Laboratory of Clinical and Epidemiological Virology
Clinical and Molecular Aspects of Infectious Diseases Group

Bogotá, Colombia / Leuven, Belgium
2014

Andrea-Clemencia Pineda-Peña

THE DYNAMICS OF LOCAL HIV-1 EPIDEMICS
The Colombian and Belgian Cohorts

August 2014

KU LEUVEN

UNIVERSIDAD DEL ROSARIO

THE DYNAMICS OF LOCAL HIV-1 EPIDEMICS

The Colombian and Belgian Cohorts

Supervisory Committee:
Prof. Anne-Mieke Vandamme
Prof. Kristel Van-Laethem
Prof. Arley Gomez-Lopez
Prof. Otto Sussmann

Andrea-Clemencia Pineda-Peña

Doctoral Thesis in Biomedical Sciences

KU Leuven
Biomedical Sciences Group
Faculty of Medicine
Department of Microbiology and Immunology
Rega Institute for Medical Research
Laboratory of Clinical and Epidemiological Virology
Leuven, Belgium

Universidad del Rosario
Faculty of Sciences and Mathematics
Faculty of Medicine
Biomedical Doctoral School
Clinical and Molecular Aspects of Infectious Diseases Group
Bogotá, Colombia

THE DYNAMICS OF LOCAL HIV-1 EPIDEMICS
The Colombian and Belgian cohorts

Andrea-Clemencia Pineda-Peña
Double Degree KU Leuven and Universidad del Rosario

Promoter of KU Leuven:	**Prof. Dr. Anne-Mieke Vandamme**
Co-promoter of KU Leuven:	**Prof. Dr. Kristel Van Laethem**
Promoter of Universidad del Rosario:	**Prof. Dr. Arley Gomez Lopez**
Co-promoter of Universidad del Rosario:	**Prof. Dr. Otto A. Sussmann**

Dissertation presented in partial fulfillment of the requirements for the double degree of Doctor in Biomedical Sciences

Leuven and Bogotá, 2014

Andrea-Clemencia Pineda-Peña was supported by a Doctoral Research Training Program ''Francisco Jose de Caldas'' by the *Departamento Administrativo de Ciencia, Tecnología e Innovación, COLCIENCIAS*, Republic of Colombia and ERACOL academic exchange in medicine and the health sciences, between Europe and Latin America. The research was supported in part by the European Community's Seventh Framework Programme (FP7/2007–2013) under the project ''Collaborative HIV and Anti-HIV Drug Resistance Network (CHAIN)'' grant agreement n° 223131 and ERC grant agreement no. 260864; by the Fonds voor Wetenschappelijk Onderzoek – Flanders (FWO) grant G.0611.09 and grant G069214N; by the AIDS Reference Laboratory of Leuven that receives support from the Belgian Ministry of Social Affairs through a fund within the Health Insurance System; by the University of Leuven (Program Financing no. PF/10/018. The computational resources and services used in this work were provided by the Hercules Foundation and the Flemish Government – department EWI-FWO Krediet aan Navorsers (Theys, KAN2012 1.5.249.12). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscripts.

KU Leuven
Biomedical Sciences Group
Faculty of Medicine
Department of Microbiology and Immunology
Rega Institute for Medical Research
Laboratory of Clinical and Epidemiological Virology
Leuven, Belgium

Universidad del Rosario
Faculty of Sciences and Mathematics
Faculty of Medicine
Biomedical Doctoral School
Clinical and Molecular Aspects of Infectious Diseases Group
Bogotá, Colombia

THE DYNAMICS OF LOCAL HIV-1 EPIDEMICS
The Colombian and Belgian cohorts

Andrea-Clemencia Pineda-Peña
Double Degree KU Leuven and Universidad del Rosario

Jury:

Promoter of KU Leuven:	**Prof. Dr. Anne-Mieke Vandamme**
Co-promoter of KU Leuven:	**Prof. Dr. Kristel Van Laethem**
Promoter of Universidad del Rosario:	**Prof. Dr. Arley Gomez Lopez**
Co-promoter of Universidad del Rosario:	**Prof. Dr. Otto A. Sussmann**
Jury members:	**Prof. Dr. Inge Derdelinckx**
	Prof. Dr. Geert Molenberghs
	Prof. Dr. Eduardo Gotuzzo
	Prof. Dr. Chris Verhofstede

Dissertation presented in partial fulfillment of the requirements for the double degree of Doctor in Biomedical Sciences

Leuven and Bogotá, 2014

"*Take the first step in faith. You don't have to see the whole staircase, just take the first step*"

Martin Luther King, Jr.

This is the first step of…

ACKNOWLEDGMENTS / AGRADECIMIENTOS

Some traditions consider gratefulness one of the keys of the life. Since I do believe it, I would like to thank you everybody here and there. Thank you for sharing this path and teaching me how life is worthy every second!

Algunas tradiciones consideran la gratitud como una de las claves de la vida. Como lo creo, quisiera agradecer a todos aquí y allá. Gracias por compartir ese camino y enseñarme como la vida vale cada segundo!

Primero quiero agradecer a **Dios** quien ha estado conmigo siempre y me ha dado esta oportunidad, y a toda mi familia en el cielo que me ha acompañado y animado en este camino.

I would like to thank to my **promoters** Prof. Anne-Mieke Vandamme, Prof. Kristel Van Laethem, Prof. Arley Gomez and Prof. Otto Sussmann for their support in this doctorate. To Annemie and Kristel, thank you for your guidance and fruitful scientific discussions. Your patience and knowledge were valuable for this thesis. To Annemie, I hope you continue with your leadership, open heart, adventurous attitude and free mind. To Kristel, keep your warm-hearted personality, reliability, and structured mind (plus all I've told you before). Arley muchas gracias por brindarme tu incondicional apoyo como profesor y amigo, tus palabras siempre fueron inspiradoras y oportunas. Continúa con fe en todos tus propósitos. Otto gracias por introducirme en este medio del VIH, transmitirme tu pasión por las Infecciosas, tu confianza e impulsarme a continuar con mis metas. Continúa con tu buen humor y pasión por tu profesión.

I am grateful with my **Jury Members** Prof. Inge Derdelinckx, Prof. Eduardo Gotuzzo, Prof. Geert Molenberghs, Prof. Chris Verhofstede and Prof. Ghislain Opdenakker for their suggestions for the improvement of this thesis and their encouragement. I am indebted with the coordinator of my Doctoral training in Colombia, Prof. Luisa Marina Matheus and my administrative supervisor, Prof. Manuel Alfonso Patarroyo.

To my colleagues and friends from the Clinical and Epidemiological Virology group: I am indebted to my colleagues for their support and the nice moments we have shared, I hope all of you reach your goals and specially happiness along the way. In alphabetical order… A Carolina Alvarez a quien la bondad y dedicación le sobran, espero sigas triunfando en todos los aspectos. Por favor recuerda nuestras charlas "potencial Oncóloga". To Lize, I hope you continue with you harmonious leadership and with your kindness, it was nice to share the workshop with you. To Ricardo, thank you for your positive attitude, serenity and scientific inspiration. To Fossie, thank you for your professional collaboration but especially for your words, humor and friendship; they were perfect for our daily lunch or our talks with Emelia. To Ricky, you are the "Capoeira" panda of the group, your kindness and pleasant attitude are incredible. To Lowie, I'm grateful for your teachings and help in all the computational stuff, your patience and your sweet heart amazed me. To Pieter, thank you for helping me with the cluster complaints, you are very dedicated and always helpful. To Sarah, thank you for your sweetness and kind words, you always had a gentle smile in every moment. To Soraya, I admire your generosity; I hope you receive several times more love that you offer to others. To Joao, I was impressed for your story of life, I hope you continue with your passion for the science and history. To Soo, I'm glad you pursue your dreams, please keep your cheerful smile and good mood. To Yoeri, thank you for your hilarious quotes and your warm personality. You and Bernard think every little detail that makes the life more special. To Kristof, I admire your change through the time; you are now plenty of confidence, and patience. To Johan, thank you for the scientific discussions y tu ejemplo de fortaleza y dedicación. To Lore, it took some time to know you but I admire your dedication and carefulness, keep your smile always with you (and your love with Jelle). To Annelies, I appreciate your encouraging

talks, but your faith and spontaneous attitudes were very important! By the way thank you for the Dutch summary! To Ria, thank you for your administrative support but more importantly, your kind personality. You and Patrick are very inspiring about love. To Sietse, thank you for your administrative support and help in the last steps.

For my **former friends and colleagues at the Rega**: To Jurgen, I was impressed for your warm personality, smile and humor in every moment, thank you for your statistical support and your affable mood. To Britta & Gertjan, I wish you all the love in your new adventure; please keep that cheerful (Britta) and amiable mood (Gertjan). To Stijn, thank you for your patience and help with the computers stuff, but more for your trust and your nice words. To Raphael, I hope you keep your lovely smile and positive attitude in this new path. To Fausta, please keep your perseverance and big smile. To Monica, I shared nice moments with you, keep seeing the positive side of the life. To Lia, I hope you keep your determination alive and accomplish your goals. To Laura, I really enjoyed our Italian-Spanish-English talks. I am grateful for little but nice moments with Abbas, Caro Palma, Fabienne, Goedele, Heidi, Joke, Kiyoshi, Lien, Supinya, Tim and Yang. To the future member Tim, good luck with the next steps.

I would like to show my gratitude to the people from the **Evolutionary and computational Lab**: Bram, Guy, Nidia, Nuno, Filip, Philippe and Tiago. Especially to Nuno, you have been my buddy all these years, thank you for your encouragement, teachings, and your kindness. To Nidia, thank you for the friendship and cheerful mood! Philippe, I am grateful for your suggestions and useful discussions. I also would like to show my gratitude for the **people of the 6th and 7th floor!.** Especialmente a Alexandra, quien me acompañó en todo momento en esta travesía y terminó siendo parte de nuestra familia :-) Muchos éxitos siempre, que sigas creciendo en todos los aspectos! And Bobby, your positive mood and advices were very helpful! Marc and Elizabeth, keep your nice humor and kindness!
Thank you to **Prof. Eric Van Wijngaerden** for his guidance in the clinical consultation and his willingness to speak several languages at the same time in order to help me understand. I am also grateful with Dr. Paul de Munter and Dr. Inge Derdelicks from the Internal Medicine group.

Quisiera agradecer a las personas de la **Universidad del Rosario**, especialmente a las de la Unidad de Microbiología. Maria Antonia, Dora y Maria Cristina, gracias por ser mi apoyo por todos esos años y transmitirme su amor por las enfermedades infecciosas. Giovanni, tú has sido mi colega y amigo por todo este tiempo, gracias por los momentos vividos en Colombia y en Lovaina… espero sigas cumpliendo tus metas junto a Mafe y tu bella familia. Sigue con tu humor y buen trabajo! A todas mis profes de Genética, Claudia Silva, amiga, profesora y colega…gracias por tus dosis de amor todos estos años. Prof. Restrepo, Dorita, Heidi y Norita, gracias por ser mi familia genetista e introducirme en esta área de la investigación. Gracias Luisa por tu apoyo como coordinadora del Doctorado

Agradezco profundamente a mis amigos y colegas de la **Fundación Instituto de Inmunología de Colombia**. Agradezco al FIDIC por su soporte académico y su apoyo especialmente en el primer semestre del doctorado. Especialmente a Manuel Alfonso, gracias por creer en mí, por darme ánimo y consejo en todo momento, ser ese apoyo y amigo incondicional. Espero logres ese gran sueño que te has propuesto!. A las chicas HPV: Marina, Milena, y Sara. Espero continúen con su excelente trabajo y dedicación. Me alegra que sean ejemplo de amistad!

Agradezco a mis colegas y amigos de **Asistencia Científica de Alta Complejidad**, quienes fueron la fuente de inspiración para continuar con este camino: Andrea, Carlos, German, Jorge, Ivonne, Marce, Milena, Nestor, Paola, Rosmery, Yesika, Yolima, Alexandra, Fany, Jenny, Nini, y Martha. Espero todos ustedes continúen con esa pasión y dedicación por su trabajo. Espero tener el gusto de verlos pronto y compartir más experiencias juntos. Especialmente Otto, gracias por tu confianza y por promocionar la

investigación en la institución y Diana, gracias por tu incondicional ayuda y entusiasmo en ese camino investigativo y de Comités.

Thank you to my **friends in Leuven**, you have been a great support in these 4 years. To my wonderful and always helpful neighbors: Ona & Uwe (a quienes admiro su dedicación y amabilidad), Isael & Maria (por su generosidad y darnos la bienvenida), Vladimir & Keith (you are very warm-hearted), Herman & Monique (your sympathy and cordiality are valuable). To an excellent teacher and lovely friend: Det. Gracias a la comunidad hispano-parlante de Lovaina por su incondicional apoyo y amistad: A Juan Carlos, gracias por tu apoyo, compañía y buenos momentos; continua con tu misión de servicio y humor. A Paul, gracias por esos momentos valiosos y ejemplo de servicio. Adriana Moreno, continúa con ese espíritu libre y carisma. Cesar, tu alegría y animo son envidiables, sigue adelante con tus metas e irradiando felicidad. Melvin, me alegra verte feliz y con mucho ánimo hacia el futuro. Italo & Oksana, gracias por los momentos compartidos y su espíritu amable. A Wouter & Elizabeth, gracias por los momentos compartidos y el apoyo-amistad que le han brindado a Jorgito. Carme & Xavi, su espontanea sonrisa y calidez han sido invaluables. Carolina & Oscar, ustedes nos recibieron con los brazos abiertos, su corazón es de puertas abiertas igual que su generosidad, realmente muchos éxitos en este camino. Natalia & Eddy, espero sigan ese camino con mucha felicidad y Nata admiro tu perseverancia y chispa. Camilo, te deseo mucha felicidad en esta nueva etapa con Ale, sigue juicioso y de buen humor. Jorge Duitama, gracias por tu amistad y generosidad. Andrea Ávila, a sumercé que le vaya bien con Pole y tus metas académicas, continúa con esa chispa que alegra. Carolina Vallecilla, continúa con ese empuje que te caracteriza y amor por Martino. A Milda & Marc, gracias por su inmensa generosidad y adicionalmente, Milda eres una mujer llena de amor. A Varima & Chris, ustedes tienen una alegria contagiosa y emocionante, continuen con ese amor que los caracteriza. A David & Tara, su apoyo principalmente a Jorge y de paso a mí ha sido importante en este camino. Aunque hemos compartido pocos pero intensos momentos, he estado asombrada de la generosidad y calidez de Constanza & familia, Sergio & Maritza, Vilma &Jimmy, Jenny & familia, Gladys & familia, Myriam & familia, y Jesus & familia.

Tengo una gran **familia** de amigos de toda la vida: Nicolas & Silvi, Zhyla & Andres, Catalina & Nicolas, Paola & Mauricio. A ustedes gracias por su compañía desde la distancia, sé que cuento en todo momento con ustedes y su incondicional apoyo. A mi familia, Rosarito, Milingos, Linis y Nacho, gracias por cada momento de apoyo, compañía, y aunque lejos… siempre procurando mi bien. Jorge, Glorita e Ilbita, gracias por creer en mí en cada momento y apoyarme en todas mis empresas. Sus palabras positivas son una motivación importante. Mamalita, gracias por tu eterna sonrisa y tu cariño inmenso. Tu amor ha cambiado mucho la vida de nuestra familia…gracias nuevamente. Ana, Jorge, Carolina, Néstor, Diana, Isabella y Gabriel, gracias por sus oraciones, palabras de apoyo, amor, ternura, ánimo y humor. Espero continúen colmados de amor, éxitos, salud y amor.

Mami, gracias por ser una soñadora y lo más importante darte el lujo de disfrutar cada momento de tu vida. Sigue adelante cumpliendo con tu misión, y tu amor hacia la gente que te rodea. Gracias Alfredito por ser esa compañía de paz para mi ma.

Jorgito, gracias por ser mi Luzh! Tu incondicional apoyo y amor han sido mi perfecta compañía durante este viaje, gracias por elegir acompañarme e inspirarme cada día a ser mejor. Admiro tu capacidad de servicio, adaptabilidad, flexibilidad e increíble buen humor. Muchos nuevos retos vendrán, los cuales espero compartir contigo. Además quisiera continuar creciendo junto a ti y escribiendo nuestra historia juntos…

Gracias a cada uno

TABLE OF CONTENTS

FIGURES

TABLES

ABBREVIATIONS

ABC:	Abacavir
ACIN:	Spanish abbreviation for Asociacin Colombiana de Infectología
AIDS:	Acquired Immunodeficiency syndrome
aLTR:	Approximate likelihood ratio test
ANRS:	French abbreviation for Agence National de Recerche sur le SIDA
APV:	Amprenavir
ART:	Antiretroviral therapy
ARC:	AIDS Reference Center
ATP:	Adenosine triphosphate
ATV:	Atazanavir
AZT:	Zidovudine
BCI:	Bayesian Credible interval
BF:	Bayes Factor
BLAST:	Basic Local Alignment Search Tool
BREACH:	Belgium Research on AIDS and HIV Consortium
BSSVS:	Bayesian Stochastic Search Variable Selection
CXR4:	CXC-chemokine receptor 4
CCR5:	CC-chemokine receptor 5
CCR5 Δ 32:	32 base pairs deletion in the allele of CCR5
CD4:	Cluster of differentiation 4
CD8:	Cluster of differentiation 8
CDC:	Center for Disease Control and Prevention
CI:	Confidence interval
COMET:	COntext-based Modeling for Expeditious Typing
CPR:	Calibrated Population Resistance Tool
CRA:	Coreceptor antagonists
CRF:	Circulating Recombinant Form
d4T:	Estavudine
ddC:	Zalcitabine
ddI:	Didanosine
DHHS:	United States Department of Health and Human Services
DNA:	Deoxyribonucleic acid
DLV:	Delavirdine
DRV:	Darunavir
DTV:	Dolutegravir
EACS:	European AIDS Clinical Society
EFV:	Efavirenz
env:	Envelope
ELISA:	Enzyme linked immunosorbent assay
ETV:	Etravirine
EWIs:	Early warning indicators
FI:	Fusion inhibitors
FTC:	Emtricitabine
FSV:	Fosamprenavir
***gag*:**	group specific antigen gene
gp:	glycoprotein
GSS:	Genotypic susceptibility score
GTR:	Generalised time reversible model
GTR+G+I:	GTR plus gamma (Γ) and proportion invariant model
HAART:	Highly active antiretroviral therapy
HPD:	Highest posterior density
HIV:	Human immunodeficiency virus
HIVdb:	Stanford HIV Drug Resistance Database
HLA:	Human leukocyte antigen
HPTN:	HIV Prevention trials network
IDV:	Indinavir
Ig:	Immunoglobulin
IN:	Integrase or p31
INI:	Integrase inhibitor
ISP-WIV:	Belgian Scientific Institute for Public Health
IQR:	Interquartile range
IVDU:	Intravenous drug users

jpHMM:	Jumping profile Hidden Markov Model
LAC:	Latin America and Caribbean
LANL:	Los Alamos database
LTR:	Long terminal repeat sequences
LPV:	Lopinavir
MA:	Matrix or p17
ML:	Maximum likelihood method
MPhy:	Manual Phylogenetic analysis
mRNA:	Messenger RNA
MCC:	Maximum clade credibility
MJ:	Markov jumps
MSM:	Men who have sex with men
MTCT:	Mother to child transmission
MVC:	Maraviroc
NCBI:	National Center for Biotechnology Information
NFV:	Nelfinavir
NGS:	Next generation sequencing
NJ:	Neighbor joining method
NNRTI:	Non-nucleoside reverse transcriptase inhibitors
NRTI:	Nucleoside/nucleotide reverse transcriptase inhibitors
nts:	Nucleotides
NU:	Nucleocapsid
NVP:	Nevirapine
OAM:	Orinoquía y Amazonas regions
OPS:	Spanish abbreviation for PAHO
OR:	Odds ratio
PAHO:	Pan American Health Organization
PBMC:	Peripheral Blood Mononuclear Cells
PCR:	Polymerase chain reaction
PEP:	Post-exposure prophylaxis
PLHIV:	People living with HIV
PI:	Protease inhibitors
PMP:	Polymorphism
***pol*:**	polymerase gene
PR:	Protease or p11
PrEP:	Pre-exposure prophylaxis
RAL:	Raltegravir
REGA:	Rega Institute for Medical Research
REGAv2:	REGA automated subtyping tool version 2
REGAv3:	REGA automated subtyping tool version 3
RNA:	Ribonucleic acid
RPV:	Rilpivirine
RT:	Reverse Transcriptase or p66/p51
RTV or /r:	Ritonavir
SCUEAL:	Subtype Classification Using Evolutionary ALgorithms
SDRM:	Surveillance of drug resistance mutations
SIV:	Simian Immunodeficiency Virus
SQV:	Saquinavir
STAR:	HIV-1 Subtype Analyzer
STD:	Sexually transmitted diseases
SU:	Surface protein or glycoprotein 120
TAMs:	Thymidine analogue mutations
TasP:	Treatment as prevention
tat:	Transactivator HIV gene expression
TC:	Transmission cluster
TDF:	Tenofovir
TDR:	Transmitted drug resistance
TDR-OT:	Onward transmission of resistance
TMRCA:	The most recent common ancestor
TPV:	Tipranavir
T20:	Enfuvirtide
U:	Unclassified subtype or recombinant
UK:	United Kingdom
UNAIDS:	Joint United Nations Programme on HIV/AIDS
URFs:	Unique recombinant forms
USA:	United States of America
WHO:	World Health Organization
ZDV:	Zidovudine
3TC:	Lamivudine

SUMMARY

Approximately 35 million of people were living with Human Immunodeficiency Virus (HIV) and 2 million were newly infected in 2012. Prevention measures and a higher coverage of antiretroviral therapy (ART) resulted in a decline of HIV incidence by more than 50% in 26 countries between 2001 and 2012. In addition, the observed impact of ART on community-level HIV incidence mainly on clinical trials has influenced clinical guidelines to start treatment early and to set up test-and-treat strategies or treatment as prevention (TasP) in some countries. However, it is a matter of debate if these policies would be cost-effectively in different settings. It is imperative that local governments become better informed with timely data about the characteristics of the local HIV epidemics, because they could influence the development of effective preventive policies. In this project we wanted to gain a deeper understanding of the dynamics of the HIV-1 epidemic in Belgium and Colombia, using viral sequences, socio-demographic and clinical data, complemented with statistical and phylogenetic approaches. These two countries share an increasing incidence and prevalence of HIV-1 infection in recent years, but differ in the HIV health care and availability of relevant data. Consequently, different research questions were triggered for each country.

Unlike Belgium, there was limited information available about the molecular epidemiology of HIV-1 in Colombia and the spread of subtypes. In *chapter 3*, we focused on viral diversity and how the virus spread geographically and temporally in Colombia. Subtype B still predominated between 2002 and 2007 in the country. Additionally, phylogenetic analyses showed that multiple introductions occurred in the early 80s and suggested a link between the Colombian and Spanish epidemics. Bogotá, the capital, was the main exporter of the HIV epidemic, and other important cities within Colombia acted as sinks. These findings suggest an important role of tourism in the spread of HIV, but further studies should investigate the complete dynamics of this epidemic.

Since Transmitted drug resistance (TDR) can impact the efficacy of first-line ART and its use in prevention strategies such as the prevention of mother-to-child transmission or TasP, we focused our efforts to retrieve any data on TDR in the Colombian epidemic and zoomed out our evaluation at regional level in *chapter 4*. A systematic review was performed using published literature and viral sequences available from Colombia and other Latin American and Caribbean countries. The overall TDR prevalence was stable and some countries had more than 5%. The reported prevalence of TDR was around 6% in Colombia, but there were no data about trends or related factors. These findings

imply that surveillance of TDR should be frequently performed in countries such as Brazil or Venezuela or warranted in settings with scarce data such as Colombia, Central America and the Andean region.

We investigated the evolution in TDR prevalence for the different drug classes and its association with other socio- demographic factors at the AIDS Reference Center and AIDS Reference Laboratory in UZ Leuven as a first step in the study of the epidemic in Belgium. The TDR and viral diversity prevalence were stable in Leuven between 1998 and 2012, while a parabolic trend was observed for TDR against non-nucleoside reverse transcriptase inhibitors. We were not able to identify a specific population significantly associated with TDR or with spread of TDR, but men who have sex with men and originating from Belgium seemed to play an important role. The high number of chronically infected individuals in the cohort and the number of patients included in transmission clusters suggest that national collaborative studies should be prioritized to formulate policies that target earlier HIV diagnosis and prevention of transmission.

As a methodological step, we evaluated the performance of automated subtyping tools with the aim to classify quickly the large amount of sequences used in different analyses in the Belgian and Colombian cohorts. Since our laboratory developed REGA version 3 (REGAv3), we compared our tool with the previous version REGA version 2 and six other tools. We found that most of the tools accurately identified subtype C and B, the most frequent clades worldwide. However, the performance was variable for other subtypes. COMET, jpHMM, REGAv3, and SCUEAL had good performance for pure subtypes using the *pol* region whereas COMET and REGAv3 also had good performance when analyzing most of the CRFs. Consequently, we recommended the use of two subtyping tools for surveillance and clinical purposes, and this was the strategy we also used to classify subtypes in other chapters.

Finally, this thesis constitutes the initial step to characterize both epidemics. The complete understanding of these two local HIV epidemics will be useful for designing prevention strategies as well as for improving HIV health care programs.

RESUMEN

Aproximadamente 35 millones de personas vivían con el Virus de la Inmunodeficiencia Humana (VIH) y 2 millones de casos de nuevas infecciones ocurrieron en el 2012. Medidas de prevención y una mayor cobertura de la terapia antirretroviral (TAR) permitió una disminución de la incidencia del VIH en más de un 50% en 26 países entre 2001 y 2012. Además, el impacto observado de la TAR en la incidencia del VIH a nivel de salud pública en los ensayos clínicos ha influido en las guías clínicas para iniciar un tratamiento antirretroviral más temprano y la creación de estrategias como "*test and treat*" (testear-diagnosticar y tratar) o tratamiento como prevención (TasP; del Ingles *Treatment as Prevention*) en algunos países. Sin embargo, es debatible si estas estrategias serían costo-efectivas en diferentes entornos. Por lo tanto, es imperativo que los gobiernos a nivel local estén mejor informados sobre las características específicas de la epidemia por VIH con datos actualizados, ya que pueden influir en el desarrollo de políticas de prevención eficaces. En este proyecto nuestro objetivo es describir la dinámica de la epidemia del VIH-1 en Bélgica y Colombia, utilizando secuencias virales, datos socio-demográficos y clínicos, complementados con métodos estadísticos y filogenéticos. Estos dos países comparten un incremento de la incidencia y prevalencia de la infección por VIH-1 en los últimos años, pero difieren en la prestación de servicios de salud y la disponibilidad de información pertinente. Dado estas diferencias, las preguntas de investigación fueron diferentes para cada país

A diferencia de Bélgica, había poca información disponible acerca de la epidemiología molecular del VIH-1 en Colombia y la consecuente propagación de los subtipos. En el *capítulo 3*, nos hemos centrado en la diversidad viral y cómo el virus se propagó temporalmente y geográficamente en Colombia. El subtipo B predominó en el país entre los años 2002 y 2007. Además, los análisis filogenéticos mostraron que múltiples introducciones se produjeron entre finales de los años 70s y principios de los 80s e indicaron que las epidemias de Colombia y España estaban relacionadas. A nivel nacional, Bogotá, la capital, fue el principal exportador de la epidemia del VIH, y otras ciudades importantes actuaron como importadores del virus. Estos hallazgos sugieren un papel importante del turismo en la propagación del VIH, pero nuevos estudios son necesarios para investigar la dinámica de esta epidemia.

La farmacorresistencia transmitida (TDR; del Ingles *Transmitted drug resistance*) puede afectar la eficacia de la TAR de primera línea y su uso en las estrategias de prevención, tales como la prevención de la transmisión de madre a hijo o TasP. Por lo consiguiente, nuestro objetivo era obtener todos los datos de TDR disponibles en Colombia y extenderlos a nivel regional en el *capítulo 4*. Una revisión sistemática fue realizada con la literatura publicada y secuencias virales disponibles en bases

de datos públicas que incluyeran información acerca de Colombia y otros países de América Latina y el Caribe. La prevalencia de TDR se mantuvo estable y en algunos países la prevalencia fue mayor al 5%. Para Colombia, la prevalencia de TDR fue alrededor del 6% en Colombia, pero no hubo datos sobre tendencias o factores socioeconómicos o clínicos relacionados. Estos resultados implican que la vigilancia de TDR se debe realizar con frecuencia en países como Brasil o Venezuela o es requerida urgentemente en países con escasez de datos, tales como Colombia, América Central y la región Andina.

Como un primer paso en el estudio de la epidemia en Bélgica, investigamos la evolución de la prevalencia de TDR para las diferentes clases de drogas y su asociación con otros factores socio-demográficos en el Centro de Referencia del SIDA y del Laboratorio de Referencia de SIDA en UZ Leuven. La prevalencia de la diversidad viral y TDR fueron estables en Lovaina entre los años 1998 y 2012, mientras que se observó una tendencia parabólica de resistencia contra los inhibidores no nucleósidos de la transcriptasa reversa. Aunque no se identificó significamente una población específica asociada con TDR o con la propagación de TDR, se evidenció una posible asociación con los hombres que tienen relaciones sexuales con hombres, nacidos en Bélgica. El elevado número de personas con infección crónica en la cohorte y que participan en grupos de transmisión, sugiere que se debe priorizar a nivel nacional estudios colaborativos los cuales ayuden a la formulación de políticas de salud pública para fortalecer el diagnóstico y prevención de la transmisión del VIH.

Como paso metodológico, se compararon las herramientas automatizadas que clasifican los subtipos virales, con el propósito de identificar rápidamente la gran cantidad de secuencias utilizadas en los diferentes análisis realizados en las cohortes de Bélgica y Colombia. Como nuestro laboratorio diseñó REGA versión 3 (REGAv3), nuestro objetivo era comparar nuestra herramienta con la anterior versión y otras seis herramientas usadas frecuentemente o reportadas para clasificar subtipos. Como resultado se evidenció que la mayoría de herramientas identificaron los subtipos B y C, los cuales son los más frecuentes a nivel mundial. Sin embargo, la clasificación de otros subtipos fue variable. Cuando la región *pol* fue analizada, COMET, jpHMM, REGAv3 y SCUEAL identificaron los subtipos virales mientras que COMET y REGAv3 clasificaron las formas recombinantes circulantes. Por lo tanto, nosotros recomendamos el uso de por lo menos dos herramientas para identificar subtipos virales en la práctica clínica o en estudios de vigilancia epidemiológica molecular. Esta estrategia fue también usada para clasificar los subtipos en otros capítulos de esta tesis.

Por último, esta tesis constituye el primer paso para caracterizar las dos epidemias. La descripción detallada de estas dos epidemias de VIH a nivel local será de utilidad para el diseño de estrategias de prevención, así como para mejorar los programas de atención de salud del VIH.

SAMENVATTING

In 2012 waren er wereldwijd ongeveer 35 miljoen mensen met een HIV besmetting en liepen er 2 miljoen mensen die besmetting op. Preventieve maatregelen en een groter bereik van antiretrovirale behandeling (ART) resulteerde tussen 2001 en 2012 in een daling van de HIV incidentie met meer dan

50% in 26 landen. Het geobserveerde effect van ART op de incidentie van HIV in de bevolking zorgde voor een bijsturing van klinische richtlijnen in de richting van vroegtijdige behandeling en het opzetten van een test-en-behandel strategie of behandelen als preventie ("*treatment as prevention*", (TasP)) in sommige landen. Het is echter nog niet duidelijk of deze aanpak in verschillende omstandigheden kosteneffectief zal zijn. Daarom moeten lokale overheden beter geinformeerd worden met actuele gegevens over de kernmerken van de lokale HIV epidemies, omdat zij de ontwikkeling van doeltreffende preventieve maatregelen kunnen beïnvloeden. Het doel van deze studie was een beter begrip van de dynamiek van de HIV-1 epidemie in België en Colombia, met behulp van virale sequentie, socio-demografische en klinische gegevens, aangevuld met statistische en fylogenetische benaderingen. In beide landen is in de voorbije jaren zowel de incidentie als de prevalentie van HIV-1 infecties gestegen, maar zijn er verschillen in desbetreffende gezondheidszorg en de beschikbaarheid van relevante gegevens. Deze verschillen resulteerden in verschillende onderzoeksvragen voor beide landen.

Anders dan in België, was er in Colombia slechts beperkt informatie beschikbaar over de moleculaire epidemiologie van HIV-1 en de verspreiding van subtypes. In *hoofdstuk 3* ligt de focus op virale diversiteit in Colombia en de verspreiding van het virus in tijd en ruimte aldaar. Tussen 2002 rn 2007 had subtype B er nog de overhand. Fylogenetische analyse toonde aan dat meerdere introducties plaatsvonden in het begin van de jaren 1980 en suggereerde een verband tussen de epidemies in Colombia en in Spanje. De hoofdstad Bogota was the voornaamste exporteur van de HIV epidemie, terwijl andere belangrijke steden in Colombia als importeur functioneerden. Deze resulaten suggereren dat toerisme een belangrijke rol speelde in de verspreiding van HIV, maar meer onderzoek is nodig om de volledige dynamiek van deze epidemie te ontrafelen.

Omdat TDR de effectiviteit van eerstelijns ART kan beïnvloeden en aldus ook zijn toepassing in strategieën zoals preventie van verticale transmissie of TasP, hebben we er ons op toegelegd om zo veel mogelijk gegevens over de aanwezigheid van TDR in de Colombiaanse epidemie te verzamelen en daarna het probleem op grotere, regionale schaal te bekijken in *hoofdstuk 4*. We voerden een systematische review door van de gepubliceerde literatuur en de beschikbare virale sequenties van

Colombia en andere Latijns-Amerikaanse en Caraïbische landen. De prevalentie van totale TDR was stabiel, maar overtrof in sommige landen de 5%. De gerapporteerde prevalentie van TDR was in Colombia rond de 6%, maar er waren geen gegevens over trends of geassocieerde factoren. Deze vaststellingen betekenen dat monitoring van TDR frequent moet uitgevoerd worden in landen zoals Brazilië of Venezuela, of aanbevolen moet worden op plaatsen met slechts beperkte gegevens, zoals in Colombia, in Centraal Amerika en in het Andesgebied.

We onderzochten de evolutie van de prevalentie van overdracht van resistent HIV-1 ('Transmission of drug resistance" or TDR) voor de verschillende klassen van remmers en zijn associatie met andere socio-demografische factoren in het AIDS Referentie Centrum and AIDS Referentie Laboratorium van UZ Leuven, als een eerste stap in de studie van de epidemie in België. TDR en virale diversiteit waren tussen 1998 en 2012 stabiel in België, maar voor TDR tegen niet-nucleoside reverse transcriptase inhibitoren observeerden we een parabolische trend. We konden geen specifieke populatie identificeren die significant geassocieerd was met TDR of de verspreiding van TDR, maar de combinatie van MSM en Belgische origine leek een belangrijke rol te spelen. Het grote aantal chronisch geïnfecteerden in de cohorte en het aantal patiënten in transmissie clusters suggereren dat de voorrang moet verleend worden aan collaboratief onderzoek op nationale schaal, om zo maatregelen te formuleren met als doelwit vroegere diagnose en preventie van transmissie van HIV.

