male PWIDs (OR 0.89 [0.45-1.77]).

In the multivariate analysis, subtype, age group, and risk group remained independently associated with clustering. As well as MSM, female heterosexuals remained significantly less likely to cluster compared to heterosexual males (adjusted OR (aOR) 0.51 [0.32-0.81]). When separating sex and risk group, both remained independently associated with clustering, with males more likely to cluster than females (aOR 1.94 [1.27-2.95]) and MSM more likely to cluster than heterosexuals (aOR 4.31 [2.40-7.75]).

### **Cluster Analysis**

The 283 samples which clustered formed 93 subtype A1 (Fig 1) and 13 subtype B (Fig 2) clusters. Almost half (48%) clustered in pairs. The remainder formed 35 clusters of sizes 3-5, and three larger clusters; one of 6 subtype B sequences, and two subtype A1 clusters of 9 and 11 sequences. Over a third of clusters, including the largest three, were predominantly MSM, or MSM and heterosexual male. Categories PWID and PWID and heterosexual combined accounted for 38 clusters, mainly pairs, and 18 clusters (17 pairs) were classified as heterosexual. The remainder (6 clusters) were missing data (Fig 3).

The scatterplots (Fig 4) showed that association of risk groups within clusters was greatest between MSM and heterosexual male, followed by heterosexual and PWID risk groups. Separation was greatest between MSM and samples from females.

There was no evidence of clustering among sequences containing drug resistance mutations.

## DISCUSSION

In this analysis, we have characterised recent transmission patterns in Kiev, showing high levels of clustering, in well-supported branches with low intra-cluster variability.

It has been widely reported that the epidemic among PWID in Ukraine is shifting, with increasing numbers of infections being attributable to sexual contact (10, 11, 25). Our findings support this, with lower levels of clustering in PWID compared to heterosexual males, and PWID less likely to have a recent infection than individuals reporting heterosexual contact (17), although levels of clustering in heterosexual males may be overstated due to non-disclosure from MSM or PWID.

Clustering in heterosexual females was significantly lower than that in heterosexual males, and the majority appeared in small clusters, with male PWID or other female heterosexuals. High numbers of females have previously reported sexual contact with PWID (18), and while clustering is lower in PWID than other risk groups, there is still evidence of bridging between male PWID and their female partners. PWID remain a hard to reach group in Ukraine, with the criminal nature of drug taking leading to non-disclosure and interventions reaching only a small proportion of estimated numbers (26).

Clustering among MSM was strikingly high, with the majority of clusters, including the largest three, containing mainly MSM, or combinations of MSM and heterosexual males. It is likely that this reflects high levels on ongoing transmission among MSM. A 2006 epidemiological review estimated Ukraine to have around 27% HIV prevalence among MSM compared to less than 2% in other adults (27), although official estimates are considerably less and may reflect under-reporting (28). Recent MSM HIV incidence estimates in Kiev are between 2290 and 6869 per 100,000 compared to 21.5 per 100,000 in the general population (17). There are also high levels of recent infection among MSM with approximately one in four new diagnoses in Odessa and Kiev shown to be recently infected, and MSM disproportionately represented in both new diagnoses and recent infection (17, 18, 29). A 2011 bio-behavioural

study of MSM in Ukraine showed low HIV testing rates, limited access to prevention programs, and high levels of sexual mixing with female heterosexuals, and with sex workers of both sexes (30).

Subtype B clustering was higher than subtype A1 after accounting for risk group. The subtype B variant in Ukraine (IDU-B) was less epidemiologically successful than A1, and although it remains prevalent, has not spread widely to other countries of the FSU (3).

Subtype B samples were also less likely to cluster with background FSU samples compared to subtype A1. This suggests greater coverage of subtype B in our data which could explain some additional clustering. Being mainly MSM and heterosexual male, subtype B samples may include more recent transmissions than subtype A1. However, as all of the additional clustering in subtype B was among heterosexual males, it may be that the main contributory factor is high levels of undisclosed MSM in this group.

Younger age of sexual activity has been associated with an increased risk of HIV infection (27) and we found higher levels of clustering in younger age groups. A recent study in Odessa showed a higher risk of recent infection among younger people (29) and younger adults in our data were more likely to be recently-infected (17)

We found low levels of TDR (2.1%), which would be expected as estimated ART coverage in Ukraine is low at 22% (26). Previous TDR estimates have been higher (31, 32) but have used different definitions of TDR mutations from ours, in particular the inclusion of accessory NRTI mutation A62V, a reported polymorphism in non-B subtypes. (33, 34).

There are a number of limitations to consider. As this data is from relatively recent samples, data coverage is not complete, and sampling density may be markedly different for different populations within the dataset due to variable testing patterns. Much of the data is self-reported, and there may be significant levels of non-disclosure, particularly among MSM, indicated by high levels of clustering between MSM and males reporting heterosexual contact. In addition undiagnosed HIV rates are likely to be high and may vary between

populations. Our data are unlikely to be representative of the whole of Ukraine. Transmission patterns may vary in other areas of the country as there is considerable variability in HIV prevalence and population size of different risk groups (35). Differences between sequenced and un-sequenced data is a potential source of bias. Both MSM and heterosexual males were overrepresented in the sequenced data which may have improved our ability to detect clustering in these groups.

This study supports previous reports of the transition of the HIV-1 epidemic in Ukraine from PWID to the non-drug using population, through the sexual partners of PWID and MSM. High levels of recent infections in MSM, high clustering among MSM, and bridging between MSM and reported heterosexuals in recently-diagnosed people suggest that this sexually-transmitted epidemic may have become self-sustaining. This highlights the need for continued efforts to reach the MSM and PWID populations, and their partners, with targeted interventions to reduce the risk of onwards transmission.

## Appendix

**CASCADE Steering Committee:** Julia Del Amo (Chair), Laurence Meyer (Vice Chair), Heiner C. Bucher, Geneviève Chêne, Osamah Hamouda, Deenan Pillay, Maria Prins, Magda Rosinska, Caroline Sabin, Giota Touloumi.

**CASCADE Co-ordinating Centre:** Kholoud Porter (Project Leader), Ashley Olson, Andrea Cartier, Lorraine Fradette, Sarah Walker, Abdel Babiker.

**CASCADE Clinical Advisory Board:** Heiner C. Bucher, Andrea De Luca, Martin Fisher, Roberto Muga

**CASCADE Collaborators:** Australia PHAEDRA cohort (Tony Kelleher, David Cooper, Pat Grey, Robert Finlayson, Mark Bloch) Sydney AIDS Prospective Study and Sydney Primary HIV Infection cohort (Tony Kelleher, Tim Ramacciotti, Linda Gelgor, David Cooper, Don Smith); Austria Austrian HIV Cohort Study (Robert Zangerle); Canada South Alberta clinic (John Gill); Estonia Tartu Ülikool (Irja Lutsar); France ANRS CO3 Aquitaine cohort (Geneviève Chêne, Francois Dabis, Rodolphe Thiebaut), ANRS CO4 French Hospital Database (Dominique Costagliola, Marguerite Guiguet), Lyon Primary Infection cohort (Philippe Vanhems), French ANRS CO6 PRIMO cohort (Marie-Laure Chaix, Jade Ghosn), ANRS CO2 SEROCO cohort (Laurence Meyer, Faroudy Boufassa); Germany German HIV-1 seroconverter cohort (Osamah Hamouda, Karolin Meixenberger, Norbert Bannert, Barbara Bartmeyer); Greece AMACS (Anastasia Antoniadou, Georgios Chrysos, Georgios L. Daikos); Greek Haemophilia cohort (Giota Touloumi, Nikos Pantazis, Olga Katsarou); Italy Italian Seroconversion Study (Giovanni Rezza, Maria Dorrucchi), ICONA cohort (Antonella d'Arminio Monforte, Andrea De Luca.) Netherlands Amsterdam Cohort Studies among homosexual men and drug users (Maria Prins, Ronald Geskus, Jannie van der Helm, Hanneke Schuitemaker); Norway Oslo and Ullevål Hospital cohorts (Mette Sannes, Oddbjorn Brubakk, Anne-Marte Bakken Kran); Poland National Institute of Hygiene (Magdalena Rosinska); Spain Badalona IDU hospital cohort (Roberto Muga, Jordi Tor),



Barcelona IDU Cohort (Patricia Garcia de Olalla, Joan Cayla), CoRIS-scv (Julia del Amo, Santiago Moreno, Susana Monge); Madrid cohort (Julia Del Amo, Jorge del Romero), Valencia IDU cohort (Santiago Pérez-Hoyos); Sweden Swedish InfCare HIV Cohort, Sweden (Anders Sönnernborg); Switzerland Swiss HIV Cohort Study (Heiner C. Bucher, Huldrych Günthard, Alexandra Scherrer); Ukraine Perinatal Prevention of AIDS Initiative (Ruslan Malyuta); United Kingdom Public Health England (Gary Murphy), UK Register of HIV Seroconverters (Kholoud Porter, Anne Johnson, Andrew Phillips, Abdel Babiker), University College London (Deenan Pillay); African cohorts: Genital Shedding Study (US: Charles Morrison; Family Health International, Robert Salata, Case Western Reserve University, Uganda: Roy Mugerwa, Makerere University, Zimbabwe: Tsungai Chipato, University of Zimbabwe); International AIDS Vaccine Initiative (IAVI) Early Infections Cohort (Kenya, Rwanda, South Africa, Uganda, Zambia: Matt A. Price, IAVI, USA; Jill Gilmour, IAVI, UK; Anatoli Kamali, IAVI, Kenya; Etienne Karita, Projet San Francisco, Rwanda).