In een methodologische studie evalueerden we de prestatie van geautomatiseerde subtyping tools, met als doel een snelle klassificatie van de grote hoeveelheden sequenties die gebruikt zijn in de verschillende analyses in de cohortes uit België en Colombia. Omdat ons laboratorium REGA versie 3 (REGAv3) ontwikkeld heeft, vergeleken we onze tool met de vorige versie, REGA versie 2 en zes andere tools. We zagen dat de meeste tools subtypes C en B, de claden die wereldwijd het meest voorkomen, accuraat konden identificeren. Voor de andere subtypes verschilde de prestatie van tool tot tool. COMET, jpHMM, REGAv3 en SCUEAL presteerden goed voor zuivere subtypes wanneer de *pol* regio werd geanalyseerd, terwijl COMET en REGAv3 ook goed resultaat opleverden bij de analyse van de meeste CRFs. We raadden daarom aan om telkens twee subtyping tools te gebruiken in het kader van surveillance studies en klinische monitoring, en dat is ook de strategie die we zelf toepasten in de andere hoofdstukken.

Deze thesis is een eerste stap in de richting van de karakteristaie van beide epidemies. Een volledig begrip van deze twee lokale HIV epidemies zal bijdragen tot het ontwerpen van preventieve strategieën en tot het verbeteren van programma's in de HIV gezondheidszorg.

CHAPTER 1 INTRODUCTION

1.1 HISTORY

The first cases of human immunodeficiency virus (HIV) infections were described in 1981 when Kaposi's sarcoma and pneumonia by *Pneumocystis carinii* (now *Pneumocystis jiroveci)* were reported in the United States of America (USA) [1, 2]. With the increasing number of opportunistic infections and cancers in previous healthy people and the impairment of cell-mediated immunity, the acronym Acquired Immunodeficiency Syndrome (AIDS) was introduced one year later. Subsequently, a retrovirus was identified in 1983 by the French group headed by Luc Montagnier [3, 4]. In 1987, the hope of a treatment became reality with the use of zidovudine (ZDV), which decreased mortality and opportunistic infections [5]. Subsequently, mono- and bi-therapies were prescribed with the introduction of didanosine (ddI) and zalcitabine (ddC). The first evidence of resistance was reported in 1989 with the identification of mutations within reverse transcriptase [6, 7], together with the transient effect of mono-therapy in preventing disease progression [8]. It was not until 1996 that Highly Active Antiretroviral Therapy (HAART) was used that eventually resulted into a decrease of mortality and morbidity [9].

1.2 WORLDWIDE STATISTICS

According to the Joint United Nations Programme on HIV/AIDS (UNAIDS) and the World Health Organization (WHO) an estimated 35.3 million people were living with HIV in 2012 (Figure 1.1), 1.6 million of people died and 2.3 million were infected. The annual number of new infections has decreased by 50% in 26 countries between 2001 and 2012 [10]. The highest prevalence is still concentrated in sub-Saharan Africa, which has been the most affected region by the epidemic, followed by the Caribbean where HIV's incidence has decreased since the 2000s [10]. However, other epidemics like the ones in Belgium and Colombia are characterized with an increase in the number of new infections during the last years. Since these epidemics are the subject of this thesis, they will be discussed in detail in *chapter 6*.

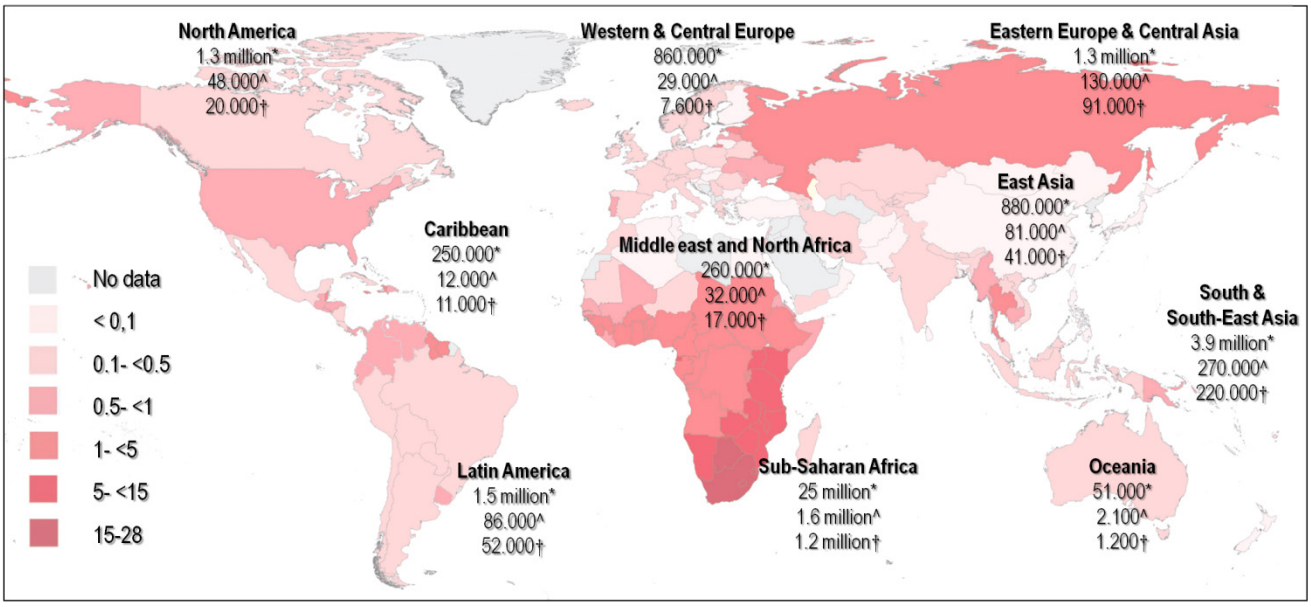


FIGURE 1. 1. PREVALENCE OF HIV WORLDWIDE

Figure 1.1: Prevalence of HIV worldwide. HIV prevalence is shown in shades of red according to UNAIDS-WHO report 2013. The estimated number of people living with HIV (*), number of new HIV infections (^) and number of AIDS-related deaths (†) in each geographical region are displayed**.** Adapted from UNAIDS [10].

1.3 THE HUMAN IMMUNODEFICIENCY VIRUS

According to the International Committee on Taxonomy of Viruses, HIV is classified into the genus *Lentivirus*, subfamily *Orthoretrovirinae*, and family *Retroviridae* [11]. Retroviruses cause chronic infections with a range of pathogenic manifestations such as cancer, immunodeficiency and central nervous system diseases. Retroviruses are characterized with a similar morphology, genomic structure 5'LTR-*gag-pol-env*-3'LTR and replicative cycle [12].

1.3.1 MORPHOLOGY AND GENOMIC STRUCTURE

The HIV mature virion is roughly spherical with a diameter of approximately 120 nm (Figure 1.2). The envelope is a lipid bilayer that is derived from the membrane of the host cell and that contains host proteins such as major histocompatibility antigens, actin and ubiquitin (yellow). At the viral surface spikes are formed by glycoproteins gp120 that interact with transmembrane proteins gp41 (blue). The inner surface of the envelope is lined with about 2,000 copies of matrix protein p17 (dark green). The cone-shaped capsid is located in the center of the virus and is formed by capsid protein p24 (brown). It

contains the essential enzymes protease (PR), reverse transcriptase (RT), integrase (IN) and two identical copies of unspliced positive single stranded ribonucleic acid (RNA) molecules that are bound to the nucleocapsid proteins p7 (red with white) [13].

The length of the HIV proviral genome is around 9,700 nucleotides. Multiple reading frames lead to nine partly overlapping genes. Three genes encode the structural and enzymatic proteins commonly found in retroviruses. The group specific antigen (*gag)* gene encodes the precursor p55 that is subsequently cleaved by PR, into the structural proteins matrix (MA or p17), capsid (CA or p24), nucleocapsid (NC or p7) and p6. The polymerase (*pol)* gene encodes PR (p11), RT (p66/p51), and IN (p31). The *env* gene encodes surface gp120 and transmembrane gp41 glycoproteins [12].

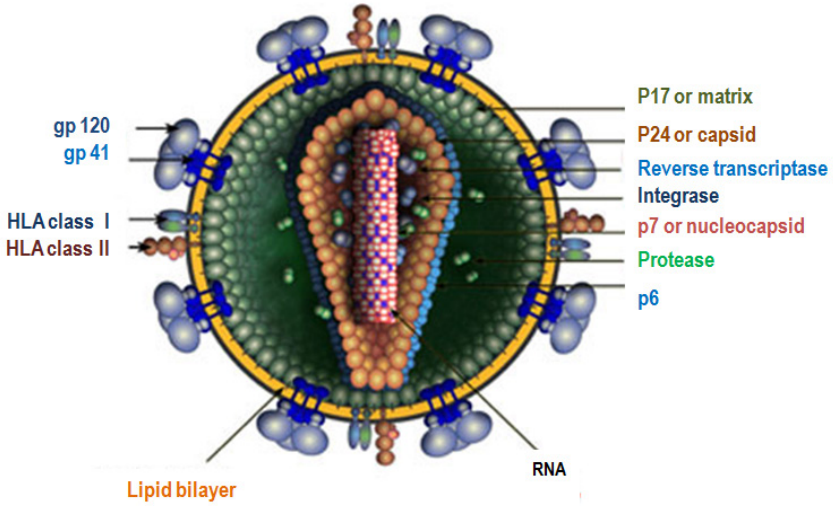


FIGURE 1. 2. MOFPHOLOGY OF HIV

Figure 1.2: Morphology of HIV. Adapted from http://www.hiv.lanl.gov

Two other genes are essential for virus propagation and replication, for instance *tat (*transactivator of HIV gene expression*)* and *rev* (regulatory factor). Four accessory or auxiliary genes are not essential *in vitro* but are necessary for spread and disease progression *in vivo*: *vif* (viral infectivity factor)*, vpr* (viral protein R in HIV-1 and its homolog *vpx* in HIV-2), *vpu* (viral protein U), and *nef (*before named negative factor*)*. Finally, the long terminal repeats (LTR) are necessary for transcription regulation [12].

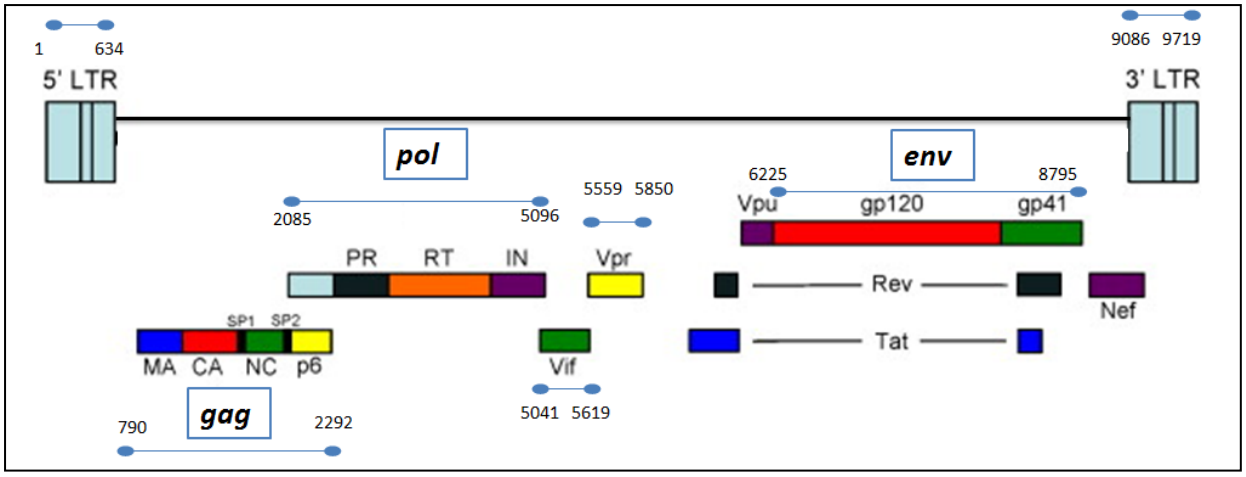


FIGURE 1. 3. GENOMIC STRUCTURE OF HIV

Figure 1.3: Genomic structure of HIV. The relative locations of the open reading frames *gag, pol, env, vif, vpr, vpu, nef, tat* and *rev* are shown. The blue lines indicate the positions according to HXB2 strain. Adapted from [14].

1.3.2 REPLICATION CYCLE

The entry of HIV requires the presence of the cluster of differentiation 4 (CD4), present on T-lymphocytes, macrophages, dendritic cells and brain microglia. The replication cycle includes ten important steps that are divided in two phases (Figure 1.4). The early phase includes binding to integration, whereas the late phase begins with transcription and ends with viral maturation [13].

EARLY PHASE

The interaction between gp120 and the amino-terminal domain of CD4 causes the **binding** between virus and target cell. Afterwards, conformational changes allow the binding between gp120 and the chemokine receptors CXC-chemokine receptor 4 (CXCR4) and the CC-chemokine receptor 5 (CCR5). The CCR5 co-receptor is generally used in early stages of the infection and CXCR4 in the late stages. Based on the use of these co-receptors, the tropism of viruses could be classified as X4, R5, R5+X4 and R5X4, which is important for therapeutic options. The interaction between CD4-gp120-CCR5/CXCR4 triggers conformational changes within gp41, which promotes the physical approach between viral and cellular membranes and **fusion** [12-14].

Entry of the HIV core into the cytoplasm causes the disintegration of the capsid with the aim to release the viral proteins. This process is known as **uncoating**. RT catalyzes the **reverse transcription** of two RNA molecules into one double-stranded deoxyribonucleic acid (DNA). The newly synthesized viral

DNA is **imported within the nucleus** as part of the preintegration complex that additionally includes viral proteins such as MA, IN, RT and Vpr [13].

Subsequently, the IN catalyzes the **integration** of proviral DNA into the host cell genome. This provirus serves as template for the synthesis of RNA that encodes the structural, enzymatic, regulatory and accessory proteins. The provirus is the basis for viral latency and reservoirs. Latency occurs when the CD4 cell turns into a resting status and limited transcription occurs. Activation of the resting cell results into complete transcription and a productive infection [15]. Viral reservoirs are a small pool of cells within lymphoid tissue and the brain, that provides a long-lived source of rebound viraemia [16].

LATE PHASE

Transcription is initiated at the LTR site, in which Tat is an essential transcriptional activator to increase viral RNA synthesis. This process generates three types of RNA: (i) unspliced or genomic RNA that is the messenger RNA (mRNA) for Gag and Gag-Pol precursors; (ii) singly spliced mRNAs that encodes Env precursor, Vpu, Vif and Vpr; and (iii) multiply spliced mRNAs that are translated into Tat, Nef and Rev [12, 17]. With the aim to translate all types of mRNA, the unspliced and singly spliced mRNAs need to be transported to the cytoplasm by Rev. Indeed, Rev acts by binding the rev responsive element, promoting the nuclear export and cycling between nucleus and cytoplasm of unspliced and singly spliced mRNA [12, 18].

Following the **translation** of the proteins, the **assembly** involves the Gag precursor polyprotein pr55, which binds to the plasma membrane and promotes the interaction of gag-gag proteins, encapsidation of RNA and the incorporation of Env proteins in the host membrane. Env precursor is translated in the rough endoplasmic reticulum and cleaved by host protease in the Golgi apparatus [17]. The **budding** of the immature virus includes interactions between host factors, and the p6 domain within Gag to hijack the endosomal machinery and release it. Viral **maturation** requires the cleavage of Gag and Pol polyproteins by PR in order to re-assemble the virion. A single cell can produce thousands of virus particles until the apoptosis of the host cell. Meanwhile new virus particles could start another replication cycle [12, 14].

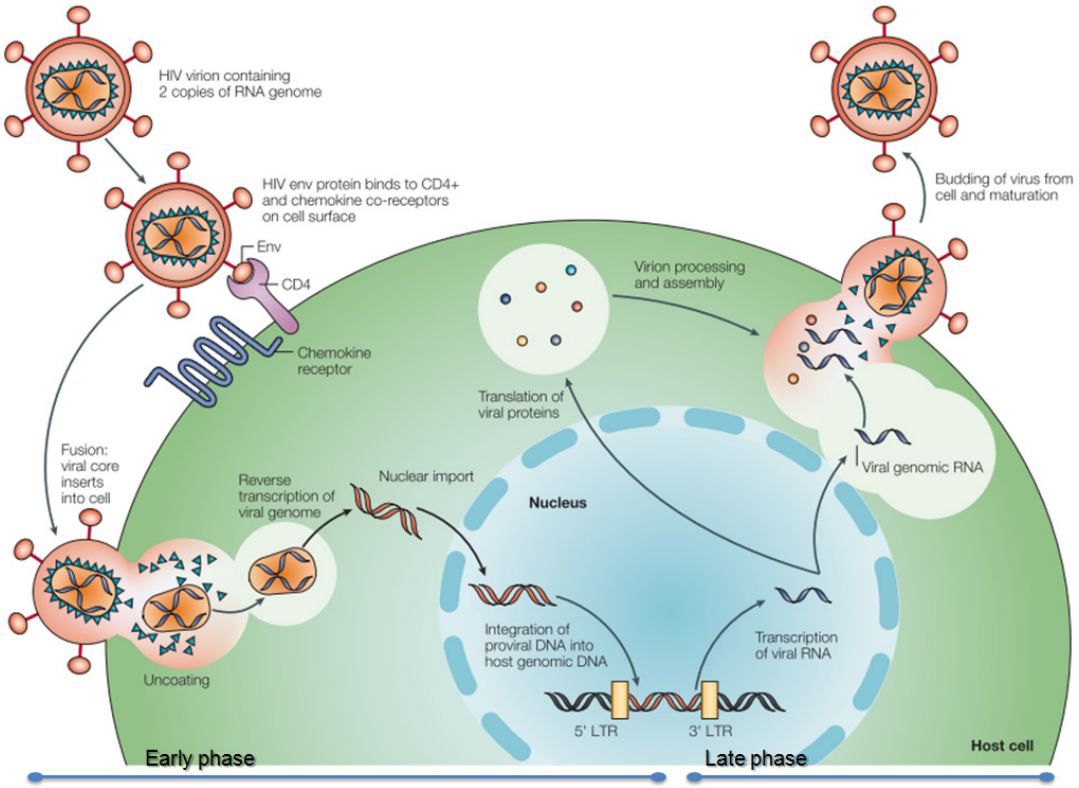


FIGURE 1. 4. REPLICATION CYCLE

Figure 1.4: Replication cycle. Adapted from [19].

1.4 MOLECULAR EPIDEMIOLOGY

The evolution of HIV is the consequence of several viral intrinsic mechanisms: the lack of proofreading activity of RT, the high mutation rate of 3.4 x 10^{-5} mutations per base pair per replication cycle [20, 21], the high viral production of 10^{10} virions per day [22, 23], the process of recombination with 7 to 30 crossovers per genome per round of replication, the flexible conformation of HIV proteins, the short generation time between 24-48 hours [24] and Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC) editing [25].

1.4.1 PHYLOGENETIC ANALYSES ON THE SURVEILLANCE OF HIV

Phylogenetics describes evolution and relationships among genes or genes fragments, by inferring their common ancestor. Therefore, it is necessary for phylogenetic analyses that sequences under investigation are from the same gene and without evidence of recombination [26]. Phylogenetic tools have been used in molecular epidemiology to investigate the origin of HIV and to classify subtypes [27] and their introduction and dissemination in different regions.

Phylogenetic analyses include the pre-processing of data from datasets or databases, such as the retrieval of particular fragments of a gene and an alignment to put homologous residues arranged in columns with the assistance of software packages such as ClustalW and Muscle [26, 28, 29]. The reconstruction of a tree could be performed by two different algorithm methods: optimality search criterion and clustering. Each one can use different kind of data such as discrete characters (morphological, physiological characteristics, nucleotides or amino acids) or a distance matrix of pairwise genetic dissimilarities [30]. Usually pairwise distance methods are complemented with evolutionary models that provide a statistical description of the substitution of nucleotides or amino acids in the sequence [26]. The most common used methods in phylogenetics of HIV are the distance pairwise method that uses clustering called Neighbor-joining (NJ) and the character state algorithm that use optimality criterion called Maximum likelihood (ML). Nonparametric bootstrap is used to determine the phylogenetic branch support in both methods [31]. However, the approximate likelihood ratio test (aLRT) is another fast alternative instead of nonparametric bootstrap when ML method is used [32]. The choice of the method depends on the kind of analyses, for instance NJ method works well for subtyping HIV sequences [33] (see section 1.4.4), whereas ML method is used in cluster transmission analyses because the exhaustive searching of different tree topologies (see section 1.7) [26, 34-39].

Whereas NJ or ML methods construct one tree, Bayesian methods integrate all plausible trees and provide confidence intervals (or Bayesian credible intervals) for any evolutionary relationship [26]. This is summarized in the maximum clade credibility (MCC) tree that also shows different traits, for instance the posterior distribution or probability. These techniques have several applications, for instance cluster transmission analyses (see section 1.7) [40-43], time-scaled phylogenies [44-48], and lately, phylogeographic diffusion models [49-51]. Phylogeography is based on the measurable imprint on the genome of rapidly evolving viruses that occurs simultaneously with the geographical dispersal over a specific period [52, 53]. Therefore it can provide insights about the viral introduction and dissemination in a region or the impact of human mobility on the spread of the HIV epidemic [51, 52].

1.4.2 ORIGIN OF HIV

Based upon clinical records, colonial literature and phylogenetic analyses of primate and human lentiviruses, the origin of HIV has been related with bushmeat hunting and keeping primates as pets [54-56]. HIV-1 group M (major) and group N (nonoutlier) originated from Simian Immunodeficiency Virus (SIV) (Figure 1.5), which was isolated from a subspecies of chimpanzee in West Africa, more specifically from *Pan troglodytes troglodytes* [57]. The origin of HIV-1 group O (outlier) and P (putative) was the SIV from gorillas in the same region [57-59].

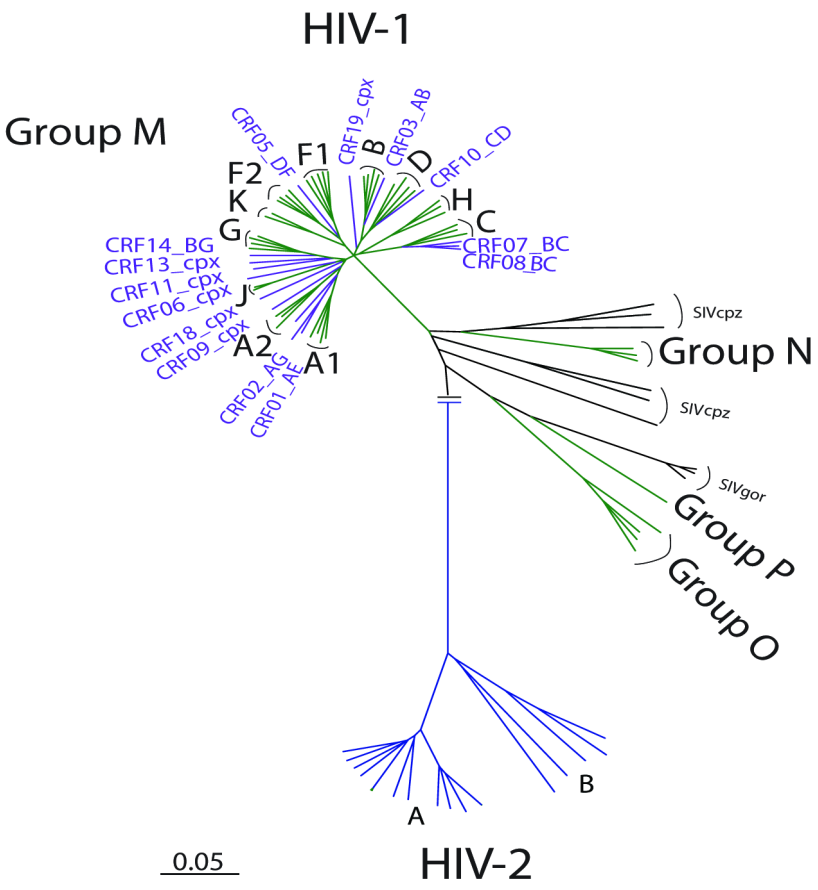


FIGURE 1. 5. CLASSIFICATION OF HIV

Figure 1.5: Classification of HIV. HIV is classified in types 1 (green) and 2 (blue). HIV-1 is divided in groups M, N, O and P. The group M includes nine subtypes, subtype A and F are divided in sub-subtypes e.g. A1 or F1. Some examples of circulating recombinant forms (CRFs) are shown in violet. HIV-2 is classified in eight groups (A-H), but only the most frequent groups A and B are shown [27, 58]. Tree adapted from [60].

The origin of HIV-1 group M was dated between 1900s and 1930s (Figure 1.6) [59, 61-64]. Two recent articles support its origin in the early 1900s. One included the phylogenetic analysis including the two oldest samples obtained from patients in Kinshasa, Democratic Republic of Congo [64]. According to a study including phylogenetic analysis, computational simulations and review of colonial literature [59], the main driver of the spread of HIV in the early 20th century was linked to sexual transmitted diseases, especially ulcerative diseases, in the heterosexual risk transmission group, whereas the increase of city population size and the low frequency of circumcision at that time had little effect in the initial viral spread [59].

HIV-2 originated from interspecies transmission between *Sooty mangabey* monkeys (*Cercocebus atys atys*) and humans. Each HIV-2 group (A-H) originated from independent cross-species transmission events from this monkey species infected with SIV [65]. HIV-2 group A originated around 1930s and group B around 1935s (Figure 1.6) [51, 63, 66, 67].

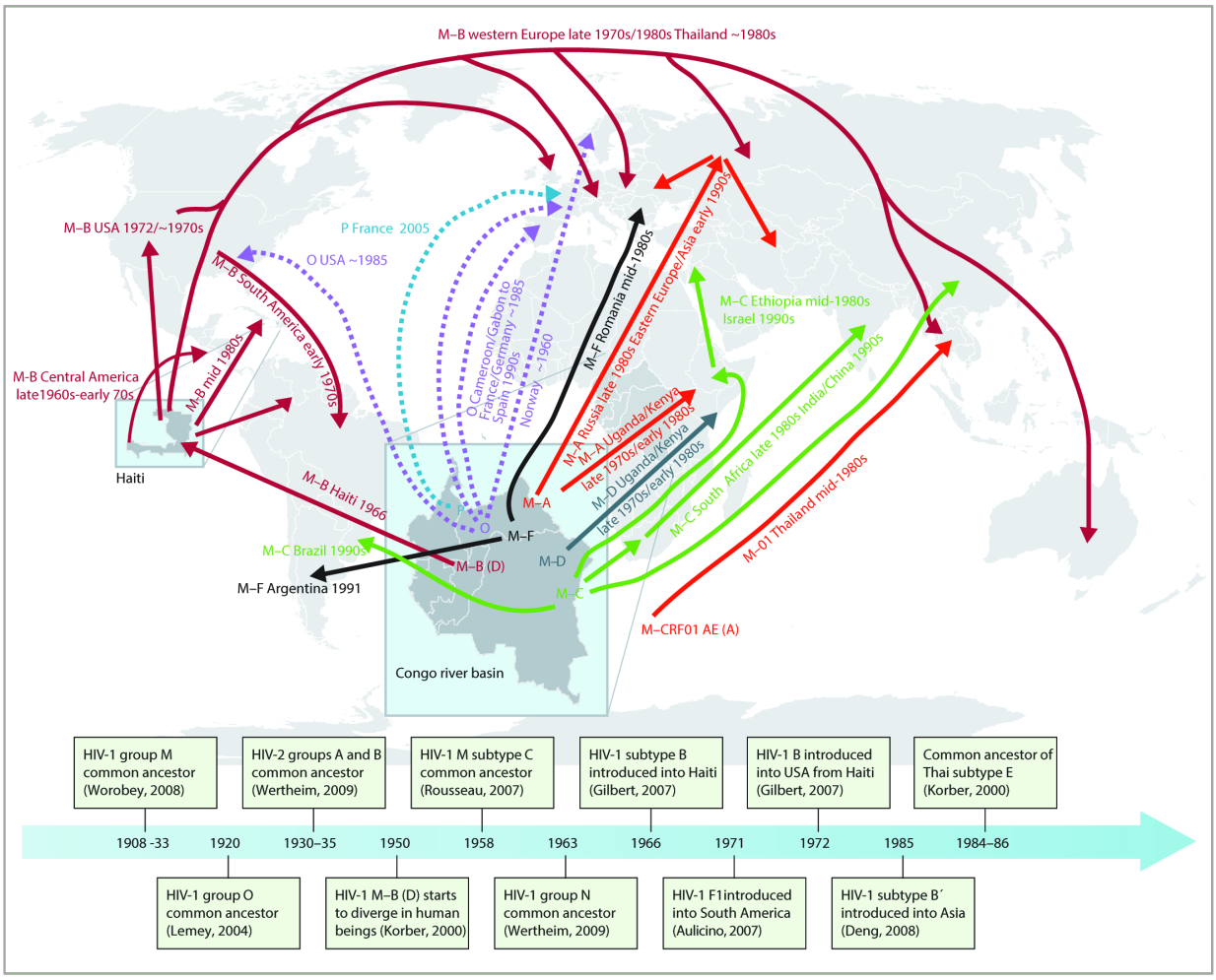


FIGURE 1. 6. ESTIMATED TIME LINE OF THE GLOBAL SPREAD OF HIV

Figure 1.6: Estimated time line of the global spread of HIV. Adapted from [60]. Additional subtype B data were adapted from [47, 48, 68-70]

1.4.3 MOLECULAR EPIDEMIOLOGY

Group M is the predominant circulating HIV-1 group, that is classified in subtypes A-D, F-H, J and K (Figure 1.5) [27]. The circulating recombinant forms (CRFs) are recombinants for which the full length sequences are fully characterized and specific breakpoints can be found in at least three people epidemiologically unlinked [27]. These are denominated according to the subtypes that compose the recombinant, and numbered according to the order of description (Figure 1.5). Up to date, there are 61 CRFs described in HIV-1 and one in HIV-2 (http://www.hiv.lanl.gov, accessed in February 2014). Additionally, inter-subtype recombination is common resulting into unique recombinant forms (URFs) within single patients [23].

The rare HIV-1 groups N, O and P have been identified mainly in Cameroon [58, 71, 72]. HIV-2 groups A and B have spread in West Africa causing 1-2 million of cases, whereas the other HIV-2 groups (C-H) have been rarely isolated [10, 73, 74].

To estimate HIV diversity large amounts of published and unpublished data were summarized and weighted according to the number of people living with HIV in each country by the WHO network of HIV Isolation and Characterization [75]. Subtype C was responsible for about 50% of HIV-1 infections worldwide between 2000 and 2007 (Figure 1.7). It was dominant in Southern and East Africa, India and Oceania. Subtype B was the most frequent subtype in Western and Central Europe and America. Regarding the molecular epidemiology of Belgium and Colombia, subtype B is predominant in both countries [76-79]. However, the percentage of subtype B is around 50% according to the last nationwide survey carried out in Belgium until 2006 [78]. In contrast, subtype B accounts for around 99% of the HIV infections in Colombia until 2002 [76, 77, 79]. The epidemiology of these two countries will be further discussed in *chapter 6*.

Several studies have reported on the origin of subtypes around the world [45, 61, 80, 81]. The most studied and prevalent subtype in America and Europe is B. The subtype B epidemic started to spread from Central Africa to Haiti around 1966 (1962-1970), before dispersing in a single migration to the USA around 1969 (1966-1972) and to Trinidad and Tobago around 1973 (1970-1976). Phylogenetic analysis showed that single chance events, ecological interactions and specific population bottlenecks and founder effects influenced the spread of subtype B in the Americas [48]. A recent study also suggested that the epidemic in South America was influenced by migration and multiple introductions from Caribbean but also by immigration to USA [69]. The spread of subtype B in Europe was related with

North American men who have sex with men (MSM) and intravenous drug users (IVDU) [82-84]. Most of the migratory pathways of subtype B in Europe are bidirectional. However, some countries are sources and others are sinks. For instance, tourism could partially explain the viral dispersal from Greece, Portugal, Serbia and Spain to Central Europe. Similarly, Austria, Belgium and Luxembourg have an imported epidemic due to the highest immigration from other European countries [83]. Concerning the origin of subtype B in Colombia and migration pathways, there is not information available (see *chapter 3*). On the other hand, multiple introductions from Africa and Asia were involved in the spread of non-B subtypes in Europe [60], and links with former colonies in Central-Africa influenced from the start the relative high prevalence of non-B subtypes in Belgium, Portugal and France [85].

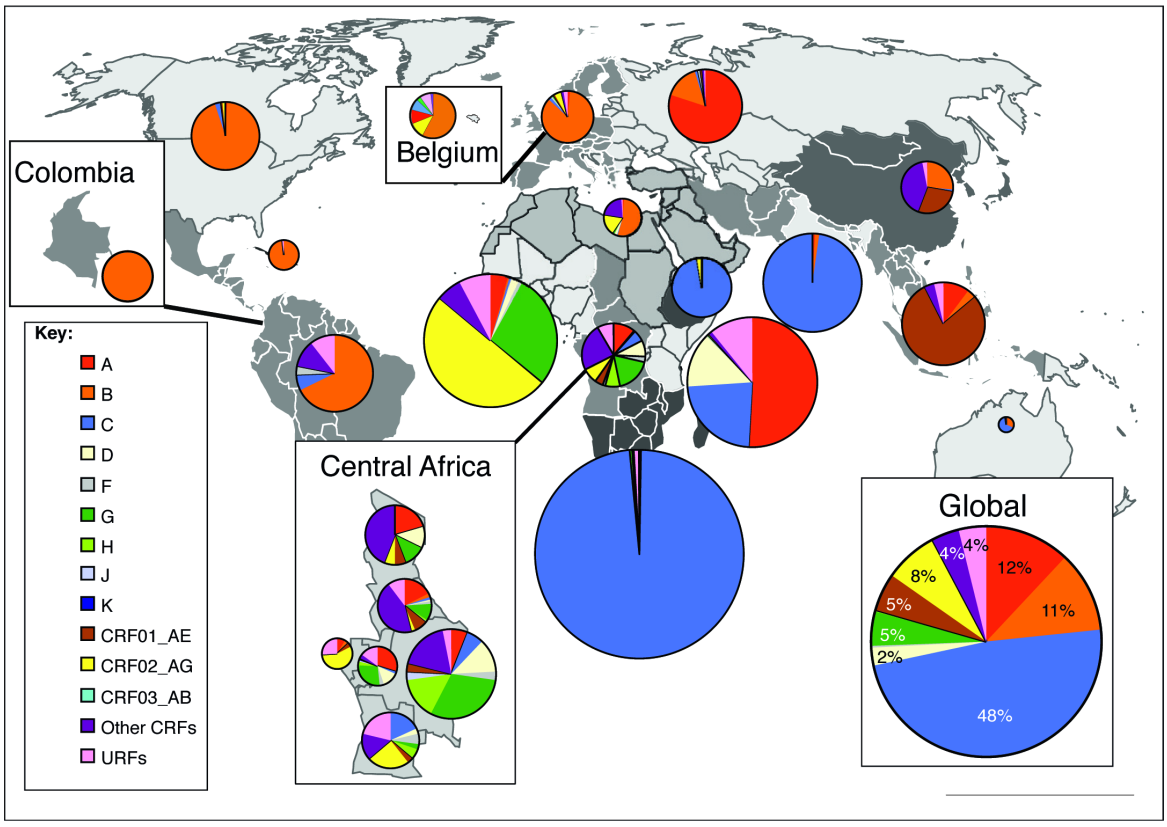


FIGURE 1. 7. EPIDEMIOLOGY OF SUBTYPES WORLDWIDE

Figure 1.7: Epidemiology of subtypes worldwide. Subtype C (blue), subtype B (orange) and subtype A (red) are the three most prevalent subtypes worldwide. The prevalence of subtypes varies with geographical localization. Adapted from [74]. Data for Colombia and Belgium were adapted from [75, 76, 78, 79].

The currently most prevalent subtype C originated around 1958 (1949-1960) [80]. A rapid spread in South Africa followed by the introductions to India and China could explain the high prevalence of this subtype (Figures 1.6 and 1.7) [60]. In the 1990s, it was also introduced to Brazil directly via East Africa [86], or via an intermediate step in UK [45]. Other multiple introductions have been reported for Italy, mainly from South America and India [87]. Although other few non-B subtypes have been studied [46, 47, 66, 88, 89], the increasing availability of more data from these subtypes is expected to unravel the routes of the HIV-1 spread.

1.4.4 METHODS OF SUBTYPING HIV

Although serological and genotypic methods have been used to identify HIV strains, the gold standard to classify subtypes is based upon the manual phylogenetic analysis of full-length viral sequences. Subtypes are phylogenetically equidistant, generating a starlike tree (Figure 1.5). The genetic variation between subtypes is usually 25-35% depending on the genes compared, whereas within a subtype it can be 15-20% [23]. Some subtypes have sub-subtypes with a smaller genetic distance (e.g. A1-A4 and F1-F2). The classification of HIV is based on phylogenetic analyses, in which the clustering pattern of viral sequences provides information about genetic similarity and evolutionary rate [27]. This classification is dynamic because it could change with the availability of new sequences. However, it is expensive to sequence the full-length viral genome and the analysis is time consuming. Since most available sequences are obtained within the framework of drug resistance testing, automatic tools have been developed with the aim to identify subtypes based upon PR and RT fragments [90-92]. In addition, the automated subtyping tools enable the quick classification of one to hundreds of sequences.

According to the methodology used to assign an HIV-1 clade to a query sequence, automatic tools can be divided into three main types. First, similarity-based tools use an alignment between the reference and the query sequence and calculate the match. Examples include the NCBI subtyping tool [93], Stanford [94], Geno2pheno [95] and EuResist (http://engine.euresist.org/data_analysis/viral_sequence/new). Second, statistical-based tools use prediction by partial matching compression algorithm like COntext-based Modeling for Expeditious Typing COMET [96], position-specific scoring matrices plus a statistical model such as STAR [97] or jumping profile Hidden Markov Models such as jpHMM [98]. Third, phylogeny-based tools like REGA [99] and SCUEAL [100]. Currently, it is not clear which tool is better to classify B or non B subtypes.

1.5 STAGES OF HIV-1 INFECTION

HIV is present in blood, genital fluids and breast milk. Transmission events are mainly the consequence of exposure to HIV at mucosal surfaces (80%) and the remaining transmissions are caused by percutaneous or intravenous inoculation. The per-act HIV transmission depends on the type of exposure. For instance, it is around 92.5% for blood transfusion, 22.5% for mother to child transmission (MTCT), 0.63% for IVDU, and 0.23% for percutaneous needle stick injury. Sexual transmission risk depends on the type of intercourse. Indeed, the per-act HIV transmission risk for receptive anal intercourse is 1.38%, for insertive anal intercourse 0.11%, for receptive penile-vaginal intercourse 0.08%, and for insertive penile-vaginal intercourse 0.04% [101]. The sexual transmission risk increases with genital ulcer diseases and high viral load, and during acute or late state of the disease, and decreases with circumcision, male condom use, and use of antiretroviral therapy (ART) [101].

Genetic diversity may affect the transmissibility [102, 103]. Subtype C seems to be more transmissible than subtypes A and D according to studies performed in pregnant women [104, 105] but not in heterosexual couples [106]. In Kenya, the MTCT rate was higher among women infected with subtype D than with subtype A [107], whereas in Tanzania subtype A was more likely to be transmitted than subtype D [104]. A study performed in discordant couples revealed a higher rate of heterosexual transmission of subtype A than subtype D [108] which seems to be in agreement with the higher number of infections of subtype A worldwide (Figure 1.7). In another study, CRF01_AE seemed also to be more transmissible than subtype B in IVDU [109].

After exposure, HIV initially replicates in mucosal, submucosal and lymphoreticular tissues and therefore it is not detected in plasma yet. The clinical manifestations of acute infection could be asymptomatic, similar to flu-like syndrome with pharyngitis, non-tender lymphadenopathies or mucocutaneous ulcerations by *Candida spp*, Epstein-Barr virus or Herpes simplex virus (acute retroviral syndrome) [110]. Other less frequent signs or symptoms can be presented such as generalized maculopapular rash, hematologic disturbances like anemia or thrombocytopenia, neurological disorders like aseptic meningitis and reactivation of Varicella-zoster virus [111].

The phases of HIV-1 infection can be categorized based on the sequential positivity of diagnostic assays (Fiebig stages I-VI, Figure 1.8A). The period between infection and the detection of viral RNA is named eclipse, and can last 7 to 21 days. At the end of the eclipse phase, HIV reaches the gut-associated lymphoid tissue, as well as other lymphoid tissues, and a peak viraemia concomitantly with

a decrease of CD4 T cells is observed [112]. In the initial phase of acute infection, the algorithm for the diagnosis of HIV-1 infection includes detection of viral RNA, p24 antigen or fourth generation combined HIV antigen/antibody immunoassay [113, 114]. The viral burden is high during acute infection. Therefore, transmission of HIV is mainly driven by recently infected individuals, who account for 5 to 50% of the transmissions [34, 102, 103, 115-117]. The seroconversion occurs in Fiebig stage III when viral-specific antibodies are detected in serum [118]. The viral load gradually decreases over 12-20 weeks and reaches a viral set point [119, 120]. Meanwhile the pressure of adaptive immune responses selects mutants and virus diversification occurs.

The Fiebig scale depends on the used test methods [121]. Ideally, these methods should be able to detect all subtypes, although some assays have a lower performance [122]. False-negative results have been reported in fourth-generation immunoassays in subtypes A, C, F, H, CRF01_AE and O. Similarly, false negatives for subtypes B, C and F have been described for gp41 immunoassays [123, 124]. These results are the consequence of differences in viral epitopes [123, 125]. Rapid HIV tests have also shown low sensitivity for D, F, H, CRF01_AG, O and HIV-2 [122, 125-128]. In general, viral load assays perform well for B and non B subtypes, but the assays based on integrase seem to perform better on viral strains belonging to M, O and N than assays using primers and probes in the *gag* gene [122, 129] .

The early chronic infection phase is frequently asymptomatic but some individuals do display clinical manifestations like persistent generalized lymphadenopathy. The duration of the clinically latent phase is variable but it could last between 1 and 10 years (Figure 1.8). In the meantime the CD4 count declines on average with 50 cells /mm^3 per year, which is influenced by the viral burden [130]. The final AIDS stage is reached when there are less than 200 CD4 cells /mm^3 or opportunistic infections and malignancies occur [111, 131]. When the CD4 count reaches 50 cells/mm^3 and ART is not provided, the median survival is 12 to 18 months [132].

Some patients are the exception of the described profile in clinical progression and are named long-term nonprogressors. They can be infected for more than 10 years without any clinical symptom and CD4 count above 500 cells/mm^3 without receiving ART [133]. Similarly, the “elite controllers” are characterized with a spontaneous virological control defined as viraemias below 50 copies/mL [134]. The genetic background and differences in immunological response, such as cytotoxic T cell response and HLA class I alleles, have been described as possible causes of this particular phenomenon [133].

Despite 30 years of research, the impact of subtypes on disease progression is still a matter of debate due to conflicting data and the presence of several confounding factors [107, 135]. Several reports from Uganda, Tanzania, Kenya and London have shown a faster progression in people infected with subtype D than with other subtypes [136-138]. The dual or X4 tropism of subtype D, the more frequent formation of syncytium and the higher replication capacity of the *pol* region are factors that could explain the faster decrease of CD4, compared with other subtypes [137, 139, 140]. A cohort study performed in London could not detect differences in disease progression between subtypes A, B, C and CRF02_AG [141]. However, a systematic review showed that subtype C is also aggressive, followed by G, CRF01_AE, and CRF02_AG mainly in developing countries [135]. Recent reports on larger study populations also showed the rapid disease progression associated with subtype C, which may be related to the increased replication capacity [142, 143]. Similarly, CRF01_AE is associated with X4 tropism and consequently low CD4 count in MSM [144, 145]. The viral load has also been compared between different cohorts. People infected with CRF01_AE had threefold higher viral load after three months of seroconversion when compared to subtype B, but no difference was found after one year [146]. Two reports showed higher viral loads in subtype D infections [137, 139] but other two did not [141, 147], which raise the question whether other factors such as comorbidities, co-infections, constitution and socio-economic conditions could impact the different response and disease progression [135, 147]. A recent systematic review pointed to the heterogeneity of the outcome measures and suggested to use subtype A as comparator and the inclusion of some typical outcome measurements such as relative risk, odds ratio (OR) or hazard ratio [135]. The standardization of future studies might enable the pooling of different datasets and provide valuable information for clinicians in developing countries where 88% of the individuals infected with non-B subtypes are living.

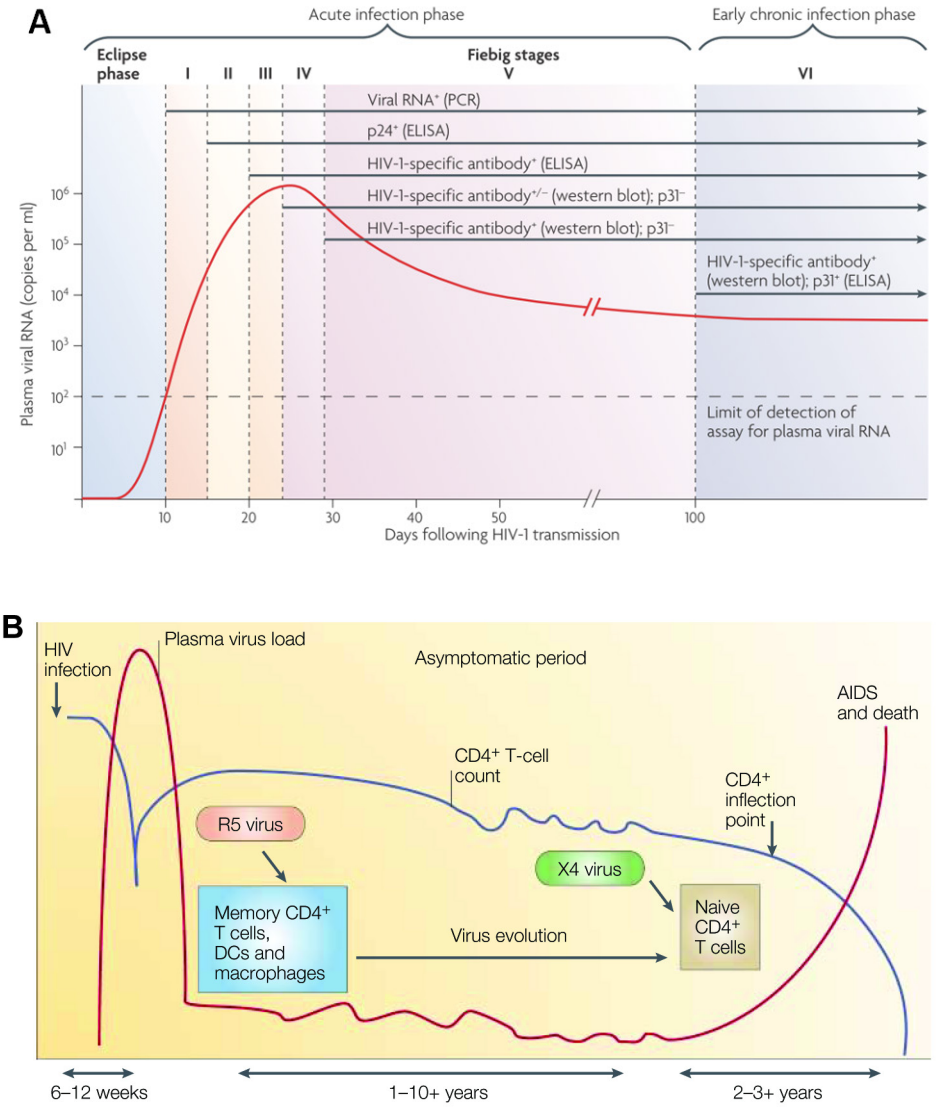


FIGURE 1. 8. PATHOGENESIS OF HIV-1

Figure 1.8: Pathogenesis of HIV-1. **A)** The Fiebig stages are represented by roman numbers and are classified according to a positive result for viral RNA measured by PCR (I), for p24 antigen (II) and for antibodies immunoglobulin M (IgM) measured by enzyme linked immunosorbent assay (ELISA) but not by western blot (III), for antibodies detected by ELISA and by western blot (still indeterminate) (IV), for antibodies detected by ELISA and by western blot (positive, but still p31 negative) (V), for antibodies detected by ELISA and by western blot (positive including for p31) (VI). **B)** The relationship between viral load (red line) and CD4+T cell (blue line) over time in an untreated patient. R5 virus is usually found during the early stages of the infection, whereas X4 viruses in the chronic stage. Abbreviation: DC: Dentritic cell, PCR: Polymerase chain reaction. Figures adapted from [120, 148].

1.6 ANTIRETROVIRAL THERAPY

Currently, there are six different classes and 28 drugs approved by USA Food and Drug Administration. The classes are nucleoside/nucleotide reverse transcriptase inhibitors (NRTI), NNRTI, protease inhibitors (PI), fusion inhibitors (FI), coreceptor antagonists (CRA) and integrase inhibitors (INI). These drugs must be combined in HAART that contains three active drugs, or at least two fully active compounds in case of salvage regimen [149, 150]. See table 1.1.

HAART could display suboptimal results due to side effects or toxicity, prior ineffective ART, infection with a drug resistant strain or suboptimal adherence [151]. The latter factor is the major reason of initial therapy failure. At least 95% of the doses must be taken with the aim to completely suppress viral replication [152, 153], although certain regimens based upon NNRTI and boosted-PIs could tolerate lapses due to their high potency and longer half-lives [154, 155]. Otherwise, residual replication could result into resistance and therapy failure. Treatment failure can be defined clinically (occurrence or recurrence of opportunistic infections or malignancies), immunologically (failure to achieve and maintain an adequate CD4 response despite virologic suppression) or virologically (inability to achieve or maintain suppression of viral load <50 copies/mL) [114, 156]. As viral load is the most important measure of response to therapy, the most recent WHO guidelines have included the assessment of viral load in the follow-up of patients treated in low-middle income countries. In this guideline, virological failure was defined as a persistent HIV viral load above 1,000 copies/mL, measured in two consecutive tests within an interval of 3 months [157]. Given these definitions and the roll-out of ART in resource-limited settings, it is essential to have access to viral load assays that can accurately quantify the extent of viral replication for all subtypes and viral variants, as lower sensitivities affect clinical decision making and could compromise subsequent therapeutic lines [122, 129].

Drug class	Drugs	Release year	Activity	Mechanisms of resistance
NRTI	Zidovudine (ZDV) Didanosine (ddI) Zalcitabine (ddC) Stavudine (d4T) Lamivudine (3TC) Abacavir (ABC) Tenofovir (TDF) Emtricitabine (FTC)	1987 1991 1992 1994 1995 1998 2001 2003	NRTI are mimetics of nucleosides/nucleotides and bind to the active site of the polymerase domain of RT, once incorporated they inhibit the synthesis of viral DNA.	(i) Thymidine analogue mutations promote ATP and pyrophosphate-mediated excision of the incorporated nucleoside analogue. (ii) Mutations interfere with the binding and/or incorporation of the nucleoside analogue into the growing viral DNA chain.
NNRTI	Nevirapine (NVP) Delavirdine (DLV) Efavirenz (EFV) Etravirine (ETR) Rilpivirine (RPV)	1996 1997 1998 2008 2011	NNRTI are designed to bind to an allosteric hydrophobic pocket within RT, modifying its structure and impairing its catalytic activities.	Mutations reduce affinity of the inhibitors for the binding site at the enzyme. Single mutations generally are sufficient to induce resistance.
PI	Saquinavir (SQV) Ritonavir (RTV) Indinavir (IDV) Nelfinavir (NFV) Amprenavir (APV) Lopinavir (LPV) Atazanavir (ATV) Fosamprenavir (FPV) Tipranavir (TPV) Darunavir (DRV)	1995 1996 1996 1997 1999 2000 2003 2003 2005 2006	Except for TPV, PI are mimetics of viral peptides and bind to the PR active site, preventing viral maturation.	Mutations reduce affinity of the inhibitors for the enzyme. PIs require accumulation of multiple mutations to reduce susceptibility. Major mutations impair binding affinity whereas minor mutations often improve replication capacity.
FI	Enfuvirtide (T20)	2003	T20 is a small peptide that binds to gp41 and prevents the fusion between viral envelope and cellular membranes.	Mutations affect a domain of gp41 whose interaction promotes membrane fusion.
CRA	Maraviroc (MVC)	2007	MVC binds to CCR5 and prevents gp120 attachment to CCR5.	Change in coreceptor use (mainly due to genotypic changes within V3 loop) or mutations that allow binding of CCR5-MVC complex.
INI	Raltegravir (RAL) Elvitegravir (ETG) Dolutegravir (DTG)	2007 2012 2013	INI bind to the viral integrase and prevent the integration of the viral double-stranded cDNA into the host cellular genome.	Mutations affect the vicinity of the INIs-binding pocket in the catalytic core domain of the enzyme.

Table 1.1: Activity and resistance mechanism of antiretroviral drugs approved by United States Food and Drug Administration. Abbreviations: ATP: Adenosine triphosphate. Adapted from [149, 150, 158].

1.6.1 TREATMENT GOALS

INDIVIDUAL LEVEL

Since ART is not able to eradicate HIV infection due to long-term viral reservoirs, the aim of therapy is the preservation or restoration of the immune system. Therefore, the main goal is the suppression of HIV viral load with the consequent reduction of HIV-associated morbidity and improvement of the quality of life and prolongation of survival [114].

At the individual-level, the viral repression requires a highly potent ART with the use of three active drugs from at least two different classes [114, 156]. The clinical evaluation of the patient, the laboratory tests and the readiness of the patient to start ART will guide clinicians to choose the first line therapy (Table 1.2). The regimen in high-income countries is individualized because upon the availability of more antiretrovirals, genotyping resistance testing and HLA-B*5701 tests. The detection of this allele is associated with the hypersensitivity to ABC, which includes mainly rash, fever, gastrointestinal and respiratory symptoms that could become severe and potentially life-threatening [159]. Moreover, the guidelines are moving towards an earlier start of ART according to the CD4 count given the increasing evidence of the association of viremia with cardiovascular, kidney and liver diseases, cancer or other neurological complications. The exceptions are USA, Brazil and France, where ART is recommended for all patients diagnosed with HIV [114, 160-162].

In middle and low-income countries where the scaling-up of ART should be optimized, the WHO recommends to start therapy with one preferred first-line regimen based on the balance between cost-effectiveness and clinical evidence (Table 1.2) [157]. Moreover, laboratory tests such as baseline resistance testing or HLA-B*5701 are not recommended. Instead, the scaling up of viral load testing is considered to be a priority. Nevertheless, every country can adapt these WHO guidelines to the local setting. For instance, the Colombian guideline includes ABC or boosted PI as first-line therapy, whereas baseline resistance testing is not recommended [163]. It is only recommended when therapeutic failure occurs (definition according to European or USA guidelines described previously).

COMMUNITY LEVEL

The community-level goal of ART is based on the prevention of HIV transmission. For instance, ART has been used since mid-90s in HIV-infected pregnant women to prevent perinatal transmission because it decreases the risk of transmission from 20% to 0.1% or 0.5% when viral load is suppressed after the 28th week of gestation [164, 165]. Several studies showed that ART reduced the community viral load since the mid-00 and they suggested that ART could therefore decrease HIV incidence [166-168]. This concept was conclusive after the HIV Prevention Trial Network (HPTN) 052 randomized study identified a 96% of reduction of HIV transmission in serodiscordant couples under ART [169]. Mathematical models also suggest that the use of ART may decrease the incidence of HIV [170]. Therefore, the recommendation to start ART regardless of the CD4 count (concept known as Treatment as Prevention (TasP)) has been implemented in USA, French and Brazilian guidelines [114, 161, 162]. However, the success of this strategy will depend on earlier HIV diagnosis or improved test and treat strategies that propose universal testing and treatment. Even more, the engagement in care and adherence are key factors to reach a decrease in new infections. For instance, an observational study developed in an upper-middle income country like China showed a 26% reduction of transmission in heterosexual couples. In this study, resistance, non-adherence, probably lower linkage and retention in care partially explained the lower effect of ART on transmission when compared with the HPTN-052 trial [171].

Guideline	AIDS or HIV-Related Symptoms	CD4+ Cell Count			Recommended First line ARV regimen				Alternative first line ARV regimen			
		< 350	350-500	> 500	NRTI	NNRTI	PI	INI	NRTI	NNRTI	PI	INI
EACS	Yes	Yes	Consider	Consider	ABC/3TC or TDF/FTC	EFV or RPV	ATV/r or DRV/r	RAL	TDF/3TC or ZDV/3TC	NVP	FPV/r or LPV/r or SQV/r	EVG‡/COBI
USA DHHS	Yes	Yes	Yes	Yes	TDF/FTC*	EFV or RPV	ATV/r or DRV/r	RAL or EVG/COBI or DTG	ABC/3TC	-	ATV/r or DRV/r or FPV/r or LPV/r	RAL
IAS-USA	Yes	Yes	Yes	Yes	ABC/3TC or TDF/FTC	EFV	ATV/r or DRV/r	RAL	ABC/3TC or TDF/FTC or ZDV/3TC	NVP	DRV/r or LPV/r	RAL or EVG/COBI or DTG
WHO	Yes	Yes	Yes	Not addressed	TDF/FTC*	EFV	-	-	TDF/FTC* or ZDV/3TC	NVP or EFV	†	-
Colombia	Yes	Yes	Yes	Not addressed	ABC/3TC or TDF/FTC	EFV	ATV/r or DRV/r	RAL	ZDV/3TC	NVP	ATV/r or LPV/r or FPV/r	-

Table 1.2: When and what to start first line ART for adults. Several considerations such as pregnancy (EFV), HIV viral load (RPV, ABC), CD4 count (NVP), HLA-B*5701 negative (ABC), other medications, comorbidities, co-infections, intolerance to other drugs should be considered to initiate ART. *3TC may be replaced for FTC or vice versa, except for EVG/Cobi which is available as fixed-dose combination with TDF/FTC and RPV/TDF/FTC, † PI and ABC are used in special circumstances such as toxicities, drug interactions, drug supply, etcetera. Abbreviations: EACS: European AIDS Clinical Society, DHHS: Department of Health and Human Services, IAS: International Antiviral Society. Adapted from [114, 156, 157, 160, 163].

ART as prevention is also used in HIV-negative people with strategies such as pre-exposure prophylaxis (PrEP), microbicides and post-exposure prophylaxis (PEP). The latter strategy was the first one to be implemented. A case-control study in the middle 90's showed that the administration of ZDV after exposure decreased the risk of HIV infection by 81% in health-care personnel [172]. However, the effect of PEP depends on the type of exposure and the host-source co-factors such as viral load and co-infections. The time of ART initiation should be as soon as possible because it aims to prevent a systemic infection that occurs approximately between 48-72 hours [173, 174]. It should last 4 weeks according to studies in animal models [175]. The efficacy of PEP has not been evaluated in-depth because randomized trials would be unethical. However, observational studies suggest that PEP in non-occupational exposure was tolerated but suboptimal adherence, unreported or subsequent exposures limit the evaluation of the public health impact of this measurement [176]. Despite the limited evidence, PEP is included in several guidelines. The specific ART combination depends on the source's clinical history, but if it is unknown the preferred ART regimen after for non-occupational exposures includes NRTI+PI like TDF/FTC or ZDV/3TV +LPV/r in Europe and USA [156, 177]. In addition, USA guideline recommends other regimens for occupational exposures. In this instance, the recommended ART is TDF+FTC+RAL but also other alternatives such as NRTI (ZDV+3TC), NNRTI (ETV or RPV), and PI (DRV/r, ATV/r, and LPV/r) can be used [178]. In contrast, WHO recommends the first-line therapy available in the country [157].

PrEP implies the intake of oral ART before the exposure in individuals who are at risk of getting HIV, which includes TDF+FTC, once daily. Studies showed a reduction of 92% of the risk of infection in MSM (iPREX study, [179]) and 90% in serodiscordant couples (Partners PrEP study, [180-182]) when levels of ART were detectable. However, this reduction is highly dependent on adherence. In the iPrEx study, a 50% reduction was reported when the adherence was ≥50% [183]. Other studies confirmed the impact of adherence and logistics problems. Low compliance resulted in a reduction of 62% in heterosexuals (TDF2study, [184]), no reduction in trials that included females (FEM-PrEP trial or VOICE, [185, 186]), or inconclusive data (West Africa trial, [187]). On the other hand, TDF showed an efficacy of 50% in IVDU. When only adherent participants with drug levels detectable in plasma were considered, the efficacy increased to 73.5% (Bankok TDF study, [188]). Given the results of these trials, PrEP has been included in the USA guideline for populations at risk of acquiring HIV infection [189], whereas the WHO guideline recommends PrEP only for serodiscordant couples or in men and transgender women when other prevention choices are needed [157]. European guideline does not address PrEP because the discrepant results of effectiveness and the need of more evidence in specific populations [176, 190].

Finally, microbicides are still in research and are not yet approved for broader use in clinical practice. Since 1987 broad spectrum of microbicides with contraceptive properties were investigated but the results showed no benefit [191]. Since the 2000s, the use of ART in a gel formulation was investigated but results were conflicting. For instance, the CAPRISA study showed that 1% TDF gel reduced the risk of HIV acquisition by 39%, and this reduction reached even 54% when the adherence was >80% [192]. However, the VOICE study that also included TDF in gel was stopped because of the low adherence and futility [186]. Another alternative that is still subject of on-going microbicide trials is dapivirine, a NNRTI that had poor oral bioavailability. The studies using diaphragms or intra-vaginal rings/tablets are however in early phases [193]. Similarly, rectal microbicides are still in early study phases (MTN-017 trial) [194].

CURE

Several initiatives were taken to find a cure for HIV without success [195]. Up until now, there is no curative treatment nor vaccine [196], although hopes for a cure have risen after the case report of the Berlin patient in 2008. This patient had an allogeneic bone marrow transplantation with a dysfunctional CC-chemokine receptor 5 (CCR5), which is crucial for HIV entry (mutation known as *CCR5* $\Delta_{32/}\Delta_{32}$ deletion) [197]. After the discontinuation of HAART, the HIV-1 virus has been undetectable in peripheral blood, bone marrow, cerebrospinal fluid, brain or rectal mucosa up to now, which implied a sterilizing cure of HIV for the first time [198]. In 2012, other two patients heterozygous for *ccr5 Δ 32* mutation had apparent functional cure or long-term control of HIV replication after allogeneic bone narrow transplantation. However, these patients continued ART despite the chemotherapy, and achieved undetectable viral load in plasma and reservoirs for more than two years [199-201]. When ART was subsequently interrupted, a virological rebound was observed in these so-called Boston patients [202]. In 2013, a case known as the Mississippi baby showed a possible functional cure after the administration of aggressive ART soon after birth, and later interruption of treatment by the mother [203, 204]. Although HIV remained undetectable in the baby for more than two years without ART, the same mother's HIV strain was detected and the CD4 decreased according with an official announcement last July. Other strategy such as ART in primary infection has led to the discovery of post-treatment controllers. ART was discontinued after a median of 36 months, but nevertheless they achieved sustainable undetectable viremia for several years [205, 206], even if some of them were characterized with poorer CD8+ T cell responses and unfavorable genetic backgrounds (e.g. HLA-B*35 and HLA-B*07 that are associated with faster disease progression) [207]. More basic and clinical research is needed to evaluate these strategies [208].

1.6.2 RESISTANCE

Although clinicians have access to potent drugs that could result in long-term viral suppression, they currently face problems with patients not responding to therapy because of drug resistant HIV. The main cause of HIV drug resistance is acquired or secondary resistance. Antiviral drug resistance arises in patients whose therapy or adherence is suboptimal, resulting in residual replication and allowing an increase in genetic variability and the selection of variants with reduced drug susceptibility. Minority drug resistant variants can ultimately replace the previously dominant susceptible population and result in resistance to ART [149, 209]. A less common source of resistance is transmitted drug resistance (TDR), also called primary resistance, when an untreated patient is infected with a resistant variant, either from a treated patient with inadequate viral suppression and acquired resistance or from another untreated patient with a virus carrying resistant strains.

DETECTION OF RESISTANCE

According to the USA and European guidelines, the resistance test should be performed in the earliest sample of chronic naive patients or if not available, before starting therapy [114, 156, 209]. In the case of acute infected individuals, the test is recommended even if ART is deferred [156] but, the USA guideline considers that it should be repeated before starting ART [114]. Resistance testing should also be performed at virologic failure, to evaluate whether resistance is the cause of failure, and to assess which drugs are still active. According to the European guidelines, resistance testing can be considered when a successful NNRTI-regimen is inappropriately interrupted [209]. The resistance test should also be used to guide PEP, on a sample from the source [114, 156, 209]. All guidelines advice to use resistance testing when a pregnant women is going to start ART or has detectable viral load [114, 209]. The test is usually successful when viral load is > 300-500 copies/mL, but specialized laboratories can perform the test with lower viral loads [156].

Drug resistance can be detected phenotypically or genotypically [114, 209]. Phenotypic assays measure the ability of a virus to grow *in vitro* in different concentrations of an inhibitor. Gene fragments, usually RT and PR, are inserted into a clone, which generates viruses that express these genes. Replication of these viruses is compared with the replication of a reference HIV strain. The drug concentration that inhibits viral replication by 50% (IC50) is calculated and the results are reported as a fold-change in drug susceptibility with respect to the reference strain [114]. On the other hand, genotypic assays identify drug resistance mutations in the viral sequence. It is the preferred test in routine practice due to easier implementation, lower costs, better clinical validation and faster turn-around time [209, 210].

Genotyping testing is widely used for RT and PR because the antiretrovirals that target those genes have been used for a long time. However, there are situations where also resistance testing for IN and *env* is recommended or considered. In the case of IN, the use of resistance test is recommended when patients are failing to these drugs [114, 209]. However, performing baseline genotypic testing is not advised because the transmission of INI-resistance is still rare [114, 209]. Regarding *env*, resistance testing is performed on gp41 to evaluate the susceptibility of a virus to T20, which is rarely used nowadays. On the other hand, resistance testing is not performed for MVC because resistance is also rare [114]. However, before starting therapy with MVC or when virological failure occurs, a co-receptor tropism assay should be performed, which can be phenotypic or genotypic [114, 156]. Phenotypic assays include the MT-2 assay, which is the co-culture of MT-2 cells with the patient's mononuclear cells, X4-tropic viruses are detected because they induce syncytia in the infected cells [211]. Another method includes the amplification of the *env* gene, the generation of recombinant viruses and the infection with these recombinants of human cell lines expressing X4 or R5 co-receptors. There are commercial and non commercial examples such as Trofile, XTrack^{C}/PhenX-R, and the Toulouse Tropism Test. Phenotypic methods, however, have the disadvantage that require high biosafety standards [211]. On the contrary, the tropism genotype testing includes the amplification of the third hypervariable loop (V3) of the co-receptor binding site using population sequencing or next-generation sequencing (NGS). This sequence is analyzed with a web-based bioinformatic algorithm that predicts the use of a co-receptor. According to the USA guideline, a phenotypic assay is usually preferred because it has better performance than the genotypic assay [114], whereas European guidelines recommend both assays given their good concordance [211].

Within the framework of HIV drug resistance surveillance, an expert committee and the WHO have developed a Surveillance Drug Resistance Mutations (SDRM) list to enable an accurate assessment of TDR and comparisons of TDR trends. The list has been updated periodically and excludes highly polymorphic amino acid variants [212]. Within the framework of clinical interpretation of resistance, several interpretation algorithms have been developed with the aim to predict susceptibility and consequently therapy response. The most commonly used genotypic interpretation algorithms are based on expert rule-based, and are HIV Drug Resistance Database (HIVdb; available at http://hivdb.stanford.edu) [94], REGA (available at http://sierra2.stanford.edu/sierra/servlet/JSierra?action=hivalgs or http://rega.kuleuven.be/cev/avd/software/software/rega-algorithm) and Agence National de Recerche sur le SIDA (ANRS; available at http://www.hivfrenchresistance.org) [213, 214]. These algorithms are based on rules designed by experts who retrieved information from publications

and databases [209], and attempt to predict virological response. The advantages of these systems include an easy clinical application, but they depend on the quality of the available scientific research and applicability of the chosen rules. Since different algorithms are based on slightly different considerations, discordances between algorithms exist, yet overall, the three algorithms perform equally well in comparative studies [215, 216]. Other interpretation algorithms are based on phenotype-genotype correlations, such as Geno2pheno (available at http://www.geno2pheno.org/index.php) [209]. This algorithm is trained on a database containing correlated genotypic and phenotypic resistance test results, complemented with information of clinical response. As such, the algorithm does not depend on the judgment of an expert which may be biased, for example by previous scientific literature [95]. Nevertheless, they also depend on the available data and the reliability may be affected by new combinations of mutations [209].