**EuroCoord Executive Board:** Fiona Burns, University College London, UK; Geneviève Chêne, University of Bordeaux, France; Dominique Costagliola (Scientific Coordinator), Institut National de la Santé et de la Recherche Médicale, France; Carlo Giaquinto, Fondazione PENTA, Italy; Jesper Grarup, Region Hovedstaden, Denmark; Ole Kirk, Region Hovedstaden, Denmark; Laurence Meyer, Institut National de la Santé et de la Recherche Médicale, France; Heather Bailey, University College London, UK; Alain Volny Anne, European AIDS Treatment Group, France; Alex Panteleev, St. Petersburg City AIDS Centre, Russian Federation; Andrew Phillips, University College London, UK, Kholoud Porter, University College London, UK; Claire Thorne, University College London, UK.

**EuroCoord Council of Partners:** Jean-Pierre Aboulker, Institut National de la Santé et de la Recherche Médicale, France; Jan Albert, Karolinska Institute, Sweden; Silvia Asandi, Romanian Angel Appeal Foundation, Romania; Geneviève Chêne, University of Bordeaux, France; Dominique Costagliola (chair), INSERM, France; Antonella d'Arminio Monforte, ICoNA Foundation, Italy; Stéphane De Wit, St. Pierre University Hospital, Belgium; Peter

Reiss, Stichting HIV Monitoring, Netherlands; Julia Del Amo, Instituto de Salud Carlos III, Spain; José Gatell, Fundació Privada Clínic per a la Recerca Biomèdica, Spain; Carlo Giaquinto, Fondazione PENTA, Italy; Osamah Hamouda, Robert Koch Institut, Germany; Igor Karpov, University of Minsk, Belarus; Bruno Ledergerber, University of Zurich, Switzerland; Jens Lundgren, Region Hovedstaden, Denmark; Ruslan Malyuta, Perinatal Prevention of AIDS Initiative, Ukraine; Claus Møller, Cadpeople A/S, Denmark; Kholoud Porter, University College London, United Kingdom; Maria Prins, Academic Medical Centre, Netherlands; Aza Rakhmanova, St. Petersburg City AIDS Centre, Russian Federation; Jürgen Rockstroh, University of Bonn, Germany; Magda Rosinska, National Institute of Public Health, National Institute of Hygiene, Poland; Manjinder Sandhu, Genome Research Limited; Claire Thorne, University College London, UK; Giota Touloumi, National and Kapodistrian University of Athens, Greece; Alain Volny Anne, European AIDS Treatment Group, France.

**EuroCoord External Advisory Board:** David Cooper, University of New South Wales, Australia; Nikos Dedes, Positive Voice, Greece; Kevin Fenton, Public Health England, USA; David Pizzuti, Gilead Sciences, USA; Marco Vitoria, World Health Organisation, Switzerland.

**EuroCoord Secretariat:** Silvia Faggion, Fondazione PENTA, Italy; Lorraine Fradette, University College London, UK; Richard Frost, University College London, UK; Andrea Cartier, University College London, UK; Dorthe Raben, Region Hovedstaden, Denmark; Christine Schwimmer, University of Bordeaux, France; Martin Scott, UCL European Research & Innovation Office, UK.

**Funding:** This work was supported by the European Union Seventh Framework Programme (FP7/2007-2013) under EuroCoord [grant agreement number 260694].

RBF received funding related to this study from the NIHR Biomedical Research Centre and the UCLH/UCL BRC funded NIHR Health Informatics Collaborative study.

### **Acknowledgements**

Assistance in data retrieval was provided by Lorraine Fradette, University College London, and Andrea Cartier, University College London.

## References

1. World Health Organisation. Good practices in Europe: HIV prevention for People Who Inject Drugs implemented by the International HIV/AIDS Alliance in Ukraine. **2014**.
2. Saad MD, Shcherbinskaya AM, Nadai Y, et al. Molecular epidemiology of HIV Type 1 in Ukraine: birthplace of an epidemic. *AIDS Research and Human Retroviruses* **2006**; 22(8): 709-14.
3. Bobkova M. Current status of HIV-1 diversity and drug resistance monitoring in the former USSR. *AIDS Reviews* **2013**; 15(4): 204-12.
4. Malyuta R. Curbing HIV incidence in people who inject drugs. *The Lancet HIV* **2016**; 3(10):e453-4.
5. Des Jarlais DC, Feelemyer JP, Modi SN, et al. Transitions from injection-drug-use-concentrated to self-sustaining heterosexual HIV epidemics: patterns in the international data. *PloS One* **2012**; 7(3): e31227.
6. Marcus U, Hickson F, Weatherburn P, Schmidt AJ, Network E. Estimating the size of the MSM populations for 38 European countries by calculating the survey-surveillance discrepancies (SSD) between self-reported new HIV diagnoses from the European MSM internet survey (EMIS) and surveillance-reported HIV diagnoses among MSM in 2009. *BMC Public Health* **2013**; 13: 919.
7. UNAIDS. Hidden HIV epidemic amongst MSM in Eastern Europe and Central Asia. Available at: <http://www.unaids.org/en/resources/presscentre/featurestories/2009/january/20090126msmukraine>. Accessed 23/08/2016.
8. Cakalo JI, Bozicevic I, Vitek C, Mandel JS, Salyuk T, Rutherford GW. Misclassification of men with reported HIV infection in Ukraine. *AIDS and Behavior* **2015**; 19(10): 1938-40.