EPIDEMIOLOGY OF HIV RESISTANCE

TDR can occur through sexual, parenteral and vertical routes and can depend on several factors like type of drugs used, viral subtype, ratio of wild-type to resistant variants, and viral load [217]. TDR can disappear gradually over months or years but can remain archived in proviral DNA [218]. The rate of disappearance can depend on viral fitness, selective drug pressure, type of mutations or superinfections. For instance, M184IV as a singleton is usually underestimated due to the rapid replacement and fitness cost [219, 220], in contrast to K103N or some thymidine analogue mutations (TAMs) that can persist over time [221].

The epidemiology of HIV drug resistance has been assessed in drug-naive and treated patients. The prevalence numbers can vary because several factors: intrinsic characteristics of the population analyzed (age, risk factor of transmission, geographical and socioeconomic factors), design of the study (method of sampling, type of study, inclusion or exclusion criteria), clinical characteristics (recent or chronic infection in naive patients, immunological status, viral load), determination of resistance (type of sample, method of sequencing, quality of the sequence, rules used for analysis of mutations), ART-related factors in the treated population (mono- or bitherapy history, HAART scale up, delay of start therapy, availability of new compounds, pharmacogenetics, toxicity, adherence problems, percentages of drug failure) and viral factors (subtype, viral fitness) [222-224].

The highest prevalence of acquired drug resistance was observed in Western Europe and North America where drugs had been prescribed in the format of mono- or bitherapy. Prevalence of drug resistance in these areas were between 70-80% [216, 225]. However, in the last years drug resistance in treated patients have decreased and resistance to novel drugs remains limited [226-229]. For

instance 1% of the cases in 2008 did not have other therapeutic options in seven countries from Western Europe, but nowadays almost every patient has an alternative therapy due to less extensive resistance profiles and the availability of new approved antiretrovirals [230].

Initially, a low prevalence of TDR was observed in high-income countries [231-234]. Subsequently, TDR levels increased up to 20% in the era of mono- and biotherapy [183, 235, 236]. Therefore, prospective and representative population-based studies were set up to monitor the extent and dynamics of TDR in Western Europe, USA and Canada [237-240]. They revealed lower prevalence values around 8-9% and recent studies have shown a stable trend of overall TDR in Western Europe and USA [241-243]. In contrast, according to a systematic review that included studies with sampling years until 2008 TDR decreased in Europe and increased in USA [244]. Additionally, regional differences were shown for TDR associated with the different drug classes [244]. For instance, no significant trends were reported in Belgium, and Sweden [78, 245], a slight increase of TDR related with NRTI was described in UK and Italy [246, 247], a slight NNRTI increase was described in Germany [248] and a significant decrease of TDR for PI was shown in France [249]. These regional differences highlight the importance to study local epidemics and suggest even contiguous areas may have different TDR prevalence values.

In low and middle income countries, the WHO has created a strategy for prevention and follow-up of TDR, which include an international expert committee board for training, technical assistance, standardized protocols, laboratory quality assurance, analysis of results, and recommendations for guidelines and public health actions [250]. Consequently, TDR studies with the WHO recommendations have been set up in Africa, Asia and Latin America. TDR was still under 5% in most of the African countries according to a systematic review that included articles published until 2008 [251]. Another meta-analysis using WHO data and studies up to July 2011 showed regional differences in Africa, for instance TDR increased at 29% rate per year in East Africa and 14% in Southern Africa whereas a non-significant increase was found in West and Central Africa [252]. Recent reports suggested that trends of TDR still are increasing based on new surveys and mathematical models [253]. In contrast, Asia seems to have the prevalence of TDR under 5% [244, 252], although still some countries lack surveys up to now [254]. Finally, the overall TDR in Latin America was stable and around 7% in studies with sampling dates up to 2009 [244, 252]. However, regional differences have been observed (see *chapter 4*), which suggest different TDR trends between countries that are important to evaluate.

Concerning the two countries analyzed in this thesis, Belgian cohorts have been included in pan-European studies such as SPREAD, or EUROSIDA. Therefore, the drug resistance patterns in treated populations are expected to be similar to other Western European countries as mentioned above. The

prevalence of TDR was 9.5% between 2003 and 2006 [78], and the trends seemed stable according to the pan-European studies including Belgian data [239, 241]. On the other hand, limited information is available for Colombia. The prevalence of acquired resistance was 82.1% between 2000-2007, and resistance for RT inhibitors was increasing at that time [255]. The prevalence of TDR was 5.8% in a nationwide study in 2006 [256] (see *chapter 4*), and 6.6% in the city of Cali between 2008 and 2010 [257]. A detail discussion of these studies is going to be presented in *chapter 6*.

IMPACT OF RESISTANCE

Resistance could impact the future options of ART, the CD4 count, mortality and viral fitness amongst other factors. Regarding primary resistance, patients had three times higher risk of virological failure within one year when the virus carried TDR and was resistant to at least one prescribed drug [258]. However, if the patient received a fully active ART regimen, the risk of virological failure was not different from patients without TDR, although it was increased slightly for NNRTI based regimens. Consequently, poor virological response due to TDR may influence the decline of CD4 counts [258, 259].

The impact of TDR on the clinical prognosis highlights the importance to perform routinely resistance testing in order to select three active antiretroviral drugs. However, basing individualized treatment on drug resistance testing is a strategy of high-income countries, whereas low and middle income countries use resistance surveillance studies to design treatment guidelines as a population public health approach [260]. The WHO guideline for the latter countries recommends standard first and second line therapies based on price, availability, potency of the drugs, the most likely profile of secondary resistance, and individualized monitoring of viral load [157, 261]. For instance, this guideline proposed ART sequencing strategies such as TDF+NNRTI based-regimens given that resistance to TDF does not impact ZDV activity, then ZDV+PI can be used as second-line regimen without the performance of resistance test [157, 261].

Regarding acquired resistance, high-income countries reported a decrease in resistance that furthermore could be easily treated with new antiretroviral drugs [230]. In contrast, some low and middle-income countries could face limited alternative therapeutic options together with limited resources for the follow-up of patients. The extent of HIV suppression ranged between 70% to 90% after one year of first-line treatment [262-264], which resulted in mainly NRTI and NNRTI resistance. However, it is predicted that ZDV would be active for second line therapy whereas the activity for NVP or EFV would be low (less than 30% of the patients) [263]. In a meta-analysis of 19 studies, second-line

virological failure was reported in 23% of the patients after one year, although failure may have been related with lack of adherence rather than resistance given the low frequency of mutations [265].

Finally, viral fitness is the ability of the virus to infect, replicate and produce infectious progeny in a given host environment [266]. In general, a resistant virus reduces the transmission efficiency by approximately 20% [267], but this effect depends on the mutation or the combination of mutations. For instance the M184IV reduces the viral fitness, whereas NNRTI mutations do not [268]. Under ART, the viral fitness could also be increased by accumulation of compensatory mutations [269], whereas these compensatory mutations especially in PR contribute to higher viral load and decline of CD4 counts in untreated individuals [270].

IMPACT OF GENETIC DIVERSITY

The time to develop drug resistance and the drug resistance pathways can vary between subtypes. Group O and HIV-2 display high-level of resistance to NNRTI and T20 [271-273]. Additionally, natural polymorphisms are frequently present in non-B subtypes. Although they can act as secondary mutations, they should not be considered as evidence for TDR [212] (see section of resistance). Natural polymorphisms are often the cause of discrepancies between algorithms as they can trigger certain rules in drug resistance interpretation systems [209, 274]. A global collaboration determined that 55 mutations related with subtype B drug-resistance were found in at least one non-B subtype and 80% of them were significantly associated with ART. Conversely, of 67 non-B resistance mutations only 61 were also associated with treatment in subtype B [275]. One example are the studies of single doses NVP for prevention of MTCT, that suggest more frequent resistance in patients with subtype C compared with subtype A [276, 277]. Some examples of mutations or polymorphism are shown in the table 1. 3.

Drug class	Subtype	Mutation	Observations
NRTIs	C	K65R	High frequency under ddI+d4T, it has been associated with Y181C
		T69N, V75I, V118I, L210N, T215S	
	F	L210W, Q151M	Low prevalence
	CRF06_cpx	TAMs	High propensity vs CRF02_AG
	HIV-2	L210N, T215S, K219N	Develops resistance faster than HIV-1
NNRTIs	A	Y181C	Less frequent than B
	C, CRF01_AE	V106M	Under NNRTI background
	C	G190A	Natural polymorphism
	G	A98S	Secondary resistance
	HIV-2	Y181I, Y188L, G190A, K101A, V106I, V179I	Natural cross-resistance NNRTI
	O	Y181C, A98S, K103R, V179E	Natural cross-resistance NNRTI
PIs	A, CRF37_cpx	V11I	Lower genetic barrier to DRV
	A, G, CRF02_AG	I13V	2ndR, less frequent in B (13%)
	G, CRF01_AE	K20I	2ndR, less frequent in B, A and F
	Non-B	M36I	2ndR, less frequent in B (13%)
	C	D30N	Less frequent than B
	CRF01_AE, CRF02_AG		Not observed, frequent in B
	G	I82M/T/S	The PMP V82I leads it. Not in B
	C,F,G,	M89I/V	The PMP L89M leads it. No in B
	A,C, G	L90M	More frequent than B
	F		Rare
	C	I93L	Less frequent in B (33%)
	HIV-2	I47A	Develops resistance faster for LPV
		L10I/V, K20V, V32I, M36I, M46I, L63E/K, A71V. G73A, V77T, V81IL	Lower potency of PIs than HIV-1
EIs	O,HIV-2		Natural resistance
INIs	B	R263K	Resistance to DTG
	C	G118R	Resistance to RAL if L74M is present
	C,CRF01_AE, CRF02_AG	G140S, G140C	Higher genetic barrier to INIs
	CRF01_AE, CRF02_AG	V151I	Higher genetic barrier to INIs

Table 1.3: Mutations or polymorphisms (PMP) present in non-B subtypes. Secondary resistance (2ndR) means presence of the mutation in treated patients. The nomenclature of mutations consists of the amino acid in the wild-type virus (first letter with the amino acid standard abbreviation), the position

in the genome (the number) and the mutated amino acid (second letter). Adapted from [158, 209, 277-280].

The pathways also vary between subtypes. For example the development of resistance to NFV is related with mutation D30N in subtype B and with L90M in subtype G and C [158]. The N88D seems to favor the selection of D30N in subtype B [281]. The treatment response of non-B subtypes seems to be equal to B subtypes according to several studies developed in Europe [74, 122, 209, 281-283].

1.7 DYNAMICS OF HIV EPIDEMIC

Nowadays, with the renewed interest in TasP as an effective prevention strategy [169] and its implementation in certain countries [114, 161, 162], there is an increasing concern about the generalizability and transferability of this measurement in different settings [284]. Since sociodemographic, behavioral and cultural drivers can vary between regions, the study of local HIV epidemics becomes important to design prevention policies [285]. Therefore, it is necessary to determine and evaluate testing and diagnosis rates, the origin of infections (local, national and international), engagement and follow-up in care, adherence to ART and consequent suppression of viral load. Additionally, the characteristics of the population engaged in transmission (the ones who transmit and are at risk for HIV infection) should be explored. Then, modeling of the epidemic can be carried out with relevant information to evaluate whether or not TasP is cost-effective as a public health policy [176, 286].

The continuous evaluation and improvement of the HIV care cascade is one of the main steps to obtain a maximum and sustained benefit from TasP. The care cascade of care continuum includes five major stages: HIV diagnosis, getting linked to care, retention in care, getting ART, and suppression of viral load [287]. The study of the cascade aims to inform in which step it is necessary to improve health care management, and usually is surveyed by Public Entities. For instance, the state of the HIV engagement cascade in USA started a controversy about the test-and-treat strategy implemented in this country. CDC together with other studies estimated that around 25% of one million people living with HIV (PLHIV) were undetectable or adherent to ART [287, 288], which implies that TasP will have little impact if the access to health care is not improved. However, in France and UK around 50% of PLHIV

were adherent or undetectable in 2010 [289, 290], indicating a better engagement in Europe (see *chapter 6*).

Since HIV transmission occurs mainly via sexual contacts, the study of sexual transmission networks is relevant to characterize the populations that contribute to the spread of HIV. Social networks have their roots in sociology during the 1930s [291], but they can be limited due to methodological and ethical issues such as interviews, disclosure and partner notifications [292]. Therefore, phylogenetics have become important tools in evaluating transmission networks or clusters within large datasets of HIV sequences because it is known that they can reconstruct known transmission histories [34, 35, 37-43, 117, 292-295].

Recently, the interest in phylogenetic analyses like phylogeography (see section 1.4.1) and transmission clusters has increased given the priority to know whether TasP could be applied at community-level in different settings [284, 293]. In this context, phylogenetic analyses may provide information for modeling studies, prevention trials and observational studies that evaluate the impact of this prevention strategy [296]. Phylogeography informs about the virus mobility that can give some insights about the role of migration and travel in a specific population or subtype, as has been shown for subtype B in Europe [51, 83] (see section 1.4.3). The study of transmission clusters sheds light on the transmission patterns between heterosexuals, MSM or IVDU [41, 42]; the factors involved in transmission, with or without drug resistance [38, 245, 294]; the evolutionary rate within a period of time in a specific epidemic, which is useful to study outbreaks in particular risk groups [297], and in general on the relationships between biological, demographical and social determinants [293]. The description of TDR and TDR transmission networks is of particular importance because TDR can impact the efficacy of first-line regimens and future therapy options (see section 1.6.5) [293]. High frequencies of transmission clusters and TDR have been mainly described in MSM and IVDU [39, 245, 294]. These clusters usually include NRTI drug resistance mutations like TAMs [294, 298-300], although NNRTI and PI mutations have also been described [301-303]. Differences between countries have been observed, a recent study in Belgium showed for instance that 18 out of 33 patients (55%) with TDR were involved in transmission clusters and were mainly infected with subtype B [40], whereas in Northern Greece 17 out of 46 (37%) TDR patients were part of clusters, mainly of subtype A [304]. Differences within countries can also occur. In this respect, mainly TAMs and few NNRTI mutations such as G190A and V106A were observed in Gent [40], whereas a cluster with K103N was recently reported in Namur [303].

It is important to bear in mind that the definition of clusters is still a matter of debate. The statistical criteria could differ and often include high bootstrap values (>95%) and low genetic distances (0.010-0.045) when ML trees are inferred [35, 37-39, 305]. Lower values have been used if clusters are confirmed with Bayesian techniques, with bootstrap values >70-90% but posterior distribution of 1 in the MCC tree, and higher thresholds for genetic distance (≤ 0.06) [40, 43, 304]. To evaluate the impact of cluster definition, some studies have performed sensitivity analyses [36, 306]. The inclusion of controls from the local HIV epidemic, and the exclusion of convergent evolution are also important parameters in the evaluation of transmission clusters [307, 308]. It is important to point out that transmission cluster analyses provide some epidemiological insights on the HIV epidemic, but they do not prove transmission given the underlying assumptions of phylogenetic analyses. For conclusive evidence, it would be necessary to have all controls from the local HIV epidemic, which is not feasible. The use of two different genes and two close different samples could enforce the results, but not enough data is yet available in clinical databases for this approach [307]. In conclusion, phylogenetic analyses have some limitations but can be valuable tools in the proper evaluation of public health, therapeutic or behavioral interventions when complemented with epidemiological, clinical, and demographical data [293].

Finally, the in-depth characterization of local epidemics can assist in the improvement of health care and can predict whether or not particular prevention policies could be effective. Different mathematical models have already been developed for sub-Saharan Africa or United States, where they have shown the differential impact of PrEP on TDR and HIV transmission [309-311] in different settings, and of TasP [170, 312-316] on HIV transmission. However, the assumptions on the characteristics of the epidemic may influence greatly the results [314, 317].

1.8 RATIONALE AND OBJECTIVES

With the aforementioned evidence about the impact of ART on morbidity and mortality at personal level [9], and on HIV transmission at community level [169], and the subsequent implementation of test-and-treat strategies in some countries [114, 161, 162], debate continues about the generalizability and transferability of this policy to other settings. In an attempt to resolve whether or not TasP could be applied in our local epidemics, we focused our research on Belgium and Colombia, within the framework of a collaborative project between KU Leuven, Belgium and Universidad del Rosario, Colombia. These countries share the concern with other countries that new infections are still increasing [318, 319], in contrast to several other countries where a 50% reduction of new HIV infections have been observed [10]. However, Belgium and Colombia differ in standard of care and availability of relevant data. In Belgium, the notification system for HIV provides some rough information about the HIV care cascade [320] and a European study has already revealed that international mobility constitutes one of the main sources of HIV-1 subtype B infections in Belgium [83]. The level and patterns of TDR have been evaluated until 2006 [78], but no updated data is yet available and still limited information is yet available on the characteristics of the population who transmit the virus, necessary for a targeted approach to prevention policies. Although Colombia has a notification system, underreporting may hamper an accurate assessment of the HIV care cascade [319]. Additionally, only little information is yet available for the in-depth study of the molecular HIV epidemiology or the extent of drug resistance and potential clinical impact. Therefore, the general goal of our studies was to study these two HIV-1 epidemics with the aim to understand their dynamics, using socio-demographic, clinical and virological data.

One of the main types of analysis in this thesis was the identification of subtypes and recombinants. Given the large datasets involved in surveillance studies, the most reliable manual phylogenetic analyses are too cumbersome. Automated subtyping tools can quickly subtype one to thousands of sequences, but their reliability is variable. We felt the need for a reliable strategy that was based on the use of one or more of the automatic tools and as few manual phylogenetic analyses as possible, which could be used in our subsequent studies. As our laboratory developed the REGA subtyping tool [99] and an update was recently implemented, we first set out to evaluate the performance of the REGA version 3 compared to the previous version of REGA and six other subtyping tools. To have a large amount of data and non-B subtypes, a large Portuguese database was used together with publicly available sequences. We did not include Belgian or Colombian sequences because at this stage their

number was too small for our purpose. From these results we designed a testing algorithm, which was subsequently used for subtype classification throughout this thesis.

Because data on viral diversity and viral mobility was lacking for the Colombian epidemic, we set up a collaborative project to retrieve viral sequences from treated HIV-1 patients sampled between 2002 and 2007. Therapy-naive patients were not included because baseline resistance testing is not routinely performed [163]. We included seven sampling regions, the largest sample available to our knowledge in this country. The viral diversity was investigated as well as the geographically and temporally viral spread in the country. Our results provided some insights about how the subtype B epidemic in Colombia is related with the global epidemic and how the different regions interacted in its spread within the country. We only set out to study viral diversity and viral mobility since clinical data that could have brought valuable information for prevention policies were not available.

TDR impacts on the selection of therapies for first line treatment, prevention of MTCT, and the potential use of TasP [260]. Therefore, we wanted to evaluate any available information useful for the Colombian setting. We systematically reviewed all sources with relevant information on TDR in Colombia and extended our search to other Latin American countries [163, 321]. When we started this research, the trends of TDR were still unknown in Latin American and Caribbean countries, and the fear of an increase in TDR was stated at the Conference of Opportunistic Infections and Retroviruses (CROI) in 2011 (available at www.retroconference.org) [322-325]. In this systematic review, we additionally evaluated the specific resistance patterns and factors that could influence TDR between 1996 and 2011 by using published data and viral sequences in public databases. As a result, this review also allowed to identify gaps in information and to define the future research agenda in our collaborative project in Colombia.

For the Belgian epidemic, a lot of information was already available, with one main question remaining: how does TDR spread in the local epidemic. Since we did not have access to the entire Belgian epidemic, an initial step in studying transmission of TDR in Belgium was focused on the smaller epidemic of the Leuven area. We used data from the AIDS Reference Center and AIDS Reference Laboratory in UZ Leuven, situated in the province Flemish Brabant. This center collected clinical and socio-demographic information since 1997 and the cohort included a total of ~1,184 HIV-1 patients until December of 2012. The availability of socio-demographic information, clinical records, and dense sampling allowed us to have a better understanding of the local HIV epidemic dynamics, specifically for the spread of TDR. UZ Leuven took part in the Belgian SPREAD study that evaluated the TDR

prevalence between 2003 and 2006 [78]. This survey showed stable trends of the overall TDR and per antiretroviral group. It also identified that patients infected in Belgium and infected with subtype B were more likely to have TDR. As UZ Leuven contributed only a small subset of their patient population to SPREAD (23 out of 283 patients), and HIV care and access to ART has changed substantially since then, we wanted to investigate the evolution in TDR prevalence for the different drug classes and the association with other socio-demographic factors over a period of 15 years (1998-2012). Since the identification of populations at risk for TDR could be useful for developing prevention policies, we focused our analysis on the identification of socio-demographic characteristics associated with TDR and with the spread of TDR by using statistical and phylogenetic tools, such as transmission cluster analyses.

In summary, we aimed to characterize the HIV-1 epidemic of the Belgian and Colombian cohorts using an integrated approach that includes socio-demographic information, clinical data, and viral sequences, analyzed with statistical and phylogenetic approaches. Taking into account the rational described above, our specific objectives were:

- To evaluate the performance of automated subtyping tools used in the clinical settings or surveillance of the HIV-1 molecular epidemiology, in order to set up a subtyping strategy for the rest of the thesis
- To establish the molecular epidemiology of HIV-1 and how the virus spread geographically and temporally in Colombia between 2002 and 2007
- To describe the trends and prevalence of TDR in Colombia, Latin America and Caribbean countries; and the factors associated between 1999 and 2011
- To determine the temporal trend of TDR prevalence in Belgium, the factors associated with TDR, its transmission, and clinical impact between 1998-2012

Finally, we hoped that our efforts would be useful in designing prevention strategies as well as in improving HIV care programs for these specific local settings because the extrapolation of data from other regions might result in incorrect conclusions and ineffective interventions [285].

REFERENCES

[1] Gottlieb MS, Schroff R, Schanker HM, Weisman JD, Fan PT, Wolf RA, et al. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. The New England journal of medicine. 1981 Dec 10;305(24):1425-31.

[2] CDC. Kaposi's Sarcoma and Pneumocystis Pneumonia among Homosexual men-New York city and California. MMWR 1981:305-8.

[3] Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science (New York, NY. 1983 May 20;220(4599):868-71.

[4] Gallo RC, Salahuddin SZ, Popovic M, Shearer GM, Kaplan M, Haynes BF, et al. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science (New York, NY. 1984 May 4;224(4648):500-3.

[5] Fischl MA, Richman DD, Grieco MH, Gottlieb MS, Volberding PA, Laskin OL, et al. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. The New England journal of medicine. 1987 Jul 23;317(4):185-91.

[6] Larder BA, Darby G, Richman DD. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. Science (New York, NY. 1989 Mar 31;243(4899):1731-4.

[7] Larder BA, Kemp SD, Purifoy DJ. Infectious potential of human immunodeficiency virus type 1 reverse transcriptase mutants with altered inhibitor sensitivity. Proceedings of the National Academy of Sciences of the United States of America. 1989 Jul;86(13):4803-7.

[8] Concorde. MRC/ANRS randomised double-blind controlled trial of immediate and deferred zidovudine in symptom-free HIV infection. Concorde Coordinating Committee. Lancet. 1994 Apr 9;343(8902):871-81.

[9] Palella FJ, Jr., Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. The New England journal of medicine. 1998 Mar 26;338(13):853-60.

[10] UNAIDS-WHO. GLOBAL REPORT: UNAIDS report on the global AIDS epidemic. Geneva; 2013.

[11] King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ, eds. Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses (ICTV). San Diego: Elsevier Academic Press 2011.

[12] Freed EO. HIV-1 replication. Somatic cell and molecular genetics. 2001 Nov;26(1-6):13-33.

[13] Turner BG, Summers MF. Structural biology of HIV. Journal of molecular biology. 1999 Jan 8;285(1):1-32.

[14] Adamson CS, Freed EO. Novel approaches to inhibiting HIV-1 replication. Antiviral Res. 2010 Jan;85(1):119-41.

[15] Chun TW, Fauci AS. Latent reservoirs of HIV: obstacles to the eradication of virus. Proceedings of the National Academy of Sciences of the United States of America. 1999 Sep 28;96(20):10958-61.

[16] Lafeuillade A, Stevenson M. The search for a cure for persistent HIV reservoirs. AIDS reviews. 2011 Apr-Jun;13(2):63-6.

[17] Coffin JM, Hughes SH, Vamus HE, eds. Retroviruses. New York: Cold Spring Harbor Laboratory Press 1997.

[18] Tang H, Kuhen KL, Wong-Staal F. Lentivirus replication and regulation. Annual review of genetics. 1999;33:133-70.

[19] Rambaut A, Posada D, Crandall KA, Holmes EC. The causes and consequences of HIV evolution. Nat Rev Genet. 2004 Jan;5(1):52-61.

[20] Mansky LM, Temin HM. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. Journal of virology. 1995 Aug;69(8):5087-94.

[21] Jetzt AE, Yu H, Klarmann GJ, Ron Y, Preston BD, Dougherty JP. High rate of recombination throughout the human immunodeficiency virus type 1 genome. Journal of virology. 2000 Feb;74(3):1234-40.

[22] Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science (New York, NY. 1996 Mar 15;271(5255):1582-6.

[23] Taylor BS, Hammer SM. The challenge of HIV-1 subtype diversity. The New England journal of medicine. 2008 Oct 30;359(18):1965-6.

[24] Mohammadi P, Desfarges S, Bartha I, Joos B, Zangger N, Munoz M, et al. 24 hours in the life of HIV-1 in a T cell line. PLoS pathogens. 2013 Jan;9(1):e1003161.

[25] Wei X, Ghosh SK, Taylor ME, Johnson VA, Emini EA, Deutsch P, et al. Viral dynamics in human immunodeficiency virus type 1 infection. Nature. 1995 Jan 12;373(6510):117-22.

[26] Lemey P, Salemi M, Vandamme AM, eds. The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Second ed. Cambridge, United Kingdom: Cambridge University Press 2009 2010.

[27] Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, et al. HIV-1 nomenclature proposal. Science (New York, NY. 2000 Apr 7;288(5463):55-6.

[28] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research. 2004;32(5):1792-7.

[29] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic acids research. 1994 Nov 11;22(22):4673-80.

[30] Felsenstein J. Phylogenies from molecular sequences: inference and reliability. Annual review of genetics. 1988;22:521-65.

[31] Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Int J Organic Evol. 1985;39:783-91.

[32] Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Systematic biology. 2006 Aug;55(4):539-52.

[33] Posada D, Crandall KA. Selecting models of nucleotide substitution: an application to human immunodeficiency virus 1 (HIV-1). Molecular biology and evolution. 2001 Jun;18(6):897-906.

[34] Brenner BG, Roger M, Routy JP, Moisi D, Ntemgwa M, Matte C, et al. High rates of forward transmission events after acute/early HIV-1 infection. The Journal of infectious diseases. 2007 Apr 1;195(7):951-9.

[35] Brenner BG, Roger M, Stephens D, Moisi D, Hardy I, Weinberg J, et al. Transmission clustering drives the onward spread of the HIV epidemic among men who have sex with men in Quebec. The Journal of infectious diseases. 2011 Oct 1;204(7):1115-9.
[36] Frentz D, Wensing AM, Albert J, Paraskevis D, Abecasis AB, Hamouda O, et al. Limited cross-border infections in patients newly diagnosed with HIV in Europe. Retrovirology. 2013;10:36.
[37] Hue S, Clewley JP, Cane PA, Pillay D. HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. AIDS (London, England). 2004 Mar 26;18(5):719-28.
[38] Yerly S, von Wyl V, Ledergerber B, Boni J, Schupbach J, Burgisser P, et al. Transmission of HIV-1 drug resistance in Switzerland: a 10-year molecular epidemiology survey. AIDS (London, England). 2007 Oct 18;21(16):2223-9.
[39] Yerly S, Junier T, Gayet-Ageron A, Amari EB, von Wyl V, Gunthard HF, et al. The impact of transmission clusters on primary drug resistance in newly diagnosed HIV-1 infection. AIDS (London, England). 2009 Jul 17;23(11):1415-23.
[40] Chalmet K, Staelens D, Blot S, Dinakis S, Pelgrom J, Plum J, et al. Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections. BMC infectious diseases. 2010;10:262.
[41] Hughes GJ, Fearnhill E, Dunn D, Lycett SJ, Rambaut A, Leigh Brown AJ. Molecular phylodynamics of the heterosexual HIV epidemic in the United Kingdom. PLoS pathogens. 2009 Sep;5(9):e1000590.
[42] Leigh Brown AJ, Lycett SJ, Weinert L, Hughes GJ, Fearnhill E, Dunn DT. Transmission network parameters estimated from HIV sequences for a nationwide epidemic. The Journal of infectious diseases. 2011 Nov;204(9):1463-9.
[43] Yebra G, Holguin A, Pillay D, Hue S. Phylogenetic and demographic characterization of HIV-1 transmission in Madrid, Spain. Infect Genet Evol. 2013 Mar;14:232-9.
[44] Faria NR, Suchard MA, Abecasis A, Sousa JD, Ndembi N, Bonfim I, et al. Phylodynamics of the HIV-1 CRF02_AG clade in Cameroon. Infect Genet Evol. 2011 Mar;12(2):453-60.
[45] de Oliveira T, Pillay D, Gifford RJ. The HIV-1 subtype C epidemic in South America is linked to the United Kingdom. PLoS One. 2010;5(2):e9311.
[46] Bartolo I, Abecasis AB, Borrego P, Barroso H, McCutchan F, Gomes P, et al. Origin and epidemiological history of HIV-1 CRF14_BG. PLoS One. 2011;6(9):e24130.
[47] Delatorre E, Bello G. Phylodynamics of the HIV-1 epidemic in Cuba. PLoS One. 2013;8(9):e72448.
[48] Gilbert MT, Rambaut A, Wlasiuk G, Spira TJ, Pitchenik AE, Worobey M. The emergence of HIV/AIDS in the Americas and beyond. Proceedings of the National Academy of Sciences of the United States of America. 2007 Nov 20;104(47):18566-70.
[49] Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLoS computational biology. 2009 Sep;5(9):e1000520.
[50] Lemey P, Rambaut A, Welch JJ, Suchard MA. Phylogeography takes a relaxed random walk in continuous space and time. Molecular biology and evolution. 2010 Aug;27(8):1877-85.
[51] Faria NR, Suchard MA, Rambaut A, Lemey P. Toward a quantitative understanding of viral phylogeography. Curr Opin Virol. 2011;1(5):423-9.
[52] Holmes EC. Evolutionary history and phylogeography of human viruses. Annual review of microbiology. 2008;62:307-28.
[53] Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. Nat Rev Genet. 2009 Aug;10(8):540-50.
[54] LeBreton M, Yang O, Tamoufe U, Mpoudi-Ngole E, Torimiro JN, Djoko CF, et al. Exposure to wild primates among HIV-infected persons. Emerging infectious diseases. 2007 Oct;13(10):1579-82.
[55] Wolfe ND, Daszak P, Kilpatrick AM, Burke DS. Bushmeat hunting, deforestation, and prediction of zoonoses emergence. Emerging infectious diseases. 2005 Dec;11(12):1822-7.
[56] Chen Z, Telfier P, Gettie A, Reed P, Zhang L, Ho DD, et al. Genetic characterization of new West African simian immunodeficiency virus SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop. Journal of virology. 1996 Jun;70(6):3617-27.
[57] Van Heuverswyn F, Li Y, Neel C, Bailes E, Keele BF, Liu W, et al. Human immunodeficiency viruses: SIV infection in wild gorillas. Nature. 2006 Nov 9;444(7116):164.
[58] Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemee V, et al. A new human immunodeficiency virus derived from gorillas. Nature medicine. 2009 Aug;15(8):871-2.
[59] de Sousa JD, Muller V, Lemey P, Vandamme AM. High GUD incidence in the early 20 century created a particularly permissive time window for the origin and initial spread of epidemic HIV strains. PLoS One. 2010;5(4):e9936.
[60] Tebit DM, Arts EJ. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. The Lancet infectious diseases. 2011 Jan;11(1):45-56.
[61] Korber B, Muldoon M, Theiler J, Gao F, Gupta R, Lapedes A, et al. Timing the ancestor of the HIV-1 pandemic strains. Science (New York, NY. 2000 Jun 9;288(5472):1789-96.
[62] Salemi M, Strimmer K, Hall WW, Duffy M, Delaporte E, Mboup S, et al. Dating the common ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes using a new method to uncover clock-like molecular evolution. Faseb J. 2001 Feb;15(2):276-8.
[63] Wertheim JO, Worobey M. Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2. PLoS computational biology. 2009 May;5(5):e1000377.
[64] Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, et al. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature. 2008 Oct 2;455(7213):661-4.
[65] Damond F, Worobey M, Campa P, Farfara I, Colin G, Matheron S, et al. Identification of a highly divergent HIV type 2 and proposal for a change in HIV type 2 classification. AIDS research and human retroviruses. 2004 Jun;20(6):666-72.
[66] Faria NR, Hodges-Mameletzis I, Silva JC, Rodes B, Erasmus S, Paolucci S, et al. The phylogeographic footprint of colonial history in the global dispersal of HIV-2 group A. The Journal of general virology. 2012 Jan 4.

[67] Lemey P, Pybus OG, Wang B, Saksena NK, Salemi M, Vandamme AM. Tracing the origin and history of the HIV-2 epidemic. Proceedings of the National Academy of Sciences of the United States of America. 2003 May 27;100(11):6588-92.
[68] Bello G, Guimaraes ML, Morgado MG. Evolutionary history of HIV-1 subtype B and F infections in Brazil. AIDS (London, England). 2006 Mar 21;20(5):763-8.
[69] Junqueira DM, de Medeiros RM, Matte MC, Araujo LA, Chies JA, Ashton-Prolla P, et al. Reviewing the history of HIV-1: spread of subtype B in the Americas. PLoS One. 2011;6(11):e27489.
[70] Murillo W, Veras N, Prosperi M, de Rivera IL, Paz-Bailey G, Morales-Miranda S, et al. A Single Early Introduction of HIV-1 Subtype B into Central America Accounts for Most Current Cases. Journal of virology. 2013 Jul;87(13):7463-70.
[71] Vallari A, Holzmayer V, Harris B, Yamaguchi J, Ngansop C, Makamche F, et al. Confirmation of putative HIV-1 group P in Cameroon. Journal of virology. 2010 Feb;85(3):1403-7.
[72] Ayouba A, Mauclere P, Martin PM, Cunin P, Mfoupouendoun J, Njinku B, et al. HIV-1 group O infection in Cameroon, 1986 to 1998. Emerging infectious diseases. 2001 May-Jun;7(3):466-7.
[73] UNAIDS. AIDS epidemic update 2006. Geneva, Switzerland 2006.
[74] Hemelaar J. The origin and diversity of the HIV-1 pandemic. Trends in molecular medicine. 2012 Mar;18(3):182-92.
[75] Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS (London, England). 2011 Mar 13;25(5):679-89.
[76] Montano SM, Sanchez JL, Laguna-Torres A, Cuchi P, Avila MM, Weissenbacher M, et al. Prevalences, genotypes, and risk factors for HIV transmission in South America. Journal of acquired immune deficiency syndromes (1999). 2005 Sep 1;40(1):57-64.
[77] Sanchez GI, Bautista CT, Eyzaguirre L, Carrion G, Arias S, Sateren WB, et al. Molecular epidemiology of human immunodeficiency virus-infected individuals in Medellin, Colombia. The American journal of tropical medicine and hygiene. 2006 Apr;74(4):674-7.
[78] Vercauteren J, Derdelinckx I, Sasse A, Bogaert M, Ceunen H, De Roo A, et al. Prevalence and epidemiology of HIV type 1 drug resistance among newly diagnosed therapy-naive patients in Belgium from 2003 to 2006. AIDS research and human retroviruses. 2008 Mar;24(3):355-62.
[79] Villarreal JL, Gutierrez J, Palacio L, Penuela M, Hernandez R, Lemay G, et al. Characterization of HIV type 1 envelope sequence among viral isolates circulating in the northern region of Colombia, South America. AIDS research and human retroviruses. 2012 Dec;28(12):1779-83.
[80] Rousseau CM, Learn GH, Bhattacharya T, Nickle DC, Heckerman D, Chetty S, et al. Extensive intrasubtype recombination in South African human immunodeficiency virus type 1 subtype C infections. Journal of virology. 2007 May;81(9):4492-500.
[81] Aulicino PC, Holmes EC, Rocco C, Mangano A, Sen L. Extremely rapid spread of human immunodeficiency virus type 1 BF recombinants in Argentina. Journal of virology. 2007 Jan;81(1):427-9.
[82] Casado C, Urtasun I, Saragosti S, Chaix ML, de Rossi A, Cattelan AM, et al. Different distribution of HIV type 1 genetic variants in European patients with distinct risk practices. AIDS research and human retroviruses. 2000 Feb 10;16(3):299-304.
[83] Paraskevis D, Pybus O, Magiorkinis G, Hatzakis A, Wensing AM, van de Vijver DA, et al. Tracing the HIV-1 subtype B mobility in Europe: a phylogeographic approach. Retrovirology. 2009;6:49.
[84] Lukashov VV, Kuiken CL, Vlahov D, Coutinho RA, Goudsmit J. Evidence for HIV type 1 strains of U.S. intravenous drug users as founders of AIDS epidemic among intravenous drug users in northern Europe. AIDS research and human retroviruses. 1996 Aug 10;12(12):1179-83.
[85] Perrin L, Kaiser L, Yerly S. Travel and the spread of HIV-1 genetic variants. The Lancet infectious diseases. 2003 Jan;3(1):22-7.
[86] Fontella R, Soares MA, Schrago CG. On the origin of HIV-1 subtype C in South America. AIDS (London, England). 2008 Oct 1;22(15):2001-11.
[87] Lai A, Bozzi G, Franzetti M, Binda F, Simonetti FR, Micheli V, et al. Phylogenetic analysis provides evidence of interactions between Italian heterosexual and South American homosexual males as the main source of national HIV-1 subtype C epidemics. Journal of medical virology. 2014 May;86(5):729-36.
[88] Bello G, Aulicino PC, Ruchansky D, Guimaraes ML, Lopez-Galindez C, Casado C, et al. Phylodynamics of HIV-1 circulating recombinant forms 12_BF and 38_BF in Argentina and Uruguay. Retrovirology. 2010;7:22.
[89] Gifford RJ, de Oliveira T, Rambaut A, Pybus OG, Dunn D, Vandamme AM, et al. Phylogenetic surveillance of viral genetic diversity and the evolving molecular epidemiology of human immunodeficiency virus type 1. Journal of virology. 2007 Dec;81(23):13050-6.
[90] Snoeck J, Van Dooren S, Van Laethem K, Derdelinckx I, Van Wijngaerden E, De Clercq E, et al. Prevalence and origin of HIV-1 group M subtypes among patients attending a Belgian hospital in 1999. Virus research. 2002 Apr 23;85(1):95-107.
[91] Pasquier C, Millot N, Njouom R, Sandres K, Cazabat M, Puel J, et al. HIV-1 subtyping using phylogenetic analysis of pol gene sequences. J Virol Methods. 2001 May;94(1-2):45-54.
[92] Yahi N, Fantini J, Tourres C, Tivoli N, Koch N, Tamalet C. Use of drug resistance sequence data for the systematic detection of non-B human immunodeficiency virus type 1 (HIV-1) subtypes: how to create a sentinel site for monitoring the genetic diversity of HIV-1 at a country scale. The Journal of infectious diseases. 2001 May 1;183(9):1311-7.
[93] Rozanov M, Plikat U, Chappey C, Kochergin A, Tatusova T. A web-based genotyping resource for viral sequences. Nucleic acids research. 2004 Jul 1;32(Web Server issue):W654-9.
[94] Liu TF, Shafer RW. Web resources for HIV type 1 genotypic-resistance test interpretation. Clin Infect Dis. 2006 Jun 1;42(11):1608-18.
[95] Beerenwinkel N, Daumer M, Oette M, Korn K, Hoffmann D, Kaiser R, et al. Geno2pheno: Estimating phenotypic drug resistance from HIV-1 genotypes. Nucleic acids research. 2003 Jul 1;31(13):3850-5.
[96] Struck D, Perez-Bercoff D, Devaux C, Schmit JC. COMET: A Novel approach to HIV-1 subtype prediction. 8th European HIV Drug Resistance Workshop; 2010 17- 19 March; Sorrento, Italy; 2010.

[97] Myers RE, Gale CV, Harrison A, Takeuchi Y, Kellam P. A statistical model for HIV-1 sequence classification using the subtype analyser (STAR). Bioinformatics (Oxford, England). 2005 Sep 1;21(17):3535-40.
[98] Schultz AK, Zhang M, Bulla I, Leitner T, Korber B, Morgenstern B, et al. jpHMM: improving the reliability of recombination prediction in HIV-1. Nucleic acids research. 2009 Jul;37(Web Server issue):W647-51.
[99] de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, et al. An automated genotyping system for analysis of HIV-1 and other microbial sequences. Bioinformatics (Oxford, England). 2005 Oct 1;21(19):3797-800.
[100] Kosakovsky Pond SL, Posada D, Stawiski E, Chappey C, Poon AF, Hughes G, et al. An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. PLoS computational biology. 2009 Nov;5(11):e1000581.
[101] Patel P, Borkowf CB, Brooks JT, Lasry A, Lansky A, Mermin J. Estimating per-act HIV transmission risk: a systematic review. AIDS (London, England). 2014 Jun 19;28(10):1509-19.
[102] Pilcher CD, Tien HC, Eron JJ, Jr., Vernazza PL, Leu SY, Stewart PW, et al. Brief but efficient: acute HIV infection and the sexual transmission of HIV. The Journal of infectious diseases. 2004 May 15;189(10):1785-92.
[103] Sagar M. HIV-1 transmission biology: selection and characteristics of infecting viruses. The Journal of infectious diseases. 2010 Oct 15;202 Suppl 2:S289-96.
[104] Renjifo B, Gilbert P, Chaplin B, Msamanga G, Mwakagile D, Fawzi W, et al. Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. AIDS (London, England). 2004 Aug 20;18(12):1629-36.
[105] John-Stewart GC, Nduati RW, Rousseau CM, Mbori-Ngacha DA, Richardson BA, Rainwater S, et al. Subtype C Is associated with increased vaginal shedding of HIV-1. The Journal of infectious diseases. 2005 Aug 1;192(3):492-6.
[106] Kahle E, Campbell M, Lingappa J, Donnell D, Celum C, Ondondo R, et al. HIV-1 subtype C is not associated with higher risk of heterosexual HIV-1 transmission: a multinational study among HIV-1 serodiscordant couples. AIDS (London, England). 2014 Jan 14;28(2):235-43.
[107] Yang C, Li M, Newman RD, Shi YP, Ayisi J, van Eijk AM, et al. Genetic diversity of HIV-1 in western Kenya: subtype-specific differences in mother-to-child transmission. AIDS (London, England). 2003 Jul 25;17(11):1667-74.
[108] Kiwanuka N, Laeyendecker O, Quinn TC, Wawer MJ, Shepherd J, Robb M, et al. HIV-1 subtypes and differences in heterosexual HIV transmission among HIV-discordant couples in Rakai, Uganda. AIDS (London, England). 2009 Nov 27;23(18):2479-84.
[109] Hudgens MG, Longini IM, Jr., Vanichseni S, Hu DJ, Kitayaporn D, Mock PA, et al. Subtype-specific transmission probabilities for human immunodeficiency virus type 1 among injecting drug users in Bangkok, Thailand. American journal of epidemiology. 2002 Jan 15;155(2):159-68.
[110] Hecht FM, Busch MP, Rawal B, Webb M, Rosenberg E, Swanson M, et al. Use of laboratory tests and clinical symptoms for identification of primary HIV infection. AIDS (London, England). 2002 May 24;16(8):1119-29.
[111] Zetola NM, Pilcher CD. Diagnosis and management of acute HIV infection. Infectious disease clinics of North America. 2007 Mar;21(1):19-48, vii.
[112] Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature. 1995 Jan 12;373(6510):123-6.
[113] Eshleman SH, Khaki L, Laeyendecker O, Piwowar-Manning E, Johnson-Lewis L, Husnik M, et al. Detection of individuals with acute HIV-1 infection using the ARCHITECT HIV Ag/Ab Combo assay. Journal of acquired immune deficiency syndromes (1999). 2009 Sep 1;52(1):121-4.
[114] Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents. In: Department of Health and Human Services (DHHS), ed. 2013:1-142.
[115] Hollingsworth TD, Anderson RM, Fraser C. HIV-1 transmission, by stage of infection. The Journal of infectious diseases. 2008 Sep 1;198(5):687-93.
[116] Xiridou M, Geskus R, de Wit J, Coutinho R, Kretzschmar M. Primary HIV infection as source of HIV transmission within steady and casual partnerships among homosexual men. AIDS (London, England). 2004 Jun 18;18(9):1311-20.
[117] Lewis F, Hughes GJ, Rambaut A, Pozniak A, Leigh Brown AJ. Episodic sexual transmission of HIV revealed by molecular phylodynamics. PLoS medicine. 2008 Mar 18;5(3):e50.
[118] Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, Peddada L, et al. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS (London, England). 2003 Sep 5;17(13):1871-9.
[119] Price DA, Goulder PJ, Klenerman P, Sewell AK, Easterbrook PJ, Troop M, et al. Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. Proceedings of the National Academy of Sciences of the United States of America. 1997 Mar 4;94(5):1890-5.
[120] McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF. The immune response during acute HIV-1 infection: clues for vaccine development. Nat Rev Immunol. 2010 Jan;10(1):11-23.
[121] Garrett PE, Manak M, Busch MP. Fiebig Stage and Recency: What's the relationship? *HIV Diagnostics Conference*. Orlando, USA: Abstract No. 25 2010.
[122] Bártolo I, Taveira N. HIV-1 Diversity and Its Implications in Diagnosis, Transmission, Disease Progression, and Antiretroviral Therapy In: Caliskan M, ed. *Genetic Diversity in Microorganisms*: InTech 2011.
[123] Weber B. Human immunodeficiency virus (HIV) antigen-antibody combination assays: evaluation of HIV seroconversion sensitivity and subtype detection. Journal of clinical microbiology. 2002 Nov;40(11):4402-3; author reply 3-4.
[124] Ly TD, Ebel A, Faucher V, Fihman V, Laperche S. Could the new HIV combined p24 antigen and antibody assays replace p24 antigen specific assays? J Virol Methods. 2007 Jul;143(1):86-94.
[125] Aghokeng AF, Mpoudi-Ngole E, Dimodi H, Atem-Tambe A, Tongo M, Butel C, et al. Inaccurate diagnosis of HIV-1 group M and O is a key challenge for ongoing universal access to antiretroviral treatment and HIV prevention in Cameroon. PLoS One. 2009;4(11):e7702.

[126] Chaillet P, Tayler-Smith K, Zachariah R, Duclos N, Moctar D, Beelaert G, et al. Evaluation of four rapid tests for diagnosis and differentiation of HIV-1 and HIV-2 infections in Guinea-Conakry, West Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2010 Sep;104(9):571-6.
[127] Pavie J, Rachline A, Loze B, Niedbalski L, Delaugerre C, Laforgerie E, et al. Sensitivity of five rapid HIV tests on oral fluid or finger-stick whole blood: a real-time comparison in a healthcare setting. PLoS One. 2010;5(7):e11581.
[128] Laforgerie E, Boucher B, Ly TD, Maisoneuve L, Izopet J, Delaugerre C, et al. Sensitivity of 8 CE (European Community)-approved rapid disposable tests for anti-HIV antibody detection during and after seroconversion. J Virol Methods. 2010 Apr;165(1):105-7.
[129] Swanson P, Huang S, Abravaya K, de Mendoza C, Soriano V, Devare SG, et al. Evaluation of performance across the dynamic range of the Abbott RealTime HIV-1 assay as compared to VERSANT HIV-1 RNA 3.0 and AMPLICOR HIV-1 MONITOR v1.5 using serial dilutions of 39 group M and O viruses. J Virol Methods. 2007 Apr;141(1):49-57.
[130] Lang W, Perkins H, Anderson RE, Royce R, Jewell N, Winkelstein W, Jr. Patterns of T lymphocyte changes with human immunodeficiency virus infection: from seroconversion to the development of AIDS. Journal of acquired immune deficiency syndromes (1999). 1989;2(1):63-9.
[131] Centers for Disease Control and Prevention. Revised Surveillance Case Definitions for HIV Infection Among Adults, Adolescents, and Children Aged <18 Months and for HIV Infection and AIDS Among Children Aged 18 Months to <13 Years — United States, 2008. In: MMWR, ed. Atlanta 2008:12.
[132] Yarchoan R, Venzon DJ, Pluda JM, Lietzau J, Wyvill KM, Tsiatis AA, et al. CD4 count and the risk for death in patients infected with HIV receiving antiretroviral therapy. Annals of internal medicine. 1991 Aug 1;115(3):184-9.
[133] Migueles SA, Connors M. Long-term nonprogressive disease among untreated HIV-infected individuals: clinical implications of understanding immune control of HIV. Jama. 2010 Jul 14;304(2):194-201.
[134] Pereyra F, Palmer S, Miura T, Block BL, Wiegand A, Rothchild AC, et al. Persistent low-level viremia in HIV-1 elite controllers and relationship to immunologic parameters. The Journal of infectious diseases. 2009 Sep 15;200(6):984-90.
[135] Pant Pai N, Shivkumar S, Cajas JM. Does genetic diversity of HIV-1 non-B subtypes differentially impact disease progression in treatment-naive HIV-1-infected individuals? A systematic review of evidence: 1996-2010. Journal of acquired immune deficiency syndromes (1999). 2012 Apr 1;59(4):382-8.
[136] Baeten JM, Chohan B, Lavreys L, Chohan V, McClelland RS, Certain L, et al. HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. The Journal of infectious diseases. 2007 Apr 15;195(8):1177-80.
[137] Kiwanuka N, Laeyendecker O, Robb M, Kigozi G, Arroyo M, McCutchan F, et al. Effect of human immunodeficiency virus Type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. The Journal of infectious diseases. 2008 Mar 1;197(5):707-13.
[138] Vasan A, Renjifo B, Hertzmark E, Chaplin B, Msamanga G, Essex M, et al. Different rates of disease progression of HIV type 1 infection in Tanzania based on infecting subtype. Clin Infect Dis. 2006 Mar 15;42(6):843-52.
[139] Kaleebu P, Nankya IL, Yirrell DL, Shafer LA, Kyosiimire-Lugemwa J, Lule DB, et al. Relation between chemokine receptor use, disease stage, and HIV-1 subtypes A and D: results from a rural Ugandan cohort. Journal of acquired immune deficiency syndromes (1999). 2007 May 1;45(1):28-33.
[140] Ng OT, Laeyendecker O, Redd AD, Munshaw S, Grabowski MK, Paquet AC, et al. HIV type 1 polymerase gene polymorphisms are associated with phenotypic differences in replication capacity and disease progression. The Journal of infectious diseases. 2014 Jan 1;209(1):66-73.
[141] Easterbrook PJ, Smith M, Mullen J, O'Shea S, Chrystie I, de Ruiter A, et al. Impact of HIV-1 viral subtype on disease progression and response to antiretroviral therapy. Journal of the International AIDS Society. 2010;13:4.
[142] Amornkul PN, Karita E, Kamali A, Rida WN, Sanders EJ, Lakhi S, et al. Disease progression by infecting HIV-1 subtype in a seroconverter cohort in sub-Saharan Africa. AIDS (London, England). 2013 Nov 13;27(17):2775-86.
[143] Prince JL, Claiborne DT, Carlson JM, Schaefer M, Yu T, Lahki S, et al. Role of transmitted Gag CTL polymorphisms in defining replicative capacity and early HIV-1 pathogenesis. PLoS pathogens. 2012;8(11):e1003041.
[144] Li X, Xue Y, Zhou L, Lin Y, Yu X, Wang X, et al. Evidence That HIV-1 CRF01_AE Is Associated with Low CD4+T Cell Count and CXCR4 Co-Receptor Usage in Recently Infected Young Men Who Have Sex with Men (MSM) in Shanghai, China. PLoS One. 2014;9(2):e89462.
[145] Li Y, Han Y, Xie J, Gu L, Li W, Wang H, et al. CRF01_AE subtype is associated with X4 tropism and fast HIV progression in Chinese patients infected through sexual transmission. AIDS (London, England). 2014 Feb 20;28(4):521-30.
[146] Amornkul PN, Tansuphasawadikul S, Limpakarnjanarat K, Likanonsakul S, Young N, Eampokalap B, et al. Clinical disease associated with HIV-1 subtype B' and E infection among 2104 patients in Thailand. AIDS (London, England). 1999 Oct 1;13(14):1963-9.
[147] Touloumi G, Pantazis N, Pillay D, Paraskevis D, Chaix ML, Bucher HC, et al. Impact of HIV-1 subtype on CD4 count at HIV seroconversion, rate of decline, and viral load set point in European seroconverter cohorts. Clin Infect Dis. 2013 Mar;56(6):888-97.
[148] Rowland-Jones SL. Timeline: AIDS pathogenesis: what have two decades of HIV research taught us? Nat Rev Immunol. 2003 Apr;3(4):343-8.
[149] Clavel F, Hance AJ. HIV drug resistance. The New England journal of medicine. 2004 Mar 4;350(10):1023-35.
[150] Santos AF, Soares MA. HIV Genetic Diversity and Drug Resistance. Viruses. 2010 Feb;2(2):503-31.
[151] d'Arminio Monforte A, Lepri AC, Rezza G, Pezzotti P, Antinori A, Phillips AN, et al. Insights into the reasons for discontinuation of the first highly active antiretroviral therapy (HAART) regimen in a cohort of antiretroviral naive patients. I.CO.N.A. Study Group. Italian Cohort of Antiretroviral-Naive Patients. AIDS (London, England). 2000 Mar 31;14(5):499-507.
[152] Ickovics JR, Cameron A, Zackin R, Bassett R, Chesney M, Johnson VA, et al. Consequences and determinants of adherence to antiretroviral medication: results from Adult AIDS Clinical Trials Group protocol 370. Antiviral therapy. 2002 Sep;7(3):185-93.

[153] Paterson DL, Swindells S, Mohr J, Brester M, Vergis EN, Squier C, et al. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Annals of internal medicine. 2000 Jul 4;133(1):21-30.
[154] Bangsberg DR. Less than 95% adherence to nonnucleoside reverse-transcriptase inhibitor therapy can lead to viral suppression. Clin Infect Dis. 2006 Oct 1;43(7):939-41.
[155] Parienti JJ, Ragland K, Lucht F, de la Blanchardiere A, Dargere S, Yazdanpanah Y, et al. Average adherence to boosted protease inhibitor therapy, rather than the pattern of missed doses, as a predictor of HIV RNA replication. Clin Infect Dis. 2010 Apr 15;50(8):1192-7.
[156] European AIDS Clinical Society (EACS). European Guidelines for Treatment of HIV infected adults in Europe version 7.0. Guidelines. Available in: http://www.europeanaidsclinicalsociety.org/guidelines.asp; 2013 November 2011.
[157] WHO. Consolidated Guidelines on the use of Antiretroviral Drugs for Treating and Preventing HIV infection: Recommendations for a Public Health approach. Geneva, Switzerland: World Health Organization 2013.
[158] Wainberg MA, Zaharatos GJ, Brenner BG. Development of antiretroviral drug resistance. The New England journal of medicine. 2011 Aug 18;365(7):637-46.
[159] Mallal S, Phillips E, Carosi G, Molina JM, Workman C, Tomazic J, et al. HLA-B*5701 screening for hypersensitivity to abacavir. The New England journal of medicine. 2008 Feb 7;358(6):568-79.
[160] Thompson MA, Aberg JA, Hoy JF, Telenti A, Benson C, Cahn P, et al. Antiretroviral treatment of adult HIV infection: 2012 recommendations of the International Antiviral Society-USA panel. Jama. 2012 Jul 25;308(4):387-402.
[161] Ministère des Affaires Sociales de la Santé, Conseil National du Sida, Agence National de Recherches sur le Sida et les Hépatites Virales. Prise en Charge Médicale de Personnes Vivant avec le VIH: Recommandations du Groupe d'Experts Rapport 2013. Paris, France: La documentation Française 2013.
[162] Ministério da Saúde. Protocolo Clínico e Diretrizes Terapêuticas para Manejo da Infecção pelo HIV em Adultos. In: Departamento de DST AeHV, ed. Brasilia, Brazil: Ministério da Saúde 2013.
[163] Ministerio de la Protección Social, Asociación Colombiana de Infectología (ACIN), United Nations Population Fund (UNFPA). Documentos preliminares Guías de practica clínica para la atención integral del VIH/sida para adolescentes, adultos y niños. Republica de Colombia; 2012.
[164] Townsend CL, Cortina-Borja M, Peckham CS, de Ruiter A, Lyall H, Tookey PA. Low rates of mother-to-child transmission of HIV following effective pregnancy interventions in the United Kingdom and Ireland, 2000-2006. AIDS (London, England). 2008 May 11;22(8):973-81.
[165] Tubiana R, Le Chenadec J, Rouzioux C, Mandelbrot L, Hamrene K, Dollfus C, et al. Factors associated with mother-to-child transmission of HIV-1 despite a maternal viral load <500 copies/ml at delivery: a case-control study nested in the French perinatal cohort (EPF-ANRS CO1). Clin Infect Dis. 2010 Feb 15;50(4):585-96.
[166] Montaner JS, Lima VD, Barrios R, Yip B, Wood E, Kerr T, et al. Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. Lancet. 2010 Aug 14;376(9740):532-9.
[167] Porco TC, Martin JN, Page-Shafer KA, Cheng A, Charlebois E, Grant RM, et al. Decline in HIV infectivity following the introduction of highly active antiretroviral therapy. AIDS (London, England). 2004 Jan 2;18(1):81-8.
[168] Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, Wabwire-Mangen F, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. The New England journal of medicine. 2000 Mar 30;342(13):921-9.
[169] Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 infection with early antiretroviral therapy. The New England journal of medicine. 2011 Aug 11;365(6):493-505.
[170] Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet. 2009 Jan 3;373(9657):48-57.
[171] Jia Z, Mao Y, Zhang F, Ruan Y, Ma Y, Li J, et al. Antiretroviral therapy to prevent HIV transmission in serodiscordant couples in China (2003-11): a national observational cohort study. Lancet. 2012 Oct 5;382(9899):1195-203.
[172] Cardo DM, Culver DH, Ciesielski CA, Srivastava PU, Marcus R, Abiteboul D, et al. A case-control study of HIV seroconversion in health care workers after percutaneous exposure. Centers for Disease Control and Prevention Needlestick Surveillance Group. The New England journal of medicine. 1997 Nov 20;337(21):1485-90.
[173] Pinto LA, Landay AL, Berzofsky JA, Kessler HA, Shearer GM. Immune response to human immunodeficiency virus (HIV) in healthcare workers occupationally exposed to HIV-contaminated blood. The American journal of medicine. 1997 May 19;102(5B):21-4.
[174] Spira AI, Marx PA, Patterson BK, Mahoney J, Koup RA, Wolinsky SM, et al. Cellular targets of infection and route of viral dissemination after an intravaginal inoculation of simian immunodeficiency virus into rhesus macaques. The Journal of experimental medicine. 1996 Jan 1;183(1):215-25.
[175] Tsai CC, Emau P, Follis KE, Beck TW, Benveniste RE, Bischofberger N, et al. Effectiveness of postinoculation (R)-9-(2-phosphonylmethoxypropyl) adenine treatment for prevention of persistent simian immunodeficiency virus SIVmne infection depends critically on timing of initiation and duration of treatment. Journal of virology. 1998 May;72(5):4265-73.
[176] European Centre for Disease Prevention and Control. Evaluating HIV treatment as prevention in the European context. Stockholm: ECDC; 2012.
[177] Centers for Disease Control and Prevention. Antiretroviral Postexposure Prophylaxis after Sexual, Injection-Drug Use, or other Nonoccupational Exposure to HIV in the United States. Morbidity and Mortality Weekly Report 2005;54(RR-2):1-20.
[178] Kuhar DT, Henderson DK, Struble KA, Heneine W, Thomas V, Cheever LW, et al. Updated US Public Health Service guidelines for the management of occupational exposures to human immunodeficiency virus and recommendations for postexposure prophylaxis. Infect Control Hosp Epidemiol. 2013 Sep;34(9):875-92.
[179] Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. The New England journal of medicine. 2010 Dec 30;363(27):2587-99.

[180] Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. The New England journal of medicine. 2012 Aug 2;367(5):399-410.
[181] Mujugira A, Baeten JM, Donnell D, Ndase P, Mugo NR, Barnes L, et al. Characteristics of HIV-1 serodiscordant couples enrolled in a clinical trial of antiretroviral pre-exposure prophylaxis for HIV-1 prevention. PLoS One. 2011;6(10):e25828.
[182] Murnane PM, Celum C, Mugo N, Campbell JD, Donnell D, Bukusi E, et al. Efficacy of preexposure prophylaxis for HIV-1 prevention among high-risk heterosexuals: subgroup analyses from a randomized trial. AIDS (London, England). 2013 Aug 24;27(13):2155-60.
[183] Grant RM, Hecht FM, Warmerdam M, Liu L, Liegler T, Petropoulos CJ, et al. Time trends in primary HIV-1 drug resistance among recently infected persons. Jama. 2002 Jul 10;288(2):181-8.
[184] Thigpen MC, Kebaabetswe PM, Paxton LA, Smith DK, Rose CE, Segolodi TM, et al. Antiretroviral preexposure prophylaxis for heterosexual HIV transmission in Botswana. The New England journal of medicine. 2012 Aug 2;367(5):423-34.
[185] Van Damme L, Corneli A, Ahmed K, Agot K, Lombaard J, Kapiga S, et al. Preexposure prophylaxis for HIV infection among African women. The New England journal of medicine. 2012 Aug 2;367(5):411-22.
[186] Marazzo JM, Ramjee G, Nair G, Palanee T, Mkhize B, Nakabiito C, et al. Pre-exposure Prophylaxis for HIV in Women: Daily Oral Tenofovir, Oral Tenofovir/Emtricitabine or Vaginal Tenofovir Gel in the VOICE Study. Abstract MTN 003. 20th Conference on Retroviruses and Opportunistic Infections 2013; Atlanta, USA; 2013.
[187] Peterson L, Taylor D, Roddy R, Belai G, Phillips P, Nanda K, et al. Tenofovir disoproxil fumarate for prevention of HIV infection in women: a phase 2, double-blind, randomized, placebo-controlled trial. PLoS clinical trials. 2007;2(5):e27.
[188] Choopanya K, Martin M, Suntharasamai P, Sangkum U, Mock PA, Leethochawalit M, et al. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet. 2013 Jun 15;381(9883):2083-90.
[189] CDC, DHHS. Pre-exposure Prophylaxis for the Prevention of HIV Infection in the United States: Clinical Practice Guideline 2014. In: Department of Health and Human Services (DHHS), Centers for Disease Control and Prevention (CDC), eds. 2013:1-67.
[190] Molina JM, Pintado C, Gatey C, Ponscarme D, Charbonneau P, Loze B, et al. Challenges and opportunities for oral pre-exposure prophylaxis in the prevention of HIV infection: where are we in Europe? BMC medicine. 2013;11:186.
[191] Cottrell ML, Kashuba AD. Topical microbicides and HIV prevention in the female genital tract. Journal of clinical pharmacology. 2014 Jun;54(6):603-15.
[192] Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, et al. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science (New York, NY. 2010 Sep 3;329(5996):1168-74.
[193] Romano J, Variano B, Coplan P, Van Roey J, Douville K, Rosenberg Z, et al. Safety and availability of dapivirine (TMC120) delivered from an intravaginal ring. AIDS research and human retroviruses. 2009 May;25(5):483-8.
[194] McGowan I. The development of rectal microbicides for HIV prevention. Expert opinion on drug delivery. 2014 Jan;11(1):69-82.
[195] Katlama C, Deeks SG, Autran B, Martinez-Picado J, van Lunzen J, Rouzioux C, et al. Barriers to a cure for HIV: new ways to target and eradicate HIV-1 reservoirs. Lancet. 2013 Jun 15;381(9883):2109-17.
[196] Schiffner T, Sattentau QJ, Dorrell L. Development of prophylactic vaccines against HIV-1. Retrovirology. 2013;10:72.
[197] Hutter G, Nowak D, Mossner M, Ganepola S, Mussig A, Allers K, et al. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. The New England journal of medicine. 2009 Feb 12;360(7):692-8.
[198] Allers K, Hutter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, et al. Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. Blood. 2011 Mar 10;117(10):2791-9.
[199] Henrich TJ, Hanhauser E, Sirignano M, Davis B, Lee T-H, Keating SM, et al. Long-term reduction in peripheral blood HIV-1 reservoirs following reduced-intensity conditioning allogeneic stem cell transplantation in two HIV-positive individuals. Abstract THAA0101. XIX International AIDS Conference; 2012; Washington, DC, USA; 2012.
[200] Henrich TJ, Hanhauser E, Sirignano M, Davis B, Lee T-H, Keating SM, et al. In depth investigation of peripheral and gut HIV-1 reservoirs, HIV-specific cellular immunity, and host microchimerism following allogeneic hematopoetic stem cell transplantation. Abstract WeLBA05. 7th International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention; 2013; Kuala Lumpur, Malaysia; 2013.
[201] Henrich TJ, Hu Z, Li JZ, Sciaranghella G, Busch MP, Keating SM, et al. Long-term reduction in peripheral blood HIV type 1 reservoirs following reduced-intensity conditioning allogeneic stem cell transplantation. The Journal of infectious diseases. 2013 Jun 1;207(11):1694-702.
[202] Henrich TJ, Hanhauser E, Sirignano M, Li JF, Marty F, Armand P, et al. HIV-1 Rebound Following Allogeneic Stem Cell Transplantation and Treatment Interruption. Abstract 144LB. 21st Conference on Retroviruses and Opportunistic Infections; 2014; Boston, USA; 2014.
[203] Persaud D, Gay H, Ziemniak C, Chen YH, Piatak M, Chun TW, et al. Functional HIV Cure after Very Early ART of an Infected Infant. Abstract 48LB. 20th Conference on Retroviruses and Opportunistic Infections; 2013; Atlanta, USA; 2013.
[204] Persaud D, Deveikis A, Gay H, Batra J, Chen T, Michalik DE, et al. Very Early Combination Antiretroviral Therapy in Perinatal HIV Infection: Two Case Studies. Abstract 75LB. 21st Conference on Retroviruses and Opportunistic Infections 2014; Boston, USA; 2014.
[205] Hocqueloux L, Prazuck T, Avettand-Fenoel V, Lafeuillade A, Cardon B, Viard JP, et al. Long-term immunovirologic control following antiretroviral therapy interruption in patients treated at the time of primary HIV-1 infection. AIDS (London, England). 2010 Jun 19;24(10):1598-601.
[206] Saez-Cirion A, Bacchus C, Hocqueloux L, Avettand-Fenoel V, Girault I, Lecuroux C, et al. Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study. PLoS pathogens. 2013 Mar;9(3):e1003211.
[207] Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, Walker BD, et al. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science (New York, NY. 2010 Dec 10;330(6010):1551-7.