9. Hamers FF, Batter V, Downs AM, Alix J, Cazein F, Brunet JB. The HIV epidemic associated with injecting drug use in Europe: geographic and time trends. *AIDS* **1997**; 11(11): 1365-74.
10. Hamers FF, Downs AM. HIV in central and eastern Europe. *Lancet* **2003**; 361(9362): 1035-44.
11. Mazhnaya A, Andreeva TI, Samuels S, DeHovitz J, Salyuk T, McNutt LA. The potential for bridging: HIV status awareness and risky sexual behaviour of injection drug users who have non-injecting permanent partners in Ukraine. *Journal of the International AIDS Society* **2014**; 17: 18825.
12. Novitsky VA, Montano MA, Essex M. Molecular epidemiology of an HIV-1 subtype A subcluster among injection drug users in the Southern Ukraine. *AIDS Research and Human Retroviruses* **1998**; 14(12): 1079-85.
13. Nabatov AA, Kravchenko ON, Lyulchuk MG, Shcherbinskaya AM, Lukashov VV. Simultaneous introduction of HIV type 1 subtype A and B viruses into injecting drug users in southern Ukraine at the beginning of the epidemic in the former Soviet Union. *AIDS Research and Human Retroviruses* **2002**; 18(12): 891-5.
14. Diez-Fuertes F, Cabello M, Thomson MM. Bayesian phylogeographic analyses clarify the origin of the HIV-1 subtype A variant circulating in former Soviet Union's countries. *Infection, genetics and evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* **2015**; 33: 197-205.
15. Thomson MM, de Parga EV, Vinogradova A, et al. New insights into the origin of the HIV type 1 subtype A epidemic in former Soviet Union's countries derived from sequence analyses of preepidemically transmitted viruses. *AIDS Research and Human Retroviruses* **2007**; 23(12): 1599-604.
16. Riva C, Romano L, Saladini F, et al. Identification of a possible ancestor of the subtype A1 HIV Type 1 variant circulating in the former Soviet Union. *AIDS Research and Human Retroviruses* **2008**; 24(10): 1319-25.

17. Simmons R, Malyuta R, Chentsova N, et al. HIV Incidence Estimates Using the Limiting Antigen Avidity EIA Assay at Testing Sites in Kiev City, Ukraine: 2013-2014. *PloS One* **2016**; 11(6): e0157179.
18. Simmons R, Malyuta R, Chentsova N, et al. HIV Testing and Diagnosis Rates in Kiev, Ukraine: April 2013-March 2014. *PloS One* **2015**; 10(8): e0137062.
19. Gaschen B, Kuiken C, Korber B, Foley B. Retrieval and on-the-fly alignment of sequence fragments from the HIV database. *Bioinformatics* **2001**; 17(5): 415-8.
20. Pineda-Pena AC, Faria NR, Imbrechts S, et al. Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: performance evaluation of the new REGA version 3 and seven other tools. *Infection, genetics and evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* **2013**; 19: 337-48.
21. Bennett DE, Camacho RJ, Otelea D, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. *PloS One* **2009**; 4(3): e4724.
22. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. *PloS One* **2010**; 5(3): e9490.
23. Aldous JL, Pond SK, Poon A, et al. Characterizing HIV transmission networks across the United States. *Clinical Infectious Diseases* **2012**; 55(8): 1135-43.
24. Poon AF, Joy JB, Woods CK, et al. The impact of clinical, demographic and risk factors on rates of HIV transmission: a population-based phylogenetic analysis in British Columbia, Canada. *The Journal of Infectious Diseases* **2015**; 211(6): 926-35.
25. Kelly JA, Amirkhanian YA. The newest epidemic: a review of HIV/AIDS in Central and Eastern Europe. *International journal of STD & AIDS* **2003**; 14(6): 361-71.
26. World Health Organisation. HIV/AIDS treatment and care in Ukraine: Evaluation report., **April 2013**.
27. Beyrer C. HIV epidemiology update and transmission factors: risks and risk contexts--16th International AIDS Conference epidemiology plenary. *Clinical Infectious Diseases* **2007**; 44(7): 981-7.

28. European Centre for Disease Prevention and Control, Europe. WROf. HIV/AIDS surveillance in Europe 2014. Stockholm: ECDC. **2015**.
29. Simmons R, Semenenko I, Tolpina M, et al. High percentage of recent HIV infection leading to onward transmission in Odessa, Ukraine associated with young adults. *AIDS and Behavior* **2014**; 18(2): 411-8.
30. Bolshov YS KM, Leshchynskyi YB, Trofymenko LV, Shvab IA. Analytical report: behaviour monitoring and HIV prevalence among men who have sex with men as a component of second generation surveillance based on results of the biobehavioural survey of 2011. Kyiv: International HIV/AIDS Alliance in Ukraine, **2012**. Available: [http://www.aidsalliance.org.ua/ru/library/our/2012/me/msm\\_en\\_2011.pdf](http://www.aidsalliance.org.ua/ru/library/our/2012/me/msm_en_2011.pdf) Accessed 15 Mar, 2017.
31. Chan PA, Kantor R. Transmitted drug resistance in nonsubtype B HIV-1 infection. *HIV Therapy* **2009**; 3(5): 447-65.
32. Vazquez de Parga E, Rakhmanova A, Perez-Alvarez L, et al. Analysis of drug resistance-associated mutations in treatment-naive individuals infected with different genetic forms of HIV-1 circulating in countries of the former Soviet Union. *Journal of Medical Virology* **2005**; 77(3): 337-44.
33. Lai MT, Lu M, Felock PJ, et al. Distinct mutation pathways of non-subtype B HIV-1 during in vitro resistance selection with nonnucleoside reverse transcriptase inhibitors. *Antimicrobial Agents and Chemotherapy* **2010**; 54(11): 4812-24.
34. Singh K, Flores JA, Kirby KA, et al. Drug resistance in non-B subtype HIV-1: impact of HIV-1 reverse transcriptase inhibitors. *Viruses* **2014**; 6(9): 3535-62.
35. Zaller N, Mazhnaya A, Larney S, et al. Geographic variability in HIV and injection drug use in Ukraine: implications for integration and expansion of drug treatment and HIV care. *The International Journal on Drug Policy* **2015**; 26(1): 37-42.

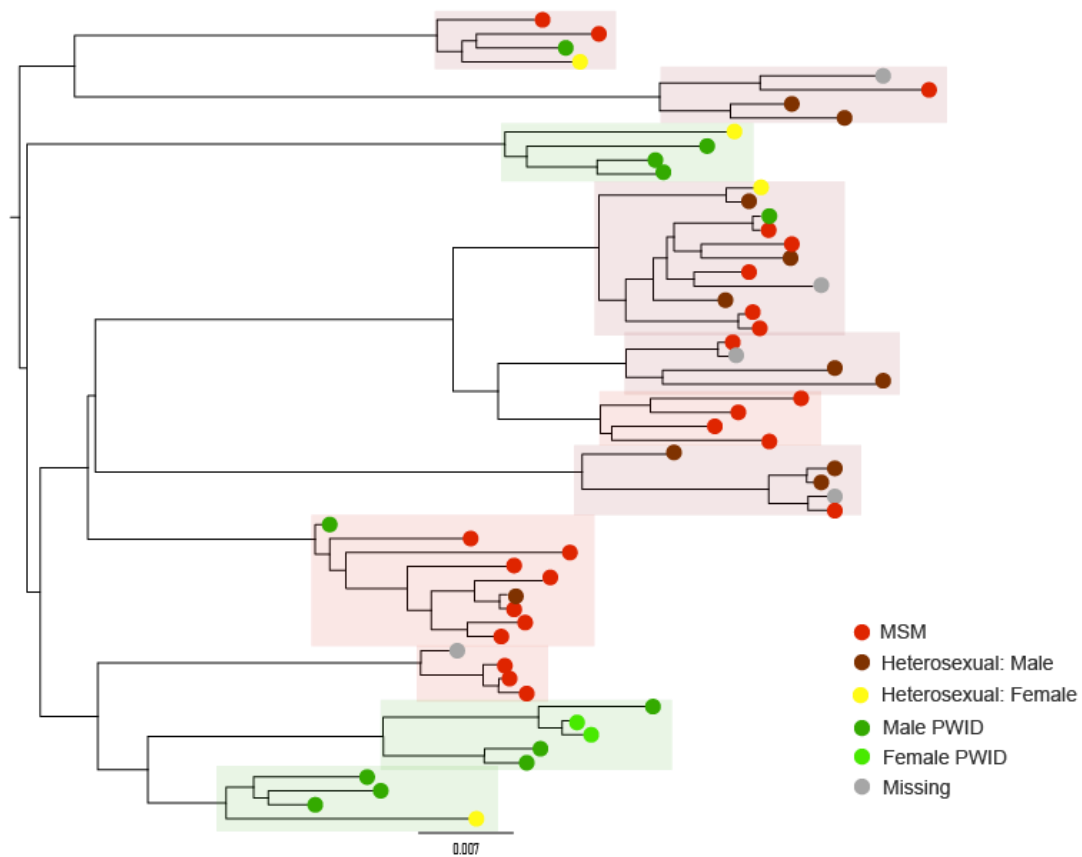
**Table 1.** Factors associated with clustering from HIV-1 pol sequences collected in Kiev April 2013-March 2015