[208] Deeks SG, Autran B, Berkhout B, Benkirane M, Cairns S, Chomont N, et al. Towards an HIV cure: a global scientific strategy. Nat Rev Immunol. 2012 Aug;12(8):607-14.
[209] Vandamme AM, Camacho RJ, Ceccherini-Silberstein F, de Luca A, Palmisano L, Paraskevis D, et al. European recommendations for the clinical use of HIV drug resistance testing: 2011 update. AIDS reviews. 2011 Apr-Jun;13(2):77-108.
[210] Vercauteren J, Vandamme AM. Algorithms for the interpretation of HIV-1 genotypic drug resistance information. Antiviral Res. 2006 Sep;71(2-3):335-42.
[211] Vandekerckhove LP, Wensing AM, Kaiser R, Brun-Vezinet F, Clotet B, De Luca A, et al. European guidelines on the clinical management of HIV-1 tropism testing. The Lancet infectious diseases. 2011 May;11(5):394-407.
[212] Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS One. 2009;4(3):e4724.
[213] Van Laethem K, De Luca A, Antinori A, Cingolani A, Perna CF, Vandamme AM. A genotypic drug resistance interpretation algorithm that significantly predicts therapy response in HIV-1-infected patients. Antiviral therapy. 2002 Jun;7(2):123-9.
[214] Vercauteren J, Beheydt G, Prosperi M, Libin P, Imbrechts S, Camacho R, et al. Clinical evaluation of Rega 8: an updated genotypic interpretation system that significantly predicts HIV-therapy response. PLoS One. 2013;8(4):e61436.
[215] Frentz D, Boucher CA, Assel M, De Luca A, Fabbiani M, Incardona F, et al. Comparison of HIV-1 genotypic resistance test interpretation systems in predicting virological outcomes over time. PLoS One. 2010;5(7):e11505.
[216] van de Vijver DA, Wensing AM, Asjo B, Bruckova M, Jorgensen LB, Camacho R, et al. HIV-1 drug-resistance patterns among patients on failing treatment in a large number of European countries. Acta dermatovenerologica Alpina, Panonica, et Adriatica. 2010 Dec;19(4):3-9.
[217] Vella S, Palmisano L. The global status of resistance to antiretroviral drugs. Clin Infect Dis. 2005 Aug 15;41 Suppl 4:S239-46.
[218] Little SJ, Frost SD, Wong JK, Smith DM, Pond SL, Ignacio CC, et al. Persistence of transmitted drug resistance among subjects with primary human immunodeficiency virus infection. Journal of virology. 2008 Jun;82(11):5510-8.
[219] Barbour JD, Hecht FM, Wrin T, Liegler TJ, Ramstead CA, Busch MP, et al. Persistence of primary drug resistance among recently HIV-1 infected adults. AIDS (London, England). 2004 Aug 20;18(12):1683-9.
[220] Jain V, Sucupira MC, Bacchetti P, Hartogensis W, Diaz RS, Kallas EG, et al. Differential persistence of transmitted HIV-1 drug resistance mutation classes. The Journal of infectious diseases. 2011 Apr 15;203(8):1174-81.
[221] Pao D, Andrady U, Clarke J, Dean G, Drake S, Fisher M, et al. Long-term persistence of primary genotypic resistance after HIV-1 seroconversion. Journal of acquired immune deficiency syndromes (1999). 2004 Dec 15;37(5):1570-3.
[222] Masquelier B, Bhaskaran K, Pillay D, Gifford R, Balestre E, Jorgensen LB, et al. Prevalence of transmitted HIV-1 drug resistance and the role of resistance algorithms: data from seroconverters in the CASCADE collaboration from 1987 to 2003. Journal of acquired immune deficiency syndromes (1999). 2005 Dec 15;40(5):505-11.
[223] Booth CL, Geretti AM. Prevalence and determinants of transmitted antiretroviral drug resistance in HIV-1 infection. J Antimicrob Chemother. 2007 Jun;59(6):1047-56.
[224] Little SJ, Smith DM. HIV treatment decisions and transmitted drug resistance. Clin Infect Dis. 2005 Jul 15;41(2):233-5.
[225] Richman DD, Morton SC, Wrin T, Hellmann N, Berry S, Shapiro MF, et al. The prevalence of antiretroviral drug resistance in the United States. AIDS (London, England). 2004 Jul 2;18(10):1393-401.
[226] Deeks SG, Gange SJ, Kitahata MM, Saag MS, Justice AC, Hogg RS, et al. Trends in multidrug treatment failure and subsequent mortality among antiretroviral therapy-experienced patients with HIV infection in North America. Clin Infect Dis. 2009 Nov 15;49(10):1582-90.
[227] von Wyl V, Yerly S, Burgisser P, Klimkait T, Battegay M, Bernasconi E, et al. Long-term trends of HIV type 1 drug resistance prevalence among antiretroviral treatment-experienced patients in Switzerland. Clin Infect Dis. 2009 Apr 1;48(7):979-87.
[228] Assoumou L, Descamps D, Yerly S, Dos Santos G, Marcelin AG, Delaugerre C, et al. Prevalence of HIV-1 drug resistance in treated patients with viral load >50 copies/mL in 2009: a French nationwide study. J Antimicrob Chemother. 2013 Jun;68(6):1400-5.
[229] Gill VS, Lima VD, Zhang W, Wynhoven B, Yip B, Hogg RS, et al. Improved virological outcomes in British Columbia concomitant with decreasing incidence of HIV type 1 drug resistance detection. Clin Infect Dis. 2010 Jan 1;50(1):98-105.
[230] De Luca A, Dunn D, Zazzi M, Camacho R, Torti C, Fanti I, et al. Declining prevalence of HIV-1 drug resistance in antiretroviral treatment-exposed individuals in Western Europe. The Journal of infectious diseases. 2013 Apr 15;207(8):1216-20.
[231] Little SJ, Daar ES, D'Aquila RT, Keiser PH, Connick E, Whitcomb JM, et al. Reduced antiretroviral drug susceptibility among patients with primary HIV infection. Jama. 1999 Sep 22-29;282(12):1142-9.
[232] Weinstock H, Respess R, Heneine W, Petropoulos CJ, Hellmann NS, Luo CC, et al. Prevalence of mutations associated with reduced antiretroviral drug susceptibility among human immunodeficiency virus type 1 seroconverters in the United States, 1993-1998. The Journal of infectious diseases. 2000 Jul;182(1):330-3.
[233] Perez-Alvarez L, Villahermosa ML, Cuevas MT, Delgado E, Manjon N, Vazquez de Parga E, et al. Single- and multidrug resistance mutations to reverse transcriptase and protease inhibitors: human immunodeficiency virus type 1-infected patients from two geographical areas in Spain. Spanish Groups for Antiretroviral Resistance Studies. Journal of human virology. 2000 May-Jun;3(3):150-6.
[234] Descamps D, Calvez V, Izopet J, Buffet-Janvresse C, Schmuck A, Colson P, et al. Prevalence of resistance mutations in antiretroviral-naive chronically HIV-infected patients in 1998: a French nationwide study. AIDS (London, England). 2001 Sep 28;15(14):1777-82.
[235] Shet A, Berry L, Mohri H, Mehandru S, Chung C, Kim A, et al. Tracking the prevalence of transmitted antiretroviral drug-resistant HIV-1: a decade of experience. Journal of acquired immune deficiency syndromes (1999). 2006 Apr 1;41(4):439-46.
[236] Van Vaerenbergh K, Debaisieux L, De Cabooter N, Declercq C, Desmet K, Fransen K, et al. Prevalence of genotypic resistance among antiretroviral drug-naive HIV-1-infected patients in Belgium. Antiviral therapy. 2001 Mar;6(1):63-70.
[237] Jayaraman GC, Archibald CP, Kim J, Rekart ML, Singh AE, Harmen S, et al. A population-based approach to determine the prevalence of transmitted drug-resistant HIV among recent versus established HIV infections: results from the Canadian HIV

strain and drug resistance surveillance program. Journal of acquired immune deficiency syndromes (1999). 2006 May;42(1):86-90.
[238] Weinstock HS, Zaidi I, Heneine W, Bennett D, Garcia-Lerma JG, Douglas JM, Jr., et al. The epidemiology of antiretroviral drug resistance among drug-naive HIV-1-infected persons in 10 US cities. The Journal of infectious diseases. 2004 Jun 15;189(12):2174-80.
[239] Vercauteren J, Wensing AM, van de Vijver DA, Albert J, Balotta C, Hamouda O, et al. Transmission of drug-resistant HIV-1 is stabilizing in Europe. The Journal of infectious diseases. 2009 Nov 15;200(10):1503-8.
[240] SPREAD. Transmission of drug-resistant HIV-1 in Europe remains limited to single classes. AIDS (London, England). 2008 Mar 12;22(5):625-35.
[241] Frentz D, Van de Vijver DA, Abecasis AB, Albert J, Hamouda O, Jorgensen LB, et al. The Epidemiology of HIV-1 Transmitted Drug Resistance: Increase in transmitted resistance to nonnucleoside reverse transcriptase inhibitors among newly diagnosed HIV-1 infections in Europe. Rotterdam, the Netherlands: Erasmus MC; 2012.
[242] Poon AF, Aldous JL, Mathews WC, Kitahata M, Kahn JS, Saag MS, et al. Transmitted drug resistance in the CFAR network of integrated clinical systems cohort: prevalence and effects on pre-therapy CD4 and viral load. PLoS One. 2011;6(6):e21189.
[243] Readhead AC, Gordon DE, Wang Z, Anderson BJ, Brousseau KS, Kouznetsova MA, et al. Transmitted antiretroviral drug resistance in New York State, 2006-2008: results from a new surveillance system. PLoS One. 2012;7(8):e40533.
[244] Frentz D, Boucher CA, van de Vijver DA. Temporal changes in the epidemiology of transmission of drug-resistant HIV-1 across the world. AIDS reviews. 2012 Jan-Mar;14(1):17-27.
[245] Karlsson A, Bjorkman P, Bratt G, Ekvall H, Gisslen M, Sonnerborg A, et al. Low prevalence of transmitted drug resistance in patients newly diagnosed with HIV-1 infection in Sweden 2003-2010. PLoS One. 2012;7(3):e33484.
[246] Bracciale L, Colafigli M, Zazzi M, Corsi P, Meraviglia P, Micheli V, et al. Prevalence of transmitted HIV-1 drug resistance in HIV-1-infected patients in Italy: evolution over 12 years and predictors. J Antimicrob Chemother. 2009 Sep;64(3):607-15.
[247] Dolling D, Sabin C, Delpech V, Smit E, Pozniak A, Asboe D, et al. Time trends in drug resistant HIV-1 infections in the United Kingdom up to 2009: multicentre observational study. BMJ (Clinical research ed. 2012;345:e5253.
[248] Bartmeyer B, Kuecherer C, Houareau C, Werning J, Keeren K, Somogyi S, et al. Prevalence of transmitted drug resistance and impact of transmitted resistance on treatment success in the German HIV-1 Seroconverter Cohort. PLoS One. 2010;5(10):e12718.
[249] Descamps D, Assoumou L, Chaix ML, Chaillon A, Pakianather S, de Rougemont A, et al. National sentinel surveillance of transmitted drug resistance in antiretroviral-naive chronically HIV-infected patients in France over a decade: 2001-2011. J Antimicrob Chemother. 2013 Jun 24;in press.
[250] Bennett DE, Bertagnolio S, Sutherland D, Gilks CF. The World Health Organization's global strategy for prevention and assessment of HIV drug resistance. Antiviral therapy. 2008;13 Suppl 2:1-13.
[251] Hamers RL, Derdelinckx I, van Vugt M, Stevens W, Rinke de Wit TF, Schuurman R. The status of HIV-1 resistance to antiretroviral drugs in sub-Saharan Africa. Antiviral therapy. 2008;13(5):625-39.
[252] Gupta RK, Jordan MR, Sultan BJ, Hill A, Davis DH, Gregson J, et al. Global trends in antiretroviral resistance in treatment-naive individuals with HIV after rollout of antiretroviral treatment in resource-limited settings: a global collaborative study and meta-regression analysis. Lancet. 2012 Oct 6;380(9849):1250-8.
[253] Hamers RL, Sigaloff KC, Kityo C, Mugyenyi P, de Wit TF. Emerging HIV-1 drug resistance after roll-out of antiretroviral therapy in sub-Saharan Africa. Curr Opin HIV AIDS. 2013 Jan;8(1):19-26.
[254] Sohn AH, Srikantiah P, Sungkanuparph S, Zhang F. Transmitted HIV drug resistance in Asia. Curr Opin HIV AIDS. 2013 Jan;8(1):27-33.
[255] Gómez SM, Olaya P, Diaz FJ. [Resistance to antiretrovirals in patients receiving HIV-AIDS therapy in Colombia]. Infectio. 2010;14(4):248-56.
[256] DiazGranados CA, Mantilla M, Lenis W. Antiretroviral drug resistance in HIV-infected patients in Colombia. Int J Infect Dis. 2010 Apr;14(4):e298-303.
[257] Galindo-Orrego P, Mueses-Marin HF, Galindo-Quintero J, Martinez-Cajas JL. [Transmitted HIV Resistance in ART-Naive Patients in Cali, Colombia]. Infectio. 2013;17(1):19-27.
[258] Wittkop L, Gunthard HF, de Wolf F, Dunn D, Cozzi-Lepri A, de Luca A, et al. Effect of transmitted drug resistance on virological and immunological response to initial combination antiretroviral therapy for HIV (EuroCoord-CHAIN joint project): a European multicohort study. The Lancet infectious diseases. 2011 May;11(5):363-71.
[259] Bezemer D, de Ronde A, Prins M, Porter K, Gifford R, Pillay D, et al. Evolution of transmitted HIV-1 with drug-resistance mutations in the absence of therapy: effects on CD4+ T-cell count and HIV-1 RNA load. Antiviral therapy. 2006;11(2):173-8.
[260] Bertagnolio S, Perno CF, Vella S, Pillay D. The Impact of HIV Drug Resistance on the Selection of First- and Second-Line ART in Resource-Limited Settings. The Journal of infectious diseases. 2013 Jun 15;207 Suppl 2:S45-8.
[261] De Luca A, Hamers RL, Schapiro JM. Antiretroviral treatment sequencing strategies to overcome HIV type 1 drug resistance in adolescents and adults in low-middle-income countries. The Journal of infectious diseases. 2013 Jun 15;207 Suppl 2:S63-9.
[262] Barth RE, van der Loeff MF, Schuurman R, Hoepelman AI, Wensing AM. Virological follow-up of adult patients in antiretroviral treatment programmes in sub-Saharan Africa: a systematic review. The Lancet infectious diseases. 2010 Mar;10(3):155-66.
[263] Hamers RL, Sigaloff KC, Wensing AM, Wallis CL, Kityo C, Siwale M, et al. Patterns of HIV-1 drug resistance after first-line antiretroviral therapy (ART) failure in 6 sub-Saharan African countries: implications for second-line ART strategies. Clin Infect Dis. 2012 Jun;54(11):1660-9.
[264] Hosseinipour MC, Gupta RK, Van Zyl G, Eron JJ, Nachega JB. Emergence of HIV drug resistance during first- and second-line antiretroviral therapy in resource-limited settings. The Journal of infectious diseases. 2013 Jun 15;207 Suppl 2:S49-56.
[265] Ajose O, Mookerjee S, Mills EJ, Boulle A, Ford N. Treatment outcomes of patients on second-line antiretroviral therapy in resource-limited settings: a systematic review and meta-analysis. AIDS (London, England). 2012 May 15;26(8):929-38.
[266] Domingo E, Holland JJ. RNA virus mutations and fitness for survival. Annual review of microbiology. 1997;51:151-78.

[267] Leigh Brown AJ, Frost SD, Mathews WC, Dawson K, Hellmann NS, Daar ES, et al. Transmission fitness of drug-resistant human immunodeficiency virus and the prevalence of resistance in the antiretroviral-treated population. The Journal of infectious diseases. 2003 Feb 15;187(4):683-6.
[268] Nicastri E, Sarmati L, d'Ettorre G, Palmisano L, Parisi SG, Uccella I, et al. Replication capacity, biological phenotype, and drug resistance of HIV strains isolated from patients failing antiretroviral therapy. Journal of medical virology. 2003 Jan;69(1):1-6.
[269] Leigh Brown AJ, Frost SD, Good B, Daar ES, Simon V, Markowitz M, et al. Genetic basis of hypersusceptibility to protease inhibitors and low replicative capacity of human immunodeficiency virus type 1 strains in primary infection. Journal of virology. 2004 Mar;78(5):2242-6.
[270] Theys K, Deforche K, Vercauteren J, Libin P, van de Vijver DA, Albert J, et al. Treatment-associated polymorphisms in protease are significantly associated with higher viral load and lower CD4 count in newly diagnosed drug-naive HIV-1 infected patients. Retrovirology. 2012;9:81.
[271] Descamps D, Collin G, Letourneur F, Apetrei C, Damond F, Loussert-Ajaka I, et al. Susceptibility of human immunodeficiency virus type 1 group O isolates to antiretroviral agents: in vitro phenotypic and genotypic analyses. Journal of virology. 1997 Nov;71(11):8893-8.
[272] Poveda E, Rodes B, Toro C, Soriano V. Are fusion inhibitors active against all HIV variants? AIDS research and human retroviruses. 2004 Mar;20(3):347-8.
[273] Smith RA, Anderson DJ, Pyrak CL, Preston BD, Gottlieb GS. Antiretroviral drug resistance in HIV-2: three amino acid changes are sufficient for classwide nucleoside analogue resistance. The Journal of infectious diseases. 2009 May 1;199(9):1323-6.
[274] Snoeck J, Kantor R, Shafer RW, Van Laethem K, Deforche K, Carvalho AP, et al. Discordances between interpretation algorithms for genotypic resistance to protease and reverse transcriptase inhibitors of human immunodeficiency virus are subtype dependent. Antimicrobial agents and chemotherapy. 2006 Feb;50(2):694-701.
[275] Kantor R, Katzenstein DA, Efron B, Carvalho AP, Wynhoven B, Cane P, et al. Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. PLoS medicine. 2005 Apr;2(4):e112.
[276] Flys TS, Chen S, Jones DC, Hoover DR, Church JD, Fiscus SA, et al. Quantitative analysis of HIV-1 variants with the K103N resistance mutation after single-dose nevirapine in women with HIV-1 subtypes A, C, and D. Journal of acquired immune deficiency syndromes (1999). 2006 Aug 15;42(5):610-3.
[277] Wainberg MA, Brenner BG. The Impact of HIV Genetic Polymorphisms and Subtype Differences on the Occurrence of Resistance to Antiretroviral Drugs. Molecular biology international. 2012;2012:256982.
[278] Martinez-Cajas JL, Pant-Pai N, Klein MB, Wainberg MA. Role of genetic diversity amongst HIV-1 non-B subtypes in drug resistance: a systematic review of virologic and biochemical evidence. AIDS reviews. 2008 Oct-Dec;10(4):212-23.
[279] Wainberg MA, Brenner BG. Role of HIV Subtype Diversity in the Development of Resistance to Antiviral Drugs. Viruses. 2010 Nov;2(11):2493-508.
[280] Martinez-Cajas JL, Pai NP, Klein MB, Wainberg MA. Differences in resistance mutations among HIV-1 non-subtype B infections: a systematic review of evidence (1996-2008). Journal of the International AIDS Society. 2009;12:11.
[281] Deforche K, Silander T, Camacho R, Grossman Z, Soares MA, Van Laethem K, et al. Analysis of HIV-1 pol sequences using Bayesian Networks: implications for drug resistance. Bioinformatics (Oxford, England). 2006 Dec 15;22(24):2975-9.
[282] Frater J. The impact of HIV-1 subtype on the clinical response on HAART. Journal of HIV therapy. 2002 Nov;7(4):92-6.
[283] Geretti AM, Harrison L, Green H, Sabin C, Hill T, Fearnhill E, et al. Effect of HIV-1 subtype on virologic and immunologic response to starting highly active antiretroviral therapy. Clin Infect Dis. 2009 May 1;48(9):1296-305.
[284] Brenner BG, Wainberg MA. Future of phylogeny in HIV prevention. Journal of acquired immune deficiency syndromes (1999). 2013 Jul;63 Suppl 2:S248-54.
[285] DeGruttola V, Smith DM, Little SJ, Miller V. Developing and evaluating comprehensive HIV infection control strategies: issues and challenges. Clin Infect Dis. 2010 May 15;50 Suppl 3:S102-7.
[286] Cambiano V, O'Connor J, Phillips AN, Rodger A, Lodwick R, Pharris A, et al. Antiretroviral therapy for prevention of HIV transmission: implications for Europe. Euro Surveill. 2013;18(48):20647.
[287] Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. Clin Infect Dis. 2011 Mar 15;52(6):793-800.
[288] CDC. Vital Signs: HIV Prevention Through Care and Treatment — United States. Morbidity and Mortality Weekly Report 2011;60(47):1618-23.
[289] Supervie V, Costagliola D. The Spectrum of Engagement in HIV Care in France: Strengths and Gaps. Abstract 1030. 20th Conference on Retroviruses and Opportunistic Infections 2013; Atlanta, USA; 2013.
[290] Delpech V, Yin Z, Kall M, Brown A. The United Kingdom's National Health Service provides excellent access to high quality HIV care: results from a national cohort. *IX International AIDS Conference: Abstract no MOPDC0301*. Washington, USA 2012.
[291] Doherty IA, Padian NS, Marlow C, Aral SO. Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections. The Journal of infectious diseases. 2005 Feb 1;191 Suppl 1:S42-54.
[292] Skar H, Hedskog C, Albert J. HIV-1 evolution in relation to molecular epidemiology and antiretroviral resistance. Annals of the New York Academy of Sciences. 2011 Aug;1230:108-18.
[293] Brenner B, Wainberg MA, Roger M. Phylogenetic inferences on HIV-1 transmission: implications for the design of prevention and treatment interventions. AIDS (London, England). 2013 Apr 24;27(7):1045-57.
[294] Brenner BG, Roger M, Moisi DD, Oliveira M, Hardy I, Turgel R, et al. Transmission networks of drug resistance acquired in primary/early stage HIV infection. AIDS (London, England). 2008 Nov 30;22(18):2509-15.
[295] Leitner T, Escanilla D, Franzen C, Uhlen M, Albert J. Accurate reconstruction of a known HIV-1 transmission history by phylogenetic tree analysis. Proceedings of the National Academy of Sciences of the United States of America. 1996 Oct 1;93(20):10864-9.
[296] Dennis AM. Phylogeny Studies in HIV Prevention Research. Imperial College, London: National Institutes of Health, National Institute of Allergy and Infectious Diseases, Division of AIDS, HIV Prevention Trials Network, Bill & Melinda Gates Foundation; 2012.

[297] Maljkovic Berry I, Ribeiro R, Kothari M, Athreya G, Daniels M, Lee HY, et al. Unequal evolutionary rates in the human immunodeficiency virus type 1 (HIV-1) pandemic: the evolutionary rate of HIV-1 slows down when the epidemic rate increases. Journal of virology. 2007 Oct;81(19):10625-35.
[298] Castor D, Low A, Evering T, Karmon S, Davis B, Figueroa A, et al. Transmitted drug resistance and phylogenetic relationships among acute and early HIV-1-infected individuals in New York City. Journal of acquired immune deficiency syndromes (1999). 2012 Sep 1;61(1):1-8.
[299] Castro E, Khonkarly M, Ciuffreda D, Burgisser P, Cavassini M, Yerly S, et al. HIV-1 drug resistance transmission networks in southwest Switzerland. AIDS research and human retroviruses. 2010 Nov;26(11):1233-8.
[300] Hue S, Gifford RJ, Dunn D, Fernhill E, Pillay D. Demonstration of sustained drug-resistant human immunodeficiency virus type 1 lineages circulating among treatment-naive individuals. Journal of virology. 2009 Mar;83(6):2645-54.
[301] Buskin SE, Ellis GM, Pepper GG, Frenkel LM, Pergam SA, Gottlieb GS, et al. Transmission cluster of multiclass highly drug-resistant HIV-1 among 9 men who have sex with men in Seattle/King County, WA, 2005-2007. Journal of acquired immune deficiency syndromes (1999). 2008 Oct 1;49(2):205-11.
[302] Dennis AM, Hue S, Hurt CB, Napravnik S, Sebastian J, Pillay D, et al. Phylogenetic insights into regional HIV transmission. AIDS (London, England). 2012 Sep 10;26(14):1813-22.
[303] Ruelle J, Ingels MG, Jnaoui K, Ausselet N, Vincent A, Sasse A, et al. Transmission Network of an HIV Type 1 Strain with K103N in Young Belgian Patients from Different Risk Groups. AIDS research and human retroviruses. 2013 Oct;29(10):1306-9.
[304] Skoura L, Metallidis S, Buckton AJ, Mbisa JL, Pilalas D, Papadimitriou E, et al. Molecular and epidemiological characterization of HIV-1 infection networks involving transmitted drug resistance mutations in Northern Greece. J Antimicrob Chemother. 2011 Dec;66(12):2831-7.
[305] Bezemer D, van Sighem A, Lukashov VV, van der Hoek L, Back N, Schuurman R, et al. Transmission networks of HIV-1 among men having sex with men in the Netherlands. AIDS (London, England). 2010 Jan 16;24(2):271-82.
[306] Winand R. Development of a phylogenetic framework for the detection of HIV-1 drug resistance transmission chains. Leuven, Belgium: KU Leuven; 2011.
[307] Abecasis AB, Geretti AM, Albert J, Power L, Weait M, Vandamme AM. Science in court: the myth of HIV fingerprinting. The Lancet infectious diseases. 2011 Feb;11(2):78-9.
[308] Pillay D, Rambaut A, Geretti AM, Brown AJ. HIV phylogenetics. BMJ (Clinical research ed. 2007 Sep 8;335(7618):460-1.
[309] Abbas UL, Glaubius R, Mubayi A, Hood G, Mellors JW. Antiretroviral therapy and pre-exposure prophylaxis: combined impact on HIV transmission and drug resistance in South Africa. The Journal of infectious diseases. 2013 Jul 15;208(2):224-34.
[310] Supervie V, Barrett M, Kahn JS, Musuka G, Moeti TL, Busang L, et al. Modeling dynamic interactions between pre-exposure prophylaxis interventions & treatment programs: predicting HIV transmission & resistance. Scientific reports. 2011;1:185.
[311] Supervie V, Garcia-Lerma JG, Heneine W, Blower S. HIV, transmitted drug resistance, and the paradox of preexposure prophylaxis. Proceedings of the National Academy of Sciences of the United States of America. 2010 Jul 6;107(27):12381-6.
[312] Baggaley RF, Garnett GP, Ferguson NM. Modelling the impact of antiretroviral use in resource-poor settings. PLoS medicine. 2006 Apr;3(4):e124.
[313] Dodd PJ, Garnett GP, Hallett TB. Examining the promise of HIV elimination by 'test and treat' in hyperendemic settings. AIDS (London, England). 2010 Mar 13;24(5):729-35.
[314] Eaton JW, Johnson LF, Salomon JA, Barnighausen T, Bendavid E, Bershteyn A, et al. HIV treatment as prevention: systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS medicine. 2012;9(7):e1001245.
[315] Lima VD, Johnston K, Hogg RS, Levy AR, Harrigan PR, Anema A, et al. Expanded access to highly active antiretroviral therapy: a potentially powerful strategy to curb the growth of the HIV epidemic. The Journal of infectious diseases. 2008 Jul 1;198(1):59-67.
[316] Velasco-Hernandez JX, Gershengorn HB, Blower SM. Could widespread use of combination antiretroviral therapy eradicate HIV epidemics? The Lancet infectious diseases. 2002 Aug;2(8):487-93.
[317] The HIV Molelling Consortium Treatment as Prevention Editorial Writing Group. HIV treatment as prevention: models, data, and questions--towards evidence-based decision-making. PLoS medicine. 2012;9(7):e1001259.
[318] ISP-WIV. Epidemiologie van AIDS en HIV infectie in België. Epidemiologie du SIDA et de l'infection à VIH en Belgique. In: Direction opérationnelle Santé publique et Surveillance, Service Maladies infectieuses au sein de la Population générale, Programme de surveillance VIH/SIDA/IST, eds. 2012.
[319] Ministerio de la Protección Social. Informe UNGASS: Seguimiento de la Declaracion de compromiso sobre el VIH/sida. Bogota, Colombia: Ministerio de la Proteccion Social; 2012.
[320] Ministry of Social Affairs and Health. Belgian National HIV/AIDS Plan 2014-2019. Belgium 2013.
[321] Ministerio de la Protección Social. Guía para el manejo de VIH/sida basada en la evidencia Colombia. In: Ministerio de la Protección Social, ed. Republica de Colombia: Available at: http://www.minproteccionsocial.gov.co/VBeContent/library/documents/DocNewsNo16914DocumentNo6250.pdf. 2006.
[322] Avila-Rios S, Garcia-Morales C, Garrido-Rodríguez D, Ormsby CE, Andrade-Villanueva J, Gonzales-Hernandez L, et al. National Prevalence and Trends of HIV Transmitted Drug Resistance in Mexico. *18th Conference on Retroviruses and Opportunistic Infections* Boston, United States 2011.
[323] De Moraes Soares CMP, Vergara TRC, Sucupira MC, Brites C, Brito JDU, Caseiro MM, et al. Prevalence of Transmitted HIV-1 Antiretroviral Resistance among Patients Initiating ART in Brazil: A Surveillance Using Dried Blood Spots. *18th Conference on Retroviruses and Opportunistic Infections*. Boston, United States 2011.
[324] Hernandez MT, Sanchez MA, Bart A, Ruiz JD, Magis C, Bravo-Garcia E, et al. Epidemiology of HIV and STI among Mexican Migrants in California 18th Conference on Retroviruses and Opportunistic Infections 2011; Boston, United States; 2011.
[325] Jordan MR, Kelley K, Saadani-Hassani A, Obeng-Aduasare Y, Bertagnolio S. Monitoring ART Clinic and Program Using World Health Organization HIV Drug Resistance Early Warning Indicators. 18th Conference on Retroviruses and Opportunistic Infections; 2011; Boston, United States; 2011.

CHAPTER 2

AUTOMATED SUBTYPING OF HIV-1 GENETIC SEQUENCES FOR CLINICAL AND SURVEILLANCE PURPOSES: PERFORMANCE EVALUATION OF THE NEW REGA VERSION 3 AND SEVEN OTHER TOOLS

Published in:

Andrea-Clemencia Pineda-Peña, Nuno Rodrigues Faria, Stijn Imbrechts, Pieter Libin, Ana Barroso Abecasis, Koen Deforche, Arley Gómez-López, Ricardo J. Camacho, Tulio de Oliveira, Anne-Mieke Vandamme. 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 2013; 19: 337–348

2.1 SUMMARY

Background: To investigate differences in pathogenesis, diagnosis and resistance pathways between HIV-1 subtypes, an accurate subtyping tool for large datasets is needed. We aimed to evaluate the performance of automated subtyping tools to classify the different subtypes and circulating recombinant forms using *pol*, the most sequenced region in clinical practice. We also present the upgraded version 3 of the REGAv3.

Methodology: HIV-1 *pol* sequences (PR+RT) for 4674 patients retrieved from the Portuguese HIV Drug Resistance Database, and 1872 *pol* sequences trimmed from full-length genomes retrieved from the Los Alamos database were classified with statistical-based tools such as COMET, jpHMM and STAR; similarity-based tools such as NCBI and Stanford; and phylogenetic-based tools such as REGA version 2 (REGAv2), REGAv3, and SCUEAL. The performance of these tools, for *pol*, and for PR and RT separately, was compared in terms of reproducibility, sensitivity and specificity with respect to the gold standard, which was manual phylogenetic analysis of the *pol* region.

Results: The sensitivity and specificity for subtypes B and C was more than 96% for seven tools, but was variable for other subtypes such as A, D, F and G. With regard to the most common CRFs, the sensitivity and specificity for CRF01_AE was ~99% with statistical-based tools, with phylogenetic-based tools and with Stanford, one of the similarity based tools. CRF02_AG was correctly identified for more than 96% by COMET, REGAv3, Stanford and STAR. All the tools reached a specificity of more than 97% for most of the subtypes and the two main CRFs (CRF01_AE and CRF02_AG). Other CRFs were identified only by COMET, REGAv2, REGAv3, and SCUEAL and with variable sensitivity. When analyzing sequences for PR and RT separately, the performance for PR was generally lower and variable between the tools. Similarity and statistical-based tools were 100% reproducible, but this was lower for phylogenetic-based tools such as REGA (~99%) and SCUEAL (~96%).

Conclusions: REGAv3 had an improved performance for subtype B and CRF02_AG compared to REGAv2 and is now able to also identify all epidemiologically relevant CRFs. In general the best performing tools, in alphabetical order, were COMET, jpHMM, REGAv3, and SCUEAL when analyzing pure subtypes in the *pol* region, and COMET and REGAv3 when analyzing most of the CRFs. Based on this study, we recommend to confirm subtyping with 2 well performing tools, and be cautious with the interpretation of short sequences.

2.2 INTRODUCTION

At the end of 2011 there were 34 million of people living with HIV [1]. Most infections are caused by HIV type 1 group Major (HIV-1 group M), which can be further classified into several clades based on genetic diversity. To date, nine distinct subtypes named A-D,F-H, J, K [2] and 58 CRFs (http://www.hiv.lanl.gov/; accessed March 2013) have been identified. While subtype B has been widely studied and is predominant in North America, Europe and Australia, it only causes approximately 10 percent of the infections globally [3], while subtype C is causing nearly half of global infections, followed by subtype A with 12% of global infections [3]. In addition, infections with recombinant forms such as CRFs and URFs have been increasing over the past decades and are now responsible for a total of 20% of the global infections. The distribution of infections caused by CRFs varies regionally; for example, while CRF02_AG (8% global infections) is mostly prevalent in West and Central Africa , CRF01_AE (5% global infections) is more frequent in South and East Asia [3]. Additionally, CRF06_cpx has been identified in West Africa and some European countries, BC recombinants such as CRF07 _BC are frequent in China, and BF recombinants, particularly CRF12_BF, prevail in Brazil and Argentina [3].

Due to the fast pace of evolution and frequent recombination of HIV-1 [4, 5], accurate subtyping of the growing arsenal of genetic data arising from epidemiological and antiretroviral resistance studies has become increasingly challenging. Importantly, different HIV-1 clades show differences in pathogenesis and present distinct resistance pathways, which in turn may lead to different clinical outcomes. For example, subtype D seems to be more transmissible and is associated with faster disease progression [6]. Moreover, subtypes A, C, F and G have some natural polymorphisms in PR and RT which contribute to resistance in subtype B [7-11]. However, it is often difficult to compare epidemiological and clinical impact studies since different subtyping methods are used and the classification of HIV-1 clades frequently seems to differ according to the method employed [12].

Although the gold standard for classification of HIV-1 is based on phylogenetic analysis of full-length genome sequences [2], this method is not widely used in clinical settings. Since most available data are derived from genotypic assays for resistance to PR and RT inhibitors and this region has proven to contain enough phylogenetic signal [13], manual phylogenetic analysis (*MPhy*) on *pol* region can be safely used to identify subtypes [14, 15]. However, for large datasets, automated tools are needed since manual subtyping is cumbersome. There are three main types of automated tools based on the method used to assign an HIV-1 clade to a query sequence. First, similarity-based tools include the

NCBI subtyping tool [16], Stanford [17], Geno2pheno [18] and EuResist (http://engine.euresist.org/data_analysis/viral_sequence/_new). Second, statistical-based tools use partial matching compression algorithms such as COntext-based Modeling for Expeditious Typing – COMET- [19], position-specific scoring matrices plus a statistical model such as STAR [20] or jumping profile Hidden Markov Models such as jpHMM [21]. Finally, there are phylogenetic-based tools such as REGA [22, 23] and SCUEAL [24].

A major objective of this paper was to compare the latest Rega subtyping tool with other available tools. The Rega subtyping tool has as philosophy to use phylogenetic analysis in order to take into account the epidemiological and evolutionary relationships among subtypes, such that it approaches the gold standard to classify subtypes [2]. The algorithms used in earlier versions of REGA have been previously described [22, 25]. REGAv2 had a high number of unassigned sequences, in part because of the limited number of CRFs included in the reference dataset [26], and the philosophy to achieve a high specificity at the cost of sensitivity. To overcome these limitations, the new REGAv3 uses an improved decision-tree algorithm geared towards increasing the recognition of pure subtypes and recombinants (see further details http://bioafrica.mrc.ac.za:8080/rega-genotype-3.0.2/hiv/typingtool/decisiontrees). The reference dataset has also been improved to include more divergent strains per subtype and to classify up to CRF47_BF (See further details http://bioafrica.mrc.ac.za:8080/rega-genotype-3.0.2/hiv/typingtool/method).

In this paper we aim to determine the performance of REGAv3 in the identification of HIV-1 clades, and to compare its sensitivity, specificity and reproducibility with its previous version REGAv2 and six other publicly available automated subtyping tools (COMET, jpHMM, NCBI, SCUEAL, Stanford and STAR). Another goal of this paper was to give guidance as to which HIV-1 subtyping tool would be better for use in a clinical and a surveillance context.