	All	(% )	Clustered	(% )	Univariate		Multivariate	
					OR (95% CI)	P value	aOR (95% CI)	P value
<b>All</b>	858		283	33%				
<b>Subtype</b>	<b>858</b>							
A1	794	93%	241	30%	1		1	
B	64	7%	42	66%	4.38 (2.56-7.50)	<0.001	2.41 (1.28-4.54)	0.007
<b>Sex</b>	<b>804</b>							
Female	234	29%	50	21%	1			
Male	570	71%	215	38%	2.23 (1.56-3.18)	<0.001		
<b>Age (Years)</b>	<b>804</b>							
<= 25	83	10%	34	41%	1.83 (1.10-3.05)	< 0.001	1.39 (0.78-2.49)	
26 - 30	199	25%	88	44%	2.09 (1.42-3.07)		1.75 (1.15-2.65)	0.033
31 - 35	249	31%	68	27%	1.00 (0.67-1.46)		1.02 (0.68-1.53)	
>= 36	273	34%	75	27%	1		1	
<b>Risk Group</b>	<b>794</b>							
Heterosexual Male	230	29%	79	34%	1		1	
Heterosexual Female	181	23%	37	20%	0.49 (0.31-0.77)		0.51 (0.32-0.81)	<0.001
PWID: Male	241	30%	67	28%	0.74 (0.50-1.09)		0.82 (0.55-1.23)	
PWID: Female	51	6%	13	25%	0.65 (0.33-1.30)		0.68 (0.34-1.38)	
MSM	91	11%	68	75%	5.65 (3.27-9.75)	<0.001	4.36 (2.48-7.67)	<0.001
<b>Recent Infection</b>	<b>795</b>							
No	728	92%	227	31%	1			
Yes	67	8%	37	55%	2.72 (1.64-4.52)	<0.001		
<b>Repeat Tester</b>	<b>796</b>							
No	678	85%	226	33%	1			
Yes	118	15%	39	33%	0.99 (0.65-1.50)	0.990		
<b>Sex worker/contact with sex work</b>	<b>794</b>							
No	746	94%	274	37%	1			
Yes	48	6%	24	50%	1.93 (1.07-3.47)	0.028		
<b>Reason for test</b>	<b>763</b>							
High risk behaviours	361	47%	112	31%	1			
Screening	46	6%	11	24%	0.70 (0.34-1.43)			
Clinical indicators	356	47%	130	37%	1.27 (0.94-1.74)	0.120		

CI Confidence interval  
aOR Adjusted odds ratio  
PWID People who inject drugs  
MSM Men who have sex with men