2.3 MATERIAL AND METHODS

2.3.1 STUDY POPULATION AND SUBTYPING TOOLS

With the objective of emphasizing the classification of prevalent non B subtypes, we used two datasets (see figure 2.1). The clinical dataset was retrieved from the Portuguese Resistance database and consisted of 4676 *pol* sequences obtained for routine resistance testing and pooled from 22 Portuguese hospitals, (mean length: 1295 nucleotides (nts); min: 993 nts, max: 1311 nts). Sequences were

obtained by population sequencing using the ViroSeq 2.0 toolkit (Abbott Laboratories, Abbott Park, IL, USA). (Sequences are available through Euresist http://www.euresist.org). The Los Alamos dataset herein named as LANL dataset, was retrieved using the following search criteria: "subtype" AND genomic region: "complete genome" AND "one sequence per patient" and we excluded CRFs that could not be shown to have epidemiological relevance and with less than 5 full length genome sequences at the time the analyses were initiated (CRF03_AB, CRF04_cpx, CRF05_DF, CRF08_BC, CRF09_cpx, CRF10_CD, CRF11_cpx, CRF13_cpx, and all CRFs later than CRF15_01B) [3]. As a result, the LANL dataset included 1872 *pol* sequences (1300 nts), that were trimmed from full genome sequences publicly available in Los Alamos database (http://www.hiv.lanl.gov/ ; Date of access: October 2011) (accession numbers are shown in the supplementary material number 2.1). In addition, each sequence of the LANL dataset was divided in PR (mean length: 300 nts) and RT (mean length: 1000 nts) with the objective of evaluating the differences in the performance for identifying PR and RT separately.

Only sequences that passed the quality control check were included: the quality of the sequences was evaluated by using the quality tool of Los Alamos database (available in http://www.hiv.lanl.gov/content/sequence/QC/index.html) and the parameters of Stanford database which are: a maximum number of four for PR and six for RT stop codons +no frame-shifts + no unpublished amino acid insertions or deletions + no highly ambiguous nucleotides (B,D,H,V,N) [27]. As a result, 2 sequences were rejected from the clinical dataset, and the final number of sequences was 4674.

Both datasets were analyzed by the following 8 subtyping tools: COMET version 2 (http://comet.retrovirology.lu), jpHMM (http://jphmm.gobics.de/submission_hiv.html), NCBI subtyping tools using the reference dataset from 2009 (http://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi), Stanford HIVdb version 6.0.10 (http://sierra2.stanford.edu/sierra/servlet/JSierra?action=sequenceInput), SCUEAL (http://www.datamonkey.org/dataupload_scueal.php), STAR (http://www.vgb.ucl.ac.uk/starn.shtml), REGAv2 (http://www.bioafrica.net/rega-genotype/html/subtypinghiv.html) and REGAv3 (http://bioafrica.mrc.ac.za:8080/rega-genotype-3.0.2/hiv/ typingtool/).

2.3.2 STANDARDIZATION OF ASSIGNMENTS AND MANUAL PHYLOGENETIC ANALYSIS

Comparison between subtyping tools required standardization of the assignments by different subtyping tools. Thus, 1) sub-subtypes were not taken into account; 2) "A-ancestral" and "A3 subtype" were assigned as A; 3) the assignment "-like" in REGAv3 which is the clustering with a pure subtype outside of the reference cluster with bootstrap > 70%, was considered as the subtype or CRF identified by the tool; 4) assignments "complex" or "recombinant" in SCUEAL and REGAv3 were considered recombinants; 5) different subtypes assigned by Stanford to the RT and PR were considered as evidence of recombination.

Each sequence from the clinical and LANL datasets was classified as concordant (all tools agreed on the assignment) or discordant (at least one tool had a different assignment than the other tools) based on the results of the 8 subtyping tools (see figure 2.1). The *MPhy* of concordant sequences was performed by using the 2008 Los Alamos curated subtypes and CRFs reference dataset (available at http://www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html), the sequences were aligned with ClustalW [28] and, if needed, the alignment was minimally edited with BioEdit [29]. Since assignment for concordant sequences is less problematic than for discordant sequences, and since the dataset is so huge, we opted for a fast NJ method with 1000 bootstrap replicates and a simple substitution model (HKY85), which we call *fast MPhy*. Such method has been proven useful for subtyping [30, 31], and it saves computation time. A query sequence was assigned to a particular clade if it clustered monophyletically inside that clade with bootstrap support >70% [14, 15, 32]. Otherwise, the query sequence was considered discordant.

The discordant sequences were further analyzed with the 2008 Los Alamos curated subtypes and CRFs reference dataset complemented with more and curated full-genome sequences available for each subtype in the database using a maximum of 15 sequences per subtype (available in http://www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/compendium.html). To optimize subtype classification of the discordant sequences [33, 34], we used as gold standard a *slow MPhy* using ML trees with 1000 bootstrap replicates and the best-fitting nucleotide substitution model (in this case GTR+I+Γ) [35, 36]. If the query sequence clustered monophyletically inside a clade with bootstrap support >70% it was assigned that clade, otherwise the sequence was further screened for recombination using SimPlot with a window size of 300 nts in steps of 20 nts [37]. For the sequences with no signal for recombination, the sequence was assigned the clade with the highest similarity in

SimPlot, and for all such sequences the majority of windows reached >70% bootstrap support. If there was a signal for recombination, the sequence was called URF, and the putative recombinant fragments were analyzed separately. A putative recombinant fragment with a phylogenetic signal >0.9 using TREEPUZZLE analysis was assigned a pure subtype or CRF if it clustered inside the respective subtype or CRF clade with >70% bootstrap support [32, 38], otherwise the fragment was called unclassified (U) [2].

The analyzed region for CRF01_AE and CRF14_BG is lacking a recombination breakpoint. We considered the *pol* region in the LANL dataset as correctly assigned to these two CRFs, since that assignment is based on the full genome. Such confirmation of breakpoints outside the *pol* region is not available for the clinical dataset, and this can cast doubt on the accurate assignment based on concordance between the tools and confirmed only by *fast MPhy* as described above. Therefore, in addition to *fast MPhy* for concordant sequences, all sequences that were assigned by any of the subtyping tools as either these CRFs or the parent pure subtype (even when concordant) were also analyzed with *slow MPhy* [39], which included all complete genomes of the CRF and parent pure subtype as reference sequences. In order to be considered CRF, the sequence should cluster inside the CRF reference cluster with more than 70% of bootstrap support [32, 38], otherwise it was considered the parent pure subtype. Finally, to verify these assignment, all sequences thus assigned subtype G or CRF14_BG were pooled with all full genome CRF14_BG and full genome subtype G sequences from LANL, and a single unrooted tree was constructed using RAxML [40](supplementary Figure 2.3). We found a big discrepancy between the different analyses for CRF14_BG and subtype G, and therefore, for the clinical dataset only, CRF14_BG and subtype G were pooled and analyzed together as a single 'subtype' called "CRF14_BG or G". Tools were considered to correctly assign these sequences when they scored either CRF14_BG or subtype G. We did not encounter problems with CRF01_AE, CRF01_AE was absent in our clinical dataset, and all subtype A sequences were confirmed not to be CRF01_AE.

2.3.3 SENSITIVITY, SPECIFICITY, REPRODUCIBILITY AND STATISTICAL ANALYSIS

The reference standard was *mPhy* of the *pol* region for the clinical dataset and the full-length genome assignment confirmed with *mPhy* of the *pol* region for the LANL dataset (the latter two were 100% concordant). Then we calculated the sensitivity with the formula TP/TP+FN and specificity with the

formula TN/TN+FP [41], where TP =true positives, FP= false positives, TN= true negatives, FN= false negatives.

To assess the reproducibility, we created a random subset of 100 sequences extracted from the clinical and LANL datasets that contained pure subtypes and CRFs and then ran this dataset 10 times with each tool (see figure 2.1). The reproducibility was, by definition, the percentage of times the same results were obtained when a subtyping tool was used 10 times on the same sequence, then the average of these percentages was calculated for the 100 sequences [41]. For G and CRF14_BG, only LANL sequences were used. If for a specific tool there were any discordant results between the runs, the entire clinical and LANL datasets were evaluated with that tool. The percentage of reproducibility was then calculated for this entire dataset, but again excluding G and CRF14_BG from the clinical dataset.

We evaluated the performance of the tools in the clinical, LANL, and clinical+ LANL datasets (herein named as overall dataset). Statistical significance of the difference between subtyping tools was evaluated with McNemar's test. The statistical analysis was calculated using R version 2.12.1.

2.4 RESULTS

2.4.1 SUBTYPE DISTRIBUTION OF THE DATASETS

Two sequences were excluded from the clinical dataset according to the quality assessment criteria [27]. The distribution of subtypes in the datasets is shown in tables 2.1, 2.2 and 2.3 according to the *MPhy* of the pooled clinical and LANL datasets (herein named as overall dataset) (the distribution of subtypes according to the clinical or the LANL dataset separately is shown in supplementary material 2.1 tables 2.2 and 2.3 respectively, phylogenetic trees for the clinical and the LANL datasets are in supplementary material figures 2.1 and 2.2). With regard to the non-B subtypes and the most common CRFs, despite the inclusion of two datasets, there was a limited number of H, J, K, CRF06_cpx, CRF07_BC, CRF12_BF and CRF14_BG sequences to reliably evaluate the performance of the subtyping tools, yet the results are still listed in the tables. The CRF13_cpx, CRF18_cpx, CRF25_cpx and CRF27_cpx were also found in the clinical dataset (see supplementary table 2.2), but these CRFs had a very low prevalence and not enough full genome sequences were found in LANL at the time of the collection of data to reliably assess the performance of the tools for these CRFs, but we included the results in supplementary materials (Supplementary table 2.2). We did not evaluate the performance

of the subtyping tools for CRFs that are not contributing substantially to the epidemic or that are poorly assigned. That is also why REGAv3 does not score all CRFs reported to date (see information about the epidemiological, geographical and recombination information in http://bioafrica.mrc.ac.za/CRFs/CRFs.php). For the clinical dataset, subtype G and CRF14_BG were pooled into a special class "G or CRF14_BG".

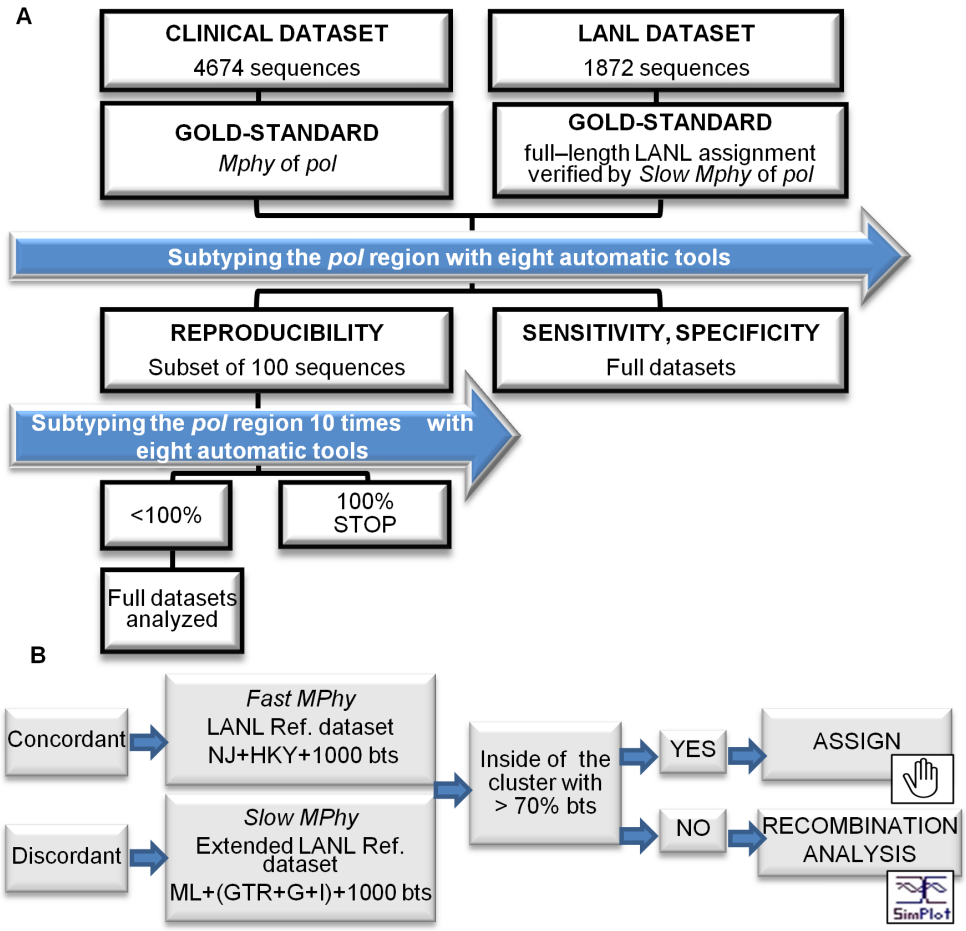


FIGURE 2. 1. METHODOLOGY OF THE STUDY

Figure 2.1: Methodology of this study: A: analyses performed on both datasets as was explained in section 2. **B**: manual phylogenetic analysis performed on both datasets. *Abbreviations*: *MPhy*: Manual Phylogenetic analysis, LANL: Los Alamos database, NJ: Neighbor joining, HKY: HKY model (Hasegawa, Kishino, Yano), GTR+G+I: General time-reversible model + Gamma + Proportion Invariant, bts: bootstrap, %: percentage.

2.4.2 PERFORMANCE OF THE SUBTYPING TOOLS FOR PURE SUBTYPE ASSIGNMENT

We evaluated the performance on the overall dataset, the clinical separately and the LANL dataset separately. Since we did not find many differences between the results of these datasets, we only show the performance of the overall dataset (see the performance of the tools for the clinical or the LANL dataset separately in supplementary tables 2.2 and 2.3 respectively). However, we found some discrepancies in the results for subtype A when we compared the two results of the clinical and the LANL datasets. For example, subtype A in the LANL dataset was 100% accurately classified by COMET, jpHMM and REGAv2 but the values in the clinical dataset were 76.5%, 86% and 73% respectively.

The sensitivity of each of the subtyping tools was more than 96% for subtype B and 98% for subtype C except for NCBI, which had lower values in both datasets (tables 2.1, 2.2, and 2.3, see details in supplementary tables 2.2 and 2.3). However the results were variable for other subtypes. For instance, COMET, jpHMM, and REGAv3 had a sensitivity of more than 90% for subtype A. jpHMM and REGAv3 obtained sensitivities of 100% for subtype D and subtype F respectively. REGAv3, Stanford and STAR classified correctly subtypes H, J, K, although the number of sequences available was limited. Noteworthy, the specificity for all pure subtypes was more than 98% (Tables 2.1, 2.2 and 2.3).

2.4.3 PERFORMANCE OF THE SUBTYPING TOOLS FOR RECOMBINANT FORMS

COMET, jpHMM, REGAv2, REGAv3, Stanford and STAR had sensitivities and specificities around 99% for the classification of CRF01_AE (see tables 2.1, 2.2 and 2.3, supplementary material table 2.3). However, the absence of a recombination breakpoint in the *pol* region makes the classification of this CRF challenging by the subtyping tools (see figure 2.2) [42]. Therefore, COMET and REGAv3 used other assignments; for example, COMET classified one sequence as “01_AE (check for 15_01B)” and REGAv3 identified 99% (168/169) as “HIV Subtype A (CRF01_AE)”. On the other hand, SCUEAL and NCBI classified independently CRF15_01B and CRF01_AE, with sensitivity dropping to 84% and 76% respectively.

Subtype	Total	Statistical-based tools											
		COMET				jpHMM				STAR			
		Sens	95% CI	Spec	95% CI	Sens	95% CI	Spec	95% CI	Sens	95% CI	Spec	95% CI
A	226	91.2	87.4-94.9	100.0	99.9-100	94.7	91.8-97.6	99.8	99.7-99.9	50.4	43.9-57.0	100.0	100-100
B	3023	99.1	98.8-99.5	99.8	99.6-99.9	99.0	98.7-99.4	99.6	99.4-99.8	97.6	97.0-98.1	99.6	99.4-99.8
C	628	99.7	99.2-100	100.0	100-100	100.0	100-100	100.0	100-100	99.8	99.5-100	100.0	100-100
D	69	97.1	93.1-100	100.0	100-100	100.0	100-100	100.0	100-100	89.9	82.7-97.0	100.0	100-100
F	129	89.1	83.8-94.5	100.0	100-100	92.2	87.6-96.9	100.0	100-100	89.1	83.8-94.5	100.0	100-100
LANL G*	34	100.0	85.1-100	99.5	99.0-99.7	97.1	84.7-99.9	99.4	98.9-99.7	73.5	55.6-87.1	99.4	98.9-99.7
H	11	90.9	73.9-100	100.0	100-100	90.9	73.9-100	100.0	100-100	100.0	100-100	100.0	100-100
J	6	50.0	10.0-90.0	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100
K†	2	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100
CRF01_AE†	169	99.4	98.3-100	100.0	100-100	100.0	100-100	100.0	100-100	98.8	97.2-100	100.0	100-100
CRF02_AG	272	96.3	94.1-98.6	100.0	99.9-100	*NA*	*NA*	*NA*	*NA*	96.0	93.6-98.3	99.7	99.5-99.8
CRF06_cpx	28	50.0	31.5-68.5	100.0	100-100	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
CRF07_BC†	10	90.0	55-100	100.0	100-100	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
CRF12_BF†	5	100.0	36.0-100	100.0	100-100	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
LANL CRF14_BG*	11	81.8	47.8-96.8	99.8	99.6- 99.9	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Clinical G+CRF14_BG‡	1571	98.7	98.2-99.3	99.9	99.8-100	98.3	97.7-99	99.2	98.9-99.5	95.3	94.2-96.3	99.6	99.4-99.8

Table 2.1: Performance of statistical-based subtyping tools. The sensitivity (Sens) and specificity (Spec) are reported for statistical-based tools. The values with 100% of performance are highlighted in dark gray; the values with more than 90% of performance are colored in light gray. * The values for G and CRF14_BG are based on the LANL dataset only. †These subtypes of CRFs only were available in the LANL dataset. ‡The 1571 sequences G and CRF14_BG of the clinical dataset were pooled as a single category. Abbreviations: n: sample, cpx: complex, LANL: Los Alamos dataset, NA: Not applicable, URF: Unique recombinant form

Subtype	Total	Similarity-based tools							
		NCBI				STANFORD			
		Sens	95% CI	Spec	95% CI	Sens	95% CI	Spec	95% CI
A	226	65.5	59.3-71.7	100.0	99.9-100	63.3	57.0-69.6	100.0	99.9-100
B	3023	84.7	83.4-85.9	98.8	98.4-99.1	98.3	97.9-98.8	99.0	98.7-99.4
C	628	92.4	90.3-94.4	100.0	100-100	98.9	98.1-99.7	100.0	99.9-100
D	69	79.7	70.2-89.2	100.0	100-100	91.3	84.7-98.0	100.0	100-100
F	129	87.6	81.9-93.3	99.9	99.9-100	71.3	63.5-79.1	100.0	100-100
LANL G*	34	47.1	29.8-64.9	100.0	99.7-100	97.1	84.7-99.9	99.4	98.9-99.7
H	11	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100
J	6	66.7	28.9-100	100.0	100-100	100.0	100-100	100.0	100-100
K†	2	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100
CRF01_AE†	169	76.3	69.9-82.7	99.7	99.4-100	100.0	100-100	99.6	99.3-99.9
CRF02_AG	272	48.5	42.6-54.5	99.8	99.7-99.9	98.9	97.7-100	98.0	97.6-98.3
CRF06_cpx	28	82.1	68.0-96.3	99.3	99.1-99.5	*NA*	*NA*	*NA*	*NA*
CRF07_BC†	10	100.0	59.0-100	100.0	100-100	*NA*	*NA*	*NA*	*NA*
CRF12_BF†	5	100.0	36.0-100	100.0	100-100	*NA*	*NA*	*NA*	*NA*
LANL CRF14_BG*	11	100.0	61.5-100	99.5	99.1-99.7	*NA*	*NA*	*NA*	*NA*
Clinical G+CRF14_BG‡	1571	99.7	99.5-100	98.8	98.4-99.2	97.5	96.7-98.3	98.9	98.5-99.2

Table 2.2: Performance of similarity-based subtyping tools. The sensitivity (Sens) and specificity (Spec) are reported for similarity-based tools. The values with 100% of performance are highlighted in dark gray; the values with more than 90% of performance are colored in light gray. * The values for G and CRF14_BG are based on the LANL dataset. †These subtypes of CRFs only were available in the LANL dataset. ‡ The 1571 sequences G and CRF14_BG of the clinical dataset were included in the total. Abbreviations: n: sample, cpx: complex, LANL: Los Alamos dataset,NA: Not applicable, URF: Unique recombinant form

Subtype	Total	Phylogenetic tools											
		REGAv2				REGAv3				SCUEAL			
		Sens	95% CI	Spec	95% CI	Sens	95% CI	Spec	95% CI	Sens	95% CI	Spec	95% CI
A	226	89.8	85.9-93.8	100.0	99.9-100	95.6	92.9-98.3	100.0	99.9-100	85.8	81.3-90.4	99.9	99.9-100
B	3023	97.3	96.7-97.8	99.8	99.7-99.9	99.2	98.9-99.5	99.3	99.0-99.5	96.3	95.7-97.0	99.9	99.5-99.9
C	628	99.8	99.5-100	100.0	100-100	100.0	100-100	100.0	100-100	99.0	98.3-99.8	100.0	100-100
D	69	84.1	75.4-92.7	100.0	100-100	88.4	80.9-93.0	100.0	100-100	95.7	90.8-100	100.0	100-100
F	129	93.8	89.6-98.0	100.0	100-100	100.0	100-100	100.0	100-100	89.9	84.7-95.1	100.0	100-100
LANL G*	34	100.0	85.1-100	99.4	98.9-99.7	100.0	85.1-100	99.4	98.9-99.7	97.1	84.7-99.9	99.9	99.6-100
H	11	90.9	73.9-100	100.0	100-100	100.0	100-100	100.0	100-100	90.9	73.9-100	100.0	100-100
J	6	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100
K†	2	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100	100.0	100-100
CRF01_AE†	169	99.4	98.3-100	100.0	100-100	99.4	98.3-100	100.0	100-100	84.0	78.5-89.5	100.0	100-100
CRF02_AG	272	64.7	59.0-70.4	100.0	99.9-100	98.9	97.7-100	100.0	100-100	33.8	28.2-39.4	100.0	100-100
CRF06_cpx	28	78.6	63.4-93.8	99.7	99.6-99.9	96.4	89.6-100	99.5	99.3-99.6	46.4	28.0-64.9	100.0	100-100
CRF07_BC†	10	100.0	59.0-100	100.0	100-100	100.0	59.0-100	100.0	100-100	40.0	9.6-70.4	100.0	100-100
CRF12_BF†	5	80.0	28.0-100	100.0	100-100	100.0	36.0-100	100.0	100-100	80.0	28.0-100	100.0	100-100
LANL CRF14_BG*	11	63.6	30.9-88.9	100.0	99.7-100	72.7	39.1-93.7	99.9	99.7-100	81.8	47.8-96.8	99.9	99.6-99.9
Clinical G+CRF14_BG‡	1571	98.8	98.3-99.3	99.8	99.6-99.9	99.8	99.6-100	98.9	98.6-99.3	97.6	96.9-98.4	99.7	99.5-99.9

Table 2.3: Performance of phylogeny-based subtyping tools. The sensitivity (Sens) and specificity (Spec) are reported for phylogenetic-based tools. The values with 100% of performance are highlighted in dark gray; the values with more than 90% of performance are colored in light gray. . * The values for G and CRF14_BG are based on the LANL dataset. †These subtypes of CRFs only were available in the LANL dataset. ‡ The 1571 sequences G and CRF14_BG of the clinical dataset were included in the total. Abbreviations: n: sample, cpx: complex, LANL: Los Alamos dataset, NA: Not applicable, URF: Unique recombinant form, REGAv3: REGA subtyping tool version 3, REGAv2: REGA subtyping tool version 2, URF: Unique recombinant form

The sensitivity and specificity was more than 96% using COMET, REGAv3, Stanford and STAR for CRF02_AG, while NCBI and SCUEAL had values below 50% for sensitivity. In most of the cases, NCBI misclassified some CRF02_AG as CRF30_0206 or CRF36_cpx whereas SCUEAL identified CRF02_AG sequences as "complex".

Regarding CRF06_cpx, REGAv3 had the highest sensitivity using 17 and 11 sequences in the clinical and LANL datasets respectively. In the clinical dataset low prevalent CRFs were also found, for instance, CRF25_cpx was identified 100% by REGAv3 in 9 sequences (see supplementary table 2.2), CRF18_cpx was classified with 100% of sensitivity with NCBI, REGAv3 and SCUEAL in 3 sequences. Only REGAv2 and REGAv3 correctly identified both sequences of CRF13_cpx and both sequences of CRF27_cpx, respectively. On the other hand, the LANL dataset included other prevalent CRFs such as CRF07_BC and CRF12_BF. A 100% of 10 sequences of CRF07_BC were identified correctly using NCBI, REGAv2 and, REGAv3. Similarly, a 100% of 5 sequences of CRF12_BF were correctly classified by COMET, NCBI and REGAv3.

A CRF was never assigned when the sequence clustered significantly with a CRF but outside of the reference clade. In this way, potential CRFs can have been assigned URF; however there is no safe way to assign such a sequence to a CRF in absence of the full genome. To overcome this potential limitation, the sequences assigned in this way as URF were further analyzed with the *slow Mphy* and the CRFs' reference dataset complemented with all curated full genome CRF sequences of Los Alamos database, and with SimPlot. We found that we had not missed any true CRFs and that all sequences that had been assigned URFs were truly URFs with unassigned fragments such as CRF06_cpx/U recombinant (23%, 79/336).

to discriminate the parent pure subtypes from the CRFs in geographic areas where the CRFs originated. **B:** An example analysis by SCUEAL. The query sequence of the genomic region *pol* has no evidence of recombination and it clusters with CRF15_01B. However, this *pol* gene is from a full-length genome sequence assigned as CRF01_AE. It is possible that it concerns here a CRF01_AE that was very closely related to the founder of CRF15_01B**. C:** An example analysis by REGAv3. The query sequence is the *pol* region trimmed from a full-length genome sequence assigned as CRF14_BG. In the pure subtype analysis it has a high support with subtype G and in the CRF analysis it has a high support with CRF14_BG. The algorithm classified it as G. This might be due to the fact that the sequence did not cluster reliably within the CRF14_BG clade. **D:** Example of the assignment "complex" by SCUEAL. The assignment of the full genome is CRF02_AG but the tool identified it as a G, CRF01_AE, CRF02_AG, A ancestral recombinant.

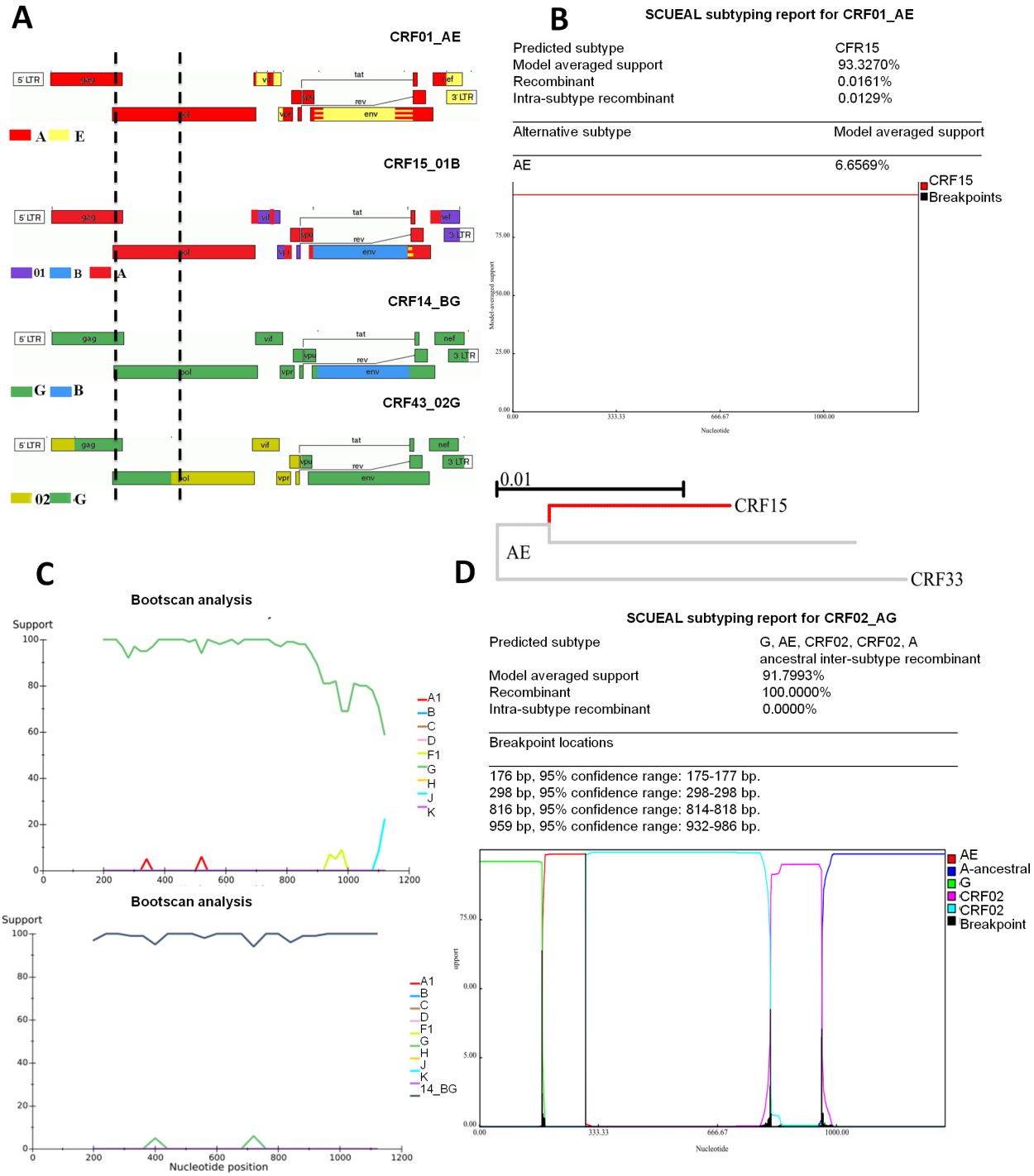


Figure 2.2: Frequent problems with the classification of CRFs. A: Adapted from the Los Alamos database available in http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html. The dashed lines delineate the region that is used for resistance testing and for the current performance analysis. The CRF01_AE and CRF15_01B are entirely subtype A in the *pol* gene, similarly CRF14_BG and CRF43_02G are entirely subtype G in this region. This complicates the identification since it is difficult

Other frequent URFs in the clinical dataset were various B/G recombinants (35%, 118/336), followed by other recombinants with unassigned fragments such as G/U recombinant (9%, 31/336) and B/U recombinant (5%, 16/336). The performance of the subtyping tools was not evaluated for URFs because of the limited number of sequences available and the complexity to evaluate recombinants when tools such as COMET, Stanford and STAR do not show the recombination breakpoints [17, 19, 20].

2.4.4 PERFORMANCE OF SUBTYPING TOOLS FOR G AND CRF14_BG

When confronted with the discrepancy in manual phylogenetic assignment for CRF14_BG and subtype G in the clinical dataset, as described in methods, we decided to only use the LANL dataset to calculate the performance on these two subtypes, because the *pol* region was trimmed from full genomes which is the only way to safely assign CRFs that lack a breakpoint in the here analyzed *pol* region (see figure 2.2) [43]. The sensitivity for subtype G was more than 97% for all the tools, except NCBI and STAR in the LANL dataset. The first had a sensitivity of 47% because of misclassification of some subtype G sequences as CRF14_BG or CRF43_02AG (See figure 2.2). STAR had 73% of sensitivity due to "unassigned" sequences. The specificity for subtype G was more than 99% for all the tools. Regarding CRF14_BG, COMET and SCUEAL classified 9 out of 11 sequences correctly followed by REGAv2 and REGAv3, which classified 8 out of 11 sequences as "Subtype G (CRF14_BG)".

To evaluate the discordances between the tools, the *fast* and *slow MPhy* procedures and the manual phylogenetic analysis of subtype G and CRF14_BG sequences in a single tree, as described in methods, we retrieved the *env* sequences from the Portuguese resistance database. We found 44 sequences (C2-V5 region of the gp120, mean length: 500 nts) from the same patients whose *pol* sequence was included in this study, 28 had their *pol* sequence classified as G using *slow MPhy*, and 16 as CRF14_BG. 4 subtypes B *env* sequences belonged to patient isolates classified as G in *pol*, similarly 6 subtype G *env* sequences belonged to patient isolates classified as CRF14_BG in *pol*. In addition, we determined how the Portuguese sequences clustered with respect to all full genome G/CRF14_BG sequences available from the Los Alamos database. All Portuguese G or CRF14_BG clinical sequences formed a monophyletic cluster within subtype G including all CRF14_BG full genomes, but the Portuguese clinical sequences, assigned as CRF14_BG and G by *slow MPhy* were paraphyletic with each other, suggesting that there may be a problem with the assignment of

CRF14_BG and this CRF may in fact consist of more than one CRF with very similar breakpoints (see supplementary material figure 2.3).

Using *slow Mphy*, 951 sequences were considered subtype G and in more than 96% of the cases these were also classified subtype G by jpHMM, REGAv2, REGAv3 and Stanford (in alphabetical order and see supplementary material table 2.2), however, the tools also assigned many of these sequences to CRF14_BG; with the exception of COMET and SCUEAL. Using *slow Mphy*, 620 sequences were considered CRF14_BG, and again, COMET and SCUEAL had the highest agreement with *slow MPhy*, but these values were just 62% and 56% respectively. Given that we did no longer consider the slow *MPhy* reliable for subtype G and CRF14_BG, these performance statistics are also not reliable. We therefore chose to pool the subtype G and CRF14_BG sequences from the Portuguese clinical database, as they were assigned by *slow MPhy*, and compute the performance of the tools on the combined class "G or CRF14_BG". The sensitivity was above 99% for NCBI and REGAv3, followed by COMET, jpHMM, and REGAv2 with more than 98% while specificity was around 99% for all the tools.

2.4.5 PERFORMANCE OF SUBTYPING TOOLS FOR PR AND RT SEPARATELY

The performance of the tools on PR and RT separately was only evaluated on the LANL dataset and the results are shown in supplementary data (supplementary table 2.4 and table 2.5). When analyzing PR separately, COMET and jpHMM showed similar performance as for *pol* with regard to pure subtypes and the CRF01_AE from the LANL dataset. However, the sensitivities varied for the other tools. For instance, Stanford and STAR had a better sensitivity for subtype A, but NCBI a worse sensitivity for subtypes A, C, F, and G. Stanford and STAR had similar sensitivities for CRF02_AG but other tools had a lower performance like COMET, NCBI, REGA and SCUEAL. With regard to RT sequences, in general, the performance of the subtyping tools was the same as for the *pol* sequences, again using the LANL dataset only. The exceptions included tools with improved sensitivity such as NCBI for subtypes B, D and G; REGAv3 for CRF02_AG, Stanford for subtypes A, F and STAR for subtype G. However, SCUEAL had decreased sensitivity for subtype CRF02_AG (supplementary table 2.4).