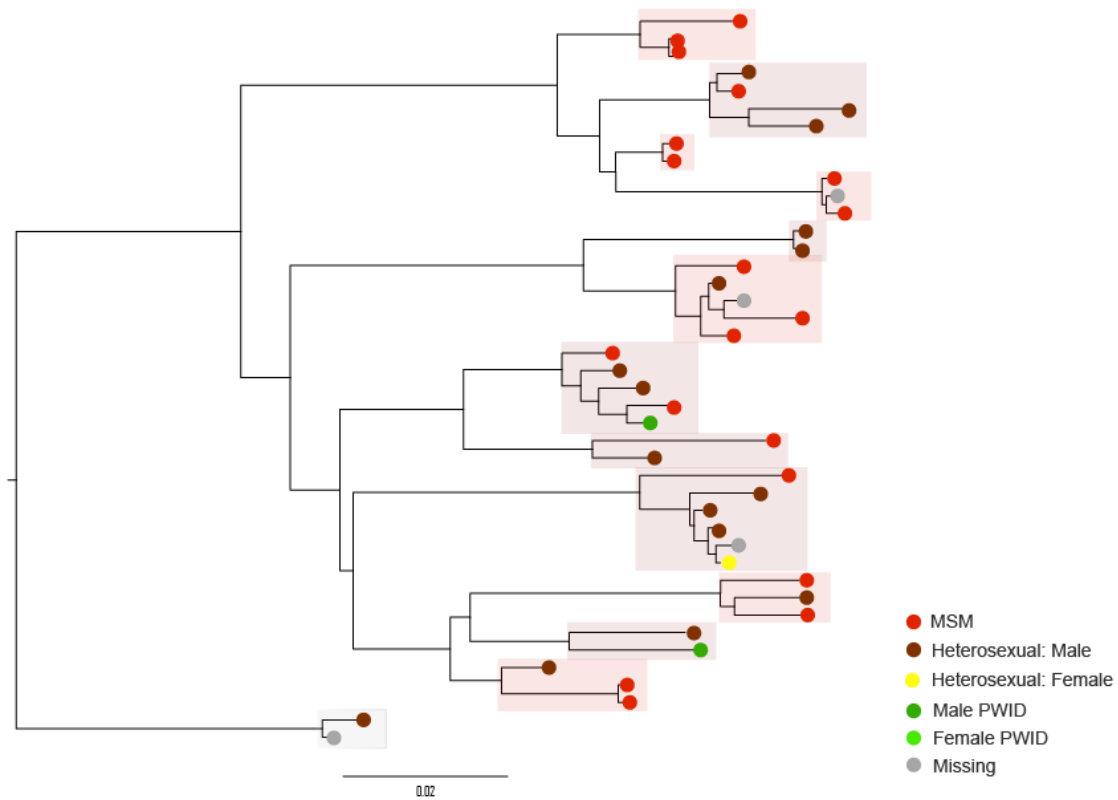
**Figure 1.**



**Figure 1.** Subtype A1 phylogeny (clusters of size 4 and above), HIV-1 pol sequences collected in Kiev April 2013-March 2015.

Subtype A1 clusters sizes 4 and above only. Individual clusters are highlighted and coloured according to the risk group of over 60% of samples in the cluster (see figure 3)

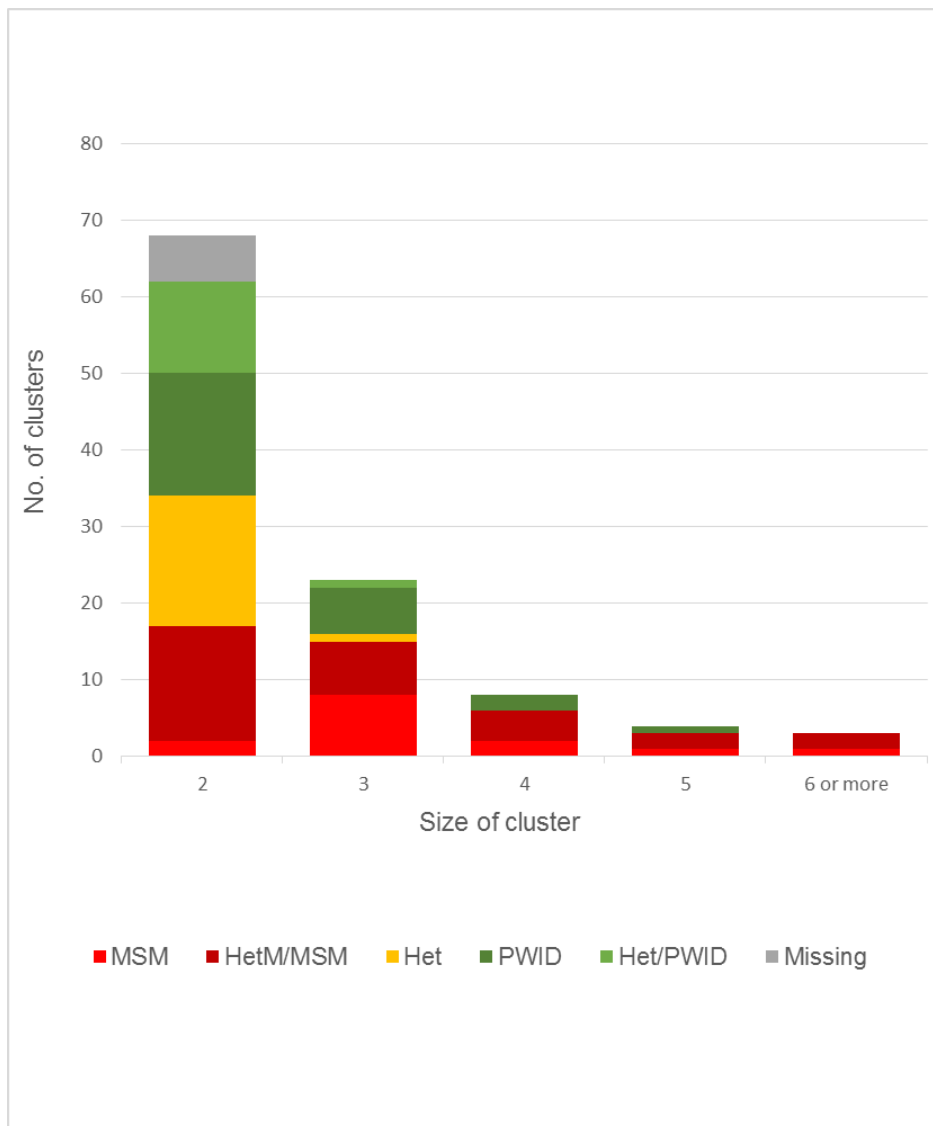
**Figure 2.**



**Figure 2.** Subtype B phylogeny (all clusters), HIV-1 pol sequences collected in Kiev April 2013-March 2015.

Subtype B clusters all sizes. Individual clusters are highlighted and coloured according to the risk group of over 60% of samples in the cluster (see figure 3)

**Figure 3.**

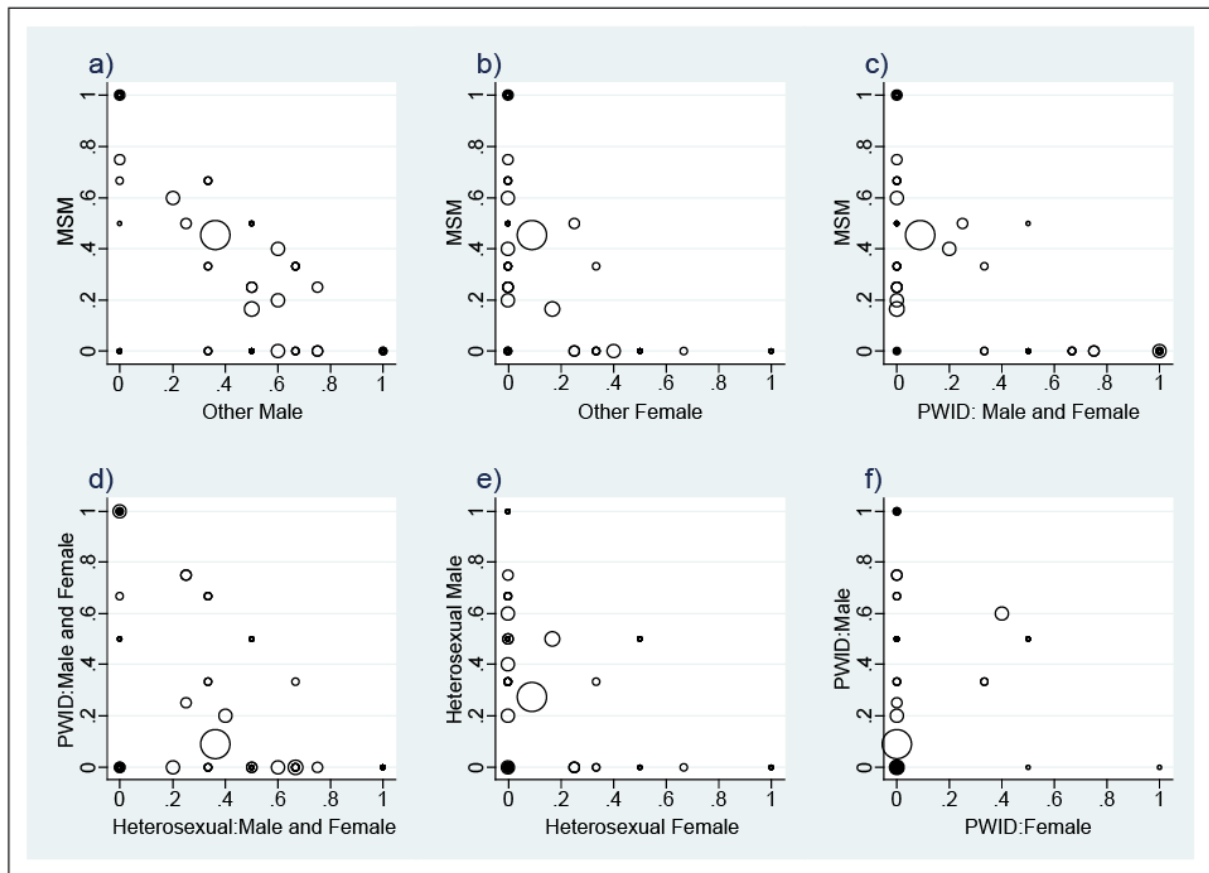


**Figure 3.** Clusters by risk group. Subtype A1 and B HIV-1 pol sequences collected in Kiev April 2013-March 2015.

Groups based on >60% of samples in the cluster.

Abbreviations: MSM, Men who have sex with men; HetM/MSM, Heterosexual males (sexual contact and PWID) and MSM; Het, Heterosexual sexual contact (male and female); PWID, People who inject drugs (male and female); Het/PWID, Heterosexual sexual contact (male and female) and PWID (male and female)

Figure 4.



**Figure 4.** Scatter plot of association of risk groups within clusters. Subtype A1 and B HIV-1 pol sequences collected in Kiev April 2013-March 2015

On each graph, points represent every cluster, with the marker scaled to cluster size. The x and y axis denote risk groups, and the location of the marker shows the proportion of sequences in the cluster belonging to each of the risk groups.

Abbreviations: MSM, Men who have sex with men; PWID, People who inject drugs