2.4.6 PERFORMANCE OF REGAV3 VERSUS THE PREVIOUS VERSION REGAV2

The performance of REGAv3 was better than REGAv2 for subtypes B (p=0.01), and CRF02_AG (p=0.001) in the genomic region *pol* in both datasets (see tables 2.1, 2.2 and 2.3 and supplementary tables 2.2 and 2.3). In the case of CRF02_AG, for instance, the changes in the decision tree for REGAv3 improved sensitivity by properly assigning CRF02_AG sequences that were classified as "check the bootscan" by REGAv2.

We compared the new term "-like" of REGAv3 with *MPhy*. "Subtype B-like" corresponded to subtype B in 28 sequences analyzed with the *MPhy*, while 8 were B/U recombinants and 2 were B/G recombinants. In the case of "subtype G-like", the *MPhy* showed 1 subtype G, 3 recombinants G and one CRF06_cpx. "Subtype A1-like" and "Subtype F1-like" were identified in 2 and 4 sequences respectively, but in both cases the *MPhy* showed these were A and F subtypes.

We also evaluated the performance for PR and RT separately but only using the LANL dataset. In the analysis of RT, REGAv3 had higher sensitivity than REGAv2 for subtype B (98.8 versus 91.9). However, the performance in PR was variable (see supplementary table 2.5) because the REGA subtyping tool algorithm is different for sequences shorter than 800 nts. In short sequences, the criteria are based on clustering only and potential recombination is not analyzed. This is because the window size for recombination analysis in REGA is chosen as 400 nts to avoid losing too much phylogenetic signal [44] and recombination is scanned in steps of 50 nts, such that no meaningful recombination signal can be obtained for such short sequences. REGAv3 correctly identified only 54% of the subtype A sequences, which is better than the 45% with REGAv2, but the sensitivity for subtype B decreased from 86% to 73% when comparing REGAv2 versus REGAv3. There was no difference between the tools for subtype C and G. The lack of phylogenetic signal in the short fragment of PR significantly reduced the sensitivity of the tool, for instance REGAv2 had a sensitivity of 27% and REGAv3 had 24% for subtype F; while none of the subtypes D, CRF01_AE and CRF02_AG sequences analyzed were properly identified.

2.4.7 REPRODUCIBILITY OF HIV-1 SUBTYPING TOOLS:

COMET, jpHMM, NCBI, Stanford, and STAR were 100% reproducible. Subtyping tools based on phylogenetic methods such as REGAv2, REGAv3 and SCUEAL were reproducible with values of 99.2% (95% CI: 99.10-99.26), 99.2% (95% CI: 99.15-99.30) and 96.4% (95% CI: 96.27-96.60)

respectively. When the clinical and LANL datasets were independently analyzed, the reproducibility did not change significantly; for instance, the reproducibility for REGAv2 was 99.1% and 99.5%, for REGAv3 98.8% and 99.7 %, and for SCUEAL 95.4% and 98.1%, respectively.

For REGAv2 and REGAv3, most of the non-reproducible results were related to subtype B. For example, when the sequence was subtype B, the tool might classify it as subtype B, or as "check the report" or "B/D recombinant" (for details see also Supplementary table 2.1). In this paper we considered "subtype–like" as belonging to the subtype (or CRF) identified by the tool for REGAv3. If "subtype-like" would be considered discordant, then the sensitivity of REGAv3 would go down (from 99.2% to 97.8% for subtype B, the assignment with the highest number of –like assignments), and the reproducibility would go up (from 98.9% to 99.2%). For SCUEAL, the non-reproducible results were related mainly to subtype B and CRF02_AG. For instance, subtype B was sometimes classified as "B/D recombinant" by the tool, and CRF02_AG was frequently assigned as "complex".

2.5 DISCUSSION

In the present study, we compared and described the performance of the phylogenetic based automated HIV typing and subtyping tool REGAv3, its previous version REGAv2, and six other commonly used automated HIV-1 subtyping tools: one other phylogenetic based tool, SCUEAL; three statistical-based tools, COMET, jpHMM and STAR; and two similarity based tools, NCBI and Stanford. We used only the *pol* (PR+RT) region that is usually sequenced for drug resistance testing [45, 46], since this generates the largest datasets for which these tools are designed. This restriction was also made since tools such as SCUEAL and Stanford cannot assign sequences outside this region [17, 24]. In addition, we analyzed with phylogenetic analysis two datasets; one dataset was derived from clinical samples and another from the Los Alamos database (available in http://www.hiv.lanl.gov/). We used the clinical data from Portugal because it is one of the European countries with the highest proportion of non B-subtypes [47, 48]. Since phylogenetic analysis of full-length genomes is the gold standard to define the current subtypes [2], we also used the assignment of the *pol* region confirmed with *MPhy* and trimmed from full-length genomes of pure subtypes and the most common CRFs from the Los Alamos database, with the aim to better evaluate the performance of automated subtyping tools on all epidemic subtypes and CRFs.

Our primary aim was to evaluate the new subtyping tool REGAv3 versus other available tools. REGAv3 identified subtype B, most of the non-B pure subtypes and the most frequent CRFs with a sensitivity

and specificity of more than 96% in the *pol* region. The classification of REGAv3 for subtype B and CRF02_AG has improved compared to REGAv2; additionally, with an updated algorithm and reference dataset, REGAv3 is designed to identify most of the epidemic CRFs. Consequently, REGAv3 performs equally well as other tools, such as COMET and jpHMM, which also had high sensitivity and specificity for classifying most of the pure subtypes and such as COMET, Stanford and STAR for classifying CRF01_AE and CRF02_AG.

Concerning some previous reports that suggested a low performance of REGAv2 compared with other subtyping tools [26, 49], it is pertinent to add that these discrepancies were almost exclusively due to unassigned reports, and not due to wrong assignments [49]. The number of unassigned sequences was reduced in REGAv3 compared to REGAv2, by introducing the term "like". Although 74% of the sequences assigned as "like" were classified as a pure subtype by the *MPhy* in our analysis, 26% of the samples showed evidence of recombination; therefore, this terminology "like" is indeed useful, as it alerts the user to further verify these sequences. This helps to reduce the number of inaccurate assignments, while also reducing the number of unassigned sequences. Thus, REGAv3 uses the "subtype-like" assignment to indicate the most likely subtype for a particular strain, and at the same time to caution for potential discrepancies, thereby increasing the usefulness of the tool both for epidemiological statistical purposes where it is important to have as few as possible unassigned sequences, and for situations where correct assignment is more important.

One of the aims of the study was to give guidance as to which tool would perform well in a context of HIV-1 surveillance activities where the overall prevalence and spread of the epidemic is important to estimate [3]. We showed that no subtyping tool is able to classify all HIV-1 clades with a 100% accuracy, and we highlighted the difference in performance of the tools according to the subtype (or CRF) or the region analyzed such as PR, RT or PR+RT, such that our results can be directly compared with other studies [26, 49, 50]. In general, our findings corroborate that phylogenetic-based, statistical-based tools and the similarity-based tool Stanford perform well for the most frequent subtypes worldwide such as B and C [3, 19-21, 24, 51]. However for other important clades such as A, D, F, G, CRF01_AE, CRF02_AG, COMET and REGAv3 correctly identified most strains while the remaining tools failed much more often.

Since we only had the *pol* region of the clinical dataset, we cannot exclude that other regions of the genome belong to other subtypes or CRFs, as has been reported in other cohorts [52, 53]. However, overall, the performance of the tools was similar for the clinical and the LANL dataset, suggesting that for most subtypes and CRFs, our performance evaluation is valid. The exception is the assignment of

subtype A and CRF06_cpx. For example, COMET correctly identified subtype A in 76% of the sequences from the clinical dataset and 100% of the LANL dataset, similarly this tool also identified only 18% of the CRF06_cpx in the clinical dataset and 100% in the LANL dataset (see supplementary tables 2.2 and 2.3). The reason for this discrepancy has not been further investigated, but we suspect that by not taking into account the evolutionary relationship of the sequences, statistical-based tools like COMET are prone to overfitting on the training dataset (LANL).

Another reason for variation in the performance of subtyping tools is the analysis of PR and RT separately [26, 50]. Most of the tools had similar performance for the RT and the *pol* region. However, with regard to the PR region separately, statistical-based tools such as COMET and jpHMM had a higher performance than the other tools. These disagreements, for instance with REGA, occurred because short sequences with low phylogenetic signal were frequently reported as "unassigned"[26]. The philosophy of REGA is to avoid assigning sequences with low phylogenetic signal, with the aim to avoid false conclusions about evolutionary relationships [44, 54]. Consequently, the user must be aware that short sequences require further examination to be correctly classified.

Current definitions of some subtypes and CRFs contribute to problems with the performance of automated tools. For example, the relationship between CRF02_AG and its parental strain A and G is still a matter of debate [24, 55, 56]. For other CRFs, such as CRF02_AG, CRF07_BC, CRF08_BC, CRF12_BF, and CRF17_BF, atypical breakpoints were found in the sequences assigned to these CRFs in the Los Alamos database [56], suggesting perhaps a wrong assignment in this database. In fact, the proper assignment of several CRFs can be disputed, for example when the recombinant region is so small that there is not sufficient phylogenetic signal for classification (e.g. around 100 nts for CRF12_BF, CRF20_BG, CRF35_A1D, CRF41_CD). Finally, several CRFs are so closely related to each other (e.g. CRF20_BG, CRF23_BG and CRF24_BG) that automated tools have great difficulty to discriminate between them. A thorough re-analysis of all CRFs is therefore urgently needed.

The absence of breakpoints in the region of study was another frequent cause of misclassifications; for instance CRF01_AE was classified as CRF15_01B [57] (see figure 2.2) and G was classified as CRF14_BG. There was no problem with the evaluation of the CRF01_AE *pol* sequences, but we had to question either the value of the manual phylogenetic analysis as gold standard for CRF14_BG, or the definition of CRF14_BG itself. The clinical dataset is derived from the country where this CRF originated, and the prevalence of the parent subtype G and of CRF14_BG as defined by manual phylogenetic analysis was very high. However we had access to *env* sequences from several patient

isolates that were included in the *pol* dataset, and for several of these, there was discordance with the *pol* assignment, as has been reported before [53]. In a joined phylogenetic analysis of all G and CRF14_BG sequences, the CRF14_BG sequences were not monophyletic but were spread among the subtype G sequences, and this was also the case for the LANL full genome CRF14_BG sequences. As a result, and only for the clinical dataset, we pooled CRF14_BG and subtype G in a single, but separate, classification 'CRF14_BG or G'.

Phylogenetic-based subtyping tools had lower reproducibility than similarity-based tools or statistical-based tools and the sequences causing this problem were not consistent across tools. The bootstrapping procedure in the tree-based algorithms is responsible for this lower reproducibility. Bootstrapping is a random process of resampling [25, 32], and each bootstrap sample is a different sample. For a more robust assignment, it would be needed to perform 1000 bootstrap samples, but this would cost too much computer time, and the phylogenetic-based tools are already slow. Other causes for the lower reproducibility were the introduction of new thresholds in the bootscan support of REGAv3 and the term "complex" in SCUEAL. This led sometimes to misidentification especially of subtype B as B/D recombinant because of the high similarity between these subtypes in the *pol* region [2, 58].

Other considerations that influences the widespread use of subtyping tools are the operational characteristics (see table 2.4). For example, the analysis of a bulk of sequences usually takes more time with phylogenetic-based tools than statistical-based tools [19, 24, 25], and this is important for any application in a context of a large dataset. On the other hand, for the surveillance of the HIV-1 epidemic it is sometimes important to have information on recombination breakpoints, which are only shown in phylogenetic-based tools and jpHMM [21, 24].

Although we included prevalent epidemiological non-B subtypes and CRFs, we acknowledge the limited number of samples available for subtypes H, J, K, CRF06_cpx, CRF07_BC and CRF12_BF, which prevents us from drawing firm conclusions for these subtypes and CRFs. We also did not evaluate all available tools, since many are based on similarity, and some have as their main objective the evaluation of antiretroviral resistance rather than subtyping [18]. We included the two most commonly used similarity-based tools, NCBI and Stanford [16, 27].

Other factors, that influence the performance of subtyping tools, are the high recombination rate of HIV-1 [4] and human migration as determinants of global HIV dynamics [59]. HIV-1 recombination increases the complexity and frequency of recombinant forms [56], while migration has driven the dissemination

of subtypes to new regions and established new epidemics [60]. As a consequence the subtyping tools should be regularly updated, especially tools which do not consider the intrinsic biologically relevant evolutionary relationships like statistical or similarity-based tools. The analysis of an epidemic where many subtypes or new CRFs are prevalent must be identified with COMET or phylogenetic-based tools that have an updated reference dataset [19, 22, 24].

2.6 CONCLUSIONS AND RECOMMENDATIONS

To our knowledge, this is the first study with an extensive comparison between subtyping tools, and manual phylogenetic analysis in PR, RT, PR+RT in two large datasets: a clinical dataset and a LANL dataset in which *pol* region was trimmed from full-length genomes. The performance of the new REGAv3 to identify subtype B and CRF02_AG in the *pol* region was much better than with REGAv2. REGAv3 had a very good performance in classifying pure subtypes, similar to that of COMET, jpHMM and SCUEAL, and it was also very good at identifying CRFs in the *pol* region, comparable to the best other tool, COMET. REGAv3 and COMET are currently the best available tools to automatically subtype HIV-1 sequences, however recombination breakpoint analysis is not possible with COMET. The performance of jpHMM is comparable but this tool has the big disadvantage that it does not classify CRFs, except for CRF01_AE [21].

We could draw some general recommendations from this analysis to use in future surveys of HIV-1 genetic diversity. First, automated tools might be useful for subtyping large *pol* datasets that are used in clinical and surveillance settings [24, 51]. Nevertheless, if accuracy is important, for example in individual patient follow-up or in detailed epidemiological analyses, it is necessary to use at least two subtyping tools whose overall performance is high in the genetic region analyzed such as COMET and REGAv3. This methodology has been previously used in different studies that required stringent analyses of large datasets [12, 61-63]. This comes at the cost of speed, which is determined by the slower of the two tools, the phylogenetic-based tool. The discordant sequences between the two tools can then be analyzed using manual phylogenetic analysis, still the gold standard. Second, for very short sequences such as PR, tools like COMET are recommended given that REGAv3 will give a considerable number of unassigned sequences, but only if accuracy is not a big issue. Therefore, we insist that for short sequences with low phylogenetic signal, such as PR, manual phylogenetic analysis is still needed [44]. Third, subtyping is often done in the context of an individual patient follow-up, using PR+RT sequences that are available from resistance genotyping. Thus it is often the case that

resistance and subtyping are analyzed together in the clinical settings. Stanford does provide this information; however, the main goal of Stanford is to provide an accurate algorithm of resistance rather than subtyping (http://hivdb.stanford.edu/DR/asi/releaseNotes/index.html#hivdb_subtyping). If this tool is used, analysis of subtypes A, F and CRFs should be complemented with other statistical-based or phylogenetic-based tools. For example, REGAv3 is also on the same website. Fourth, the use of at least two automated tools to classify subtypes in the patient follow-up could also be useful, for example, for clinical collaborators that have little experience with manual analysis. However, if superinfection is suspected, phylogenetic analysis should be carried out [64].

Characteristics	Tools							
	Phylogenetic			Similarity		Statistical		
	REGAv3	REGAv2	SCUEAL	NCBI	STANF	COMET	jpHMM	STAR
Analysis of full genomes	+	+	-	+	-	+	+	+
Exact recombination breakpoints	+	-	+	*	-	-	+	-
Intra-subtype recombination	-	-	+	-	-	-	-	-
Latest CRF that can be analyzed†	CRF47_BF	CRF14_BG	CRF43_02G	CRF43_02G	CRF02_AG	CRF49_cpx	CRF01_AE	CRF02_AG
Batch analysis online	1000	1000	500	1	>100‡	10 Mb§	5‖	500
Waiting job queue	-	-	+	-	-	-	-	-
Average time for 500 sequences ¶	~5 h	~4 h	~3 h	-	~5 min	sec	~2 - 7 h ‖	~15 min
Average time for 1 sequence **	~min	~min	~min	sec	sec	sec	sec-min	sec
Part of resistance analysis	-	-	-	-	+	-	-	-
Phylogenetic signal analysis	+	+	-	-	-	-	-	-
Summary table report	+	-	+	-	-	-	-	-
Graphical visualization of results	+	+	+	+	-	-	+	+
Position of sequence according to HXB2 reference	+	+	-	-	-	-	+	-
Download additional files (csv, txt, fasta, etc)	+	+	+	+	+	+	+	+

Table 2.4: Operational characteristics of subtyping tools. The *NCBI report shows an approximation of the breakpoints, † Some CRFs are excluded from the analysis, such as those with a limited number of strains available in LANL**,** or where the *pol* region cannot be discriminated from other CRFs (see http://bioafrica.mrc.ac.za/CRFs/CRFs.php) ‡ Stanford is able to analyze up to 100 sequences at a time (character limit: 600,000). However there are other options to analyze more sequences. § COMET accepts files with a maximum size of 10Mb (around 8000 sequences PR+RT). ‖ jpHMM has the option to download a command line program without limit of batch analysis. In addition, there is an option to speed up the program. ¶ We ran 500 sequences with pure subtypes and CRFs five times in different days. Then we calculated the average time. ** We ran 5 times a recombinant, the average time for analysis of 1 sequence is about 1 minute for Phylogenetic-based tools and seconds for the other tools. Abbreviations: (+) characteristic available in the tool, (-) characteristic not available, (h) hours, (min) minutes, (sec) seconds.

REFERENCES

[1] UNAIDS-WHO. GLOBAL REPORT: UNAIDS report on the global AIDS epidemic. Geneva; 2012.

[2] Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, et al. HIV-1 nomenclature proposal. Science (New York, NY. 2000 Apr 7;288(5463):55-6.

[3] Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS (London, England). 2011 Mar 13;25(5):679-89.

[4] Mansky LM, Temin HM. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. Journal of virology. 1995 Aug;69(8):5087-94.

[5] Jetzt AE, Yu H, Klarmann GJ, Ron Y, Preston BD, Dougherty JP. High rate of recombination throughout the human immunodeficiency virus type 1 genome. Journal of virology. 2000 Feb;74(3):1234-40.

[6] Baeten JM, Chohan B, Lavreys L, Chohan V, McClelland RS, Certain L, et al. HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. The Journal of infectious diseases. 2007 Apr 15;195(8):1177-80.

[7] Abecasis AB, Deforche K, Bacheler LT, McKenna P, Carvalho AP, Gomes P, et al. Investigation of baseline susceptibility to protease inhibitors in HIV-1 subtypes C, F, G and CRF02_AG. Antiviral therapy. 2006;11(5):581-9.

[8] Martinez-Cajas JL, Pant-Pai N, Klein MB, Wainberg MA. Role of genetic diversity amongst HIV-1 non-B subtypes in drug resistance: a systematic review of virologic and biochemical evidence. AIDS reviews. 2008 Oct-Dec;10(4):212-23.

[9] Camacho RJ, Vandamme AM. Antiretroviral resistance in different HIV-1 subtypes: impact on therapy outcomes and resistance testing interpretation. Curr Opin HIV AIDS. 2007 Mar;2(2):123-9.

[10] Brenner BG, Oliveira M, Doualla-Bell F, Moisi DD, Ntemgwa M, Frankel F, et al. HIV-1 subtype C viruses rapidly develop K65R resistance to tenofovir in cell culture. AIDS (London, England). 2006 Jun 12;20(9):F9-13.

[11] Wainberg MA, Brenner BG. Role of HIV Subtype Diversity in the Development of Resistance to Antiviral Drugs. Viruses. 2010 Nov;2(11):2493-508.

[12] Hue S, Hassan AS, Nabwera H, Sanders EJ, Pillay D, Berkley JA, et al. HIV Type 1 in a Rural Coastal Town in Kenya Shows Multiple Introductions with Many Subtypes and Much Recombination. AIDS research and human retroviruses. 2011 Jul 19;27.

[13] Snoeck J, Van Dooren S, Van Laethem K, Derdelinckx I, Van Wijngaerden E, De Clercq E, et al. Prevalence and origin of HIV-1 group M subtypes among patients attending a Belgian hospital in 1999. Virus research. 2002 Apr 23;85(1):95-107.

[14] Pasquier C, Millot N, Njouom R, Sandres K, Cazabat M, Puel J, et al. HIV-1 subtyping using phylogenetic analysis of pol gene sequences. J Virol Methods. 2001 May;94(1-2):45-54.

[15] Yahi N, Fantini J, Tourres C, Tivoli N, Koch N, Tamalet C. Use of drug resistance sequence data for the systematic detection of non-B human immunodeficiency virus type 1 (HIV-1) subtypes: how to create a sentinel site for monitoring the genetic diversity of HIV-1 at a country scale. The Journal of infectious diseases. 2001 May 1;183(9):1311-7.

[16] Rozanov M, Plikat U, Chappey C, Kochergin A, Tatusova T. A web-based genotyping resource for viral sequences. Nucleic acids research. 2004 Jul 1;32(Web Server issue):W654-9.

[17] Liu TF, Shafer RW. Web resources for HIV type 1 genotypic-resistance test interpretation. Clin Infect Dis. 2006 Jun 1;42(11):1608-18.

[18] Beerenwinkel N, Daumer M, Oette M, Korn K, Hoffmann D, Kaiser R, et al. Geno2pheno: Estimating phenotypic drug resistance from HIV-1 genotypes. Nucleic acids research. 2003 Jul 1;31(13):3850-5.

[19] Struck D, Perez-Bercoff D, Devaux C, Schmit JC. COMET: A Novel approach to HIV-1 subtype prediction. 8th European HIV Drug Resistance Workshop; 2010 17- 19 March; Sorrento, Italy; 2010.

[20] Myers RE, Gale CV, Harrison A, Takeuchi Y, Kellam P. A statistical model for HIV-1 sequence classification using the subtype analyser (STAR). Bioinformatics (Oxford, England). 2005 Sep 1;21(17):3535-40.

[21] Schultz AK, Zhang M, Bulla I, Leitner T, Korber B, Morgenstern B, et al. jpHMM: improving the reliability of recombination prediction in HIV-1. Nucleic acids research. 2009 Jul;37(Web Server issue):W647-51.

[22] de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, et al. An automated genotyping system for analysis of HIV-1 and other microbial sequences. Bioinformatics (Oxford, England). 2005 Oct 1;21(19):3797-800.

[23] Alcantara LC, Cassol S, Libin P, Deforche K, Pybus OG, Van Ranst M, et al. A standardized framework for accurate, high-throughput genotyping of recombinant and non-recombinant viral sequences. Nucleic acids research. 2009 Jul;37(Web Server issue):W634-42.

[24] Kosakovsky Pond SL, Posada D, Stawiski E, Chappey C, Poon AF, Hughes G, et al. An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. PLoS computational biology. 2009 Nov;5(11):e1000581.

[25] Abecasis AB, Wang Y, Libin P, Imbrechts S, de Oliveira T, Camacho RJ, et al. Comparative performance of the REGA subtyping tool version 2 versus version 1. Infect Genet Evol. 2010 Apr;10(3):380-5.

[26] Holguin A, Lopez M, Soriano V. Reliability of rapid subtyping tools compared to that of phylogenetic analysis for characterization of human immunodeficiency virus type 1 non-B subtypes and recombinant forms. Journal of clinical microbiology. 2008 Dec;46(12):3896-9.

[27] Rhee SY, Kantor R, Katzenstein DA, Camacho R, Morris L, Sirivichayakul S, et al. HIV-1 pol mutation frequency by subtype and treatment experience: extension of the HIVseq program to seven non-B subtypes. AIDS (London, England). 2006 Mar 21;20(5):643-51.

[28] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic acids research. 1994 Nov 11;22(22):4673-80.

[29] Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series. 1999;41:95-8.
[30] Posada D, Crandall KA. Selecting models of nucleotide substitution: an application to human immunodeficiency virus 1 (HIV-1). Molecular biology and evolution. 2001 Jun;18(6):897-906.
[31] Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular biology and evolution. 2010 Feb;27(2):221-4.
[32] Hillis DM, Bull JJ. An Empirical-Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis. Systematic biology. 1993 Jun;42(2):182-92.
[33] Kuhner MK, Felsenstein J. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Molecular biology and evolution. 1994 May;11(3):459-68.
[34] Leitner T, Escanilla D, Franzen C, Uhlen M, Albert J. Accurate reconstruction of a known HIV-1 transmission history by phylogenetic tree analysis. Proceedings of the National Academy of Sciences of the United States of America. 1996 Oct 1;93(20):10864-9.
[35] Posada D. jModelTest: phylogenetic model averaging. Molecular biology and evolution. 2008 Jul;25(7):1253-6.
[36] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and evolution. 2011 Oct;28(10):2731-9.
[37] Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. Journal of virology. 1999 Jan;73(1):152-60.
[38] Schmidt HA, Strimmer K, Vingron M, von Haeseler A. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics (Oxford, England). 2002 Mar;18(3):502-4.
[39] Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic biology. 2010 May;59(3):307-21.
[40] Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics (Oxford, England). 2006 Nov 1;22(21):2688-90.
[41] Banoo S, Bell D, Bossuyt P, Herring A, Mabey D, Poole F, et al. Evaluation of diagnostic tests for infectious diseases: general principles. Nature Reviews Microbiology. 2010 Dec;8(12 Suppl):S17-29.
[42] Carr JK, Salminen MO, Koch C, Gotte D, Artenstein AW, Hegerich PA, et al. Full-length sequence and mosaic structure of a human immunodeficiency virus type 1 isolate from Thailand. Journal of virology. 1996 Sep;70(9):5935-43.
[43] Delgado E, Thomson MM, Villahermosa ML, Sierra M, Ocampo A, Miralles C, et al. Identification of a newly characterized HIV-1 BG intersubtype circulating recombinant form in Galicia, Spain, which exhibits a pseudotype-like virion structure. Journal of acquired immune deficiency syndromes (1999). 2002 Apr 15;29(5):536-43.
[44] Strimmer K, von Haeseler A. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proceedings of the National Academy of Sciences of the United States of America. 1997 Jun 24;94(13):6815-9.
[45] Vandamme AM, Camacho RJ, Ceccherini-Silberstein F, de Luca A, Palmisano L, Paraskevis D, et al. European recommendations for the clinical use of HIV drug resistance testing: 2011 update. AIDS reviews. 2011 Apr-Jun;13(2):77-108.
[46] Thompson MA, Aberg JA, Hoy JF, Telenti A, Benson C, Cahn P, et al. Antiretroviral treatment of adult HIV infection: 2012 recommendations of the International Antiviral Society-USA panel. Jama. 2012 Jul 25;308(4):387-402.
[47] Abecasis A, Wensing AM, Vercauteren J, Paraskevis D, Van de Vijver D, Albert J, et al. HIV-1 genetic diversity in Europe and its demographic predictors. Demographic determinants of HIV-1 subtype distribution in Europe. *6th European HIV Drug Resistance Workshop*. Budapest, Hungary 2008.
[48] Abecasis AB, Wensing AM, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DA, et al. HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. Retrovirology. 2013 Jan 14;10(1):7.
[49] Yebra G, de Mulder M, Martin L, Perez-Cachafeiro S, Rodriguez C, Labarga P, et al. Sensitivity of seven HIV subtyping tools differs among subtypes/recombinants in the Spanish cohort of naive HIV-infected patients (CoRIS). Antiviral Res. 2010 Jan;89(1):19-25.
[50] Loveday C, MacRae E, and on behalf of the ICVC Clinical Collaborative Research Group. Limitations in using online tool to determine HIV-1 subtype in clinical patients: a comparison of 5 tools. XV International HIV Drug Resistance Workshop: Basic Principles & Clinical Implications; 2006 13-17 June, 2006; Sitges, Spain: Antiviral Therapy; 2006. p. S129.
[51] Gifford R, de Oliveira T, Rambaut A, Myers RE, Gale CV, Dunn D, et al. Assessment of automated genotyping protocols as tools for surveillance of HIV-1 genetic diversity. AIDS (London, England). 2006 Jul 13;20(11):1521-9.
[52] Njouom R, Pasquier C, Sandres-Saune K, Harter A, Souyris C, Izopet J. Assessment of HIV-1 subtyping for Cameroon strains using phylogenetic analysis of pol gene sequences. J Virol Methods. 2003 Jun 9;110(1):1-8.
[53] Abecasis AB, Martins A, Costa I, Carvalho AP, Diogo I, Gomes P, et al. Molecular epidemiological analysis of paired pol/env sequences from Portuguese HIV type 1 patients. AIDS research and human retroviruses. 2011 Jul;27(7):803-5.
[54] Revell LJ, Harmon LJ, Collar DC. Phylogenetic signal, evolutionary process, and rate. Systematic biology. 2008 Aug;57(4):591-601.
[55] Abecasis AB, Lemey P, Vidal N, de Oliveira T, Peeters M, Camacho R, et al. Recombination confounds the early evolutionary history of human immunodeficiency virus type 1: subtype G is a circulating recombinant form. Journal of virology. 2007 Aug;81(16):8543-51.
[56] Zhang M, Foley B, Schultz AK, Macke JP, Bulla I, Stanke M, et al. The role of recombination in the emergence of a complex and dynamic HIV epidemic. Retrovirology. 2010;7:25.
[57] Tovanabutra S, Watanaveeradej V, Viputtikul K, De Souza M, Razak MH, Suriyanon V, et al. A new circulating recombinant form, CRF15_01B, reinforces the linkage between IDU and heterosexual epidemics in Thailand. AIDS research and human retroviruses. 2003 Jul;19(7):561-7.

[58] Leitner T, Escanilla D, Marquina S, Wahlberg J, Brostrom C, Hansson HB, et al. Biological and molecular characterization of subtype D, G, and A/D recombinant HIV-1 transmissions in Sweden. Virology. 1995 May 10;209(1):136-46.
[59] Rambaut A, Posada D, Crandall KA, Holmes EC. The causes and consequences of HIV evolution. Nat Rev Genet. 2004 Jan;5(1):52-61.
[60] Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. Nat Rev Genet. 2009 Aug;10(8):540-50.
[61] Faria NR, Suchard MA, Abecasis A, Sousa JD, Ndembi N, Bonfim I, et al. Phylodynamics of the HIV-1 CRF02_AG clade in Cameroon. Infect Genet Evol. 2011 Mar;12(2):453-60.
[62] Jacobs GB, Loxton AG, Laten A, Robson B, van Rensburg EJ, Engelbrecht S. Emergence and diversity of different HIV-1 subtypes in South Africa, 2000-2001. Journal of medical virology. 2009 Nov;81(11):1852-9.
[63] Yebra G, de Mulder M, del Romero J, Rodriguez C, Holguin A. HIV-1 non-B subtypes: High transmitted NNRTI-resistance in Spain and impaired genotypic resistance interpretation due to variability. Antiviral Res. 2010 Feb;85(2):409-17.
[64] Ntemgwa M, Gill MJ, Brenner BG, Moisi D, Wainberg MA. Discrepancies in assignment of subtype/recombinant forms by genotyping programs for HIV type 1 drug resistance testing may falsely predict superinfection. AIDS research and human retroviruses. 2008 Jul;24(7):995-1002.

CHAPTER 3

MOLECULAR EPIDEMIOLOGY AND PHYLODYNAMICS OF THE HIV-1 EPIDEMIC IN COLOMBIA

Submitted:

Andrea-Clemencia Pineda-Peña, Nuno Rodriguez-Faria, Francisco-Javier Diaz, Patricia Olaya, Li Guangdi, Daniela Vanegas-Otálvaro, Casper Møller Frederiksen, Arley Gomez-Lopez, Dimitrios Paraskevis, Philippe Lemey, Anne-Mieke Vandamme. The Colombian HIV-1 epidemic is dominated by subtype B and linked with Spain. 2014

3.1 SUMMARY

Objective: To characterize the molecular epidemiology of HIV-1 in Colombia and to elucidate its spatial phylodynamics in the context of international and national migration.

Design: HIV-1 *pol* nucleotide sequences were compiled from 7 distinct geographic regions of Colombia sampled between 2002-2007.

Methods: 610 HIV-1 sequences were obtained from Bogotá, Cali, Medellin, and the Coffee, Santander, Caribbean, Orinoquía and Amazonas regions. Phylogenetic subtyping was performed and a dataset was compiled including the closest available sequence data to the Colombian epidemic, and a representative set of worldwide subtype B sequences. Phylodynamic and phylogeographic analyses revealed the most significant routes of spatial dispersal and viral flow between geographic regions was quantified.

Results: The Colombian epidemic is dominated by subtype B (99.8%) in all analyzed geographic regions. One sub-subtype F1 was isolated in Bogotá in 2003. Phylogeographic analyses suggest multiple links between the Colombian and the Spanish HIV-1 subtype B epidemic. Subtype B was imported multiple times from the late-1970s onwards. Within Colombia, we found strong statistical support for viral migration from Bogotá to Medellin, Cali, Santander and the Caribbean and Coffee regions.

Conclusions: This is the first study using a large sample, which shows the predominance of Subtype B in Colombia. We find that viral populations in Bogotá play a central role in shaping the HIV-1 epidemic also within other regions of Colombia. The Colombian HIV-1 epidemic has most likely been seeded multiple times from various countries from the late-1970s onwards, and strong links exist between the Spanish and the Colombian epidemic.

3.2 INTRODUCTION

According According to UNAIDS Colombia had a HIV-1 prevalence of approximately 0.5% in 2012 [1]. Molecular epidemiological studies up to 2002 have shown that HIV-1 subtype B clade predominates in Medellin [2] and Bogotá [3, 4] and, according to a recent study, also in the port city of Barranquilla in the northern Caribbean region [5]. Although HIV-1B predominates in the Americas [6], other HIV-1 clades circulate in Latin America. HIV-1 subtype F has been found in the men who have sex with men (MSM) population in Bogotá [3]. Subtype F also circulates in Ecuador and Peru [7] that border southern

Colombia. In other countries it was shown that countrywide HIV-1 genetic diversity is not always homogeneous. For example, subtype C clade circulates throughout southern Brazil [8, 9], and its introduction in the region has attracted considerable attention [10-12]. Subtype C also circulates in countries bordering Colombia, such as Venezuela [13], Panama [14] and in Ecuador [7].

Due to the fast evolutionary rate of HIV, phylodynamic techniques can be used to recover the evolutionary and spatial history of viral dispersal in great detail [15-17]. Molecular clock analyses have shown that subtype B finds its ultimate origins in the Central African subtype D around the 1960s [18, 19] from where it was introduced, most probably due to a single founder effect, to Haiti and made its way the US as subtype B [20, 21]. Moreover, it was recently shown that the predominance of subtype B in most of Central America was the result of a major single introduction occurring in the 1980s, a study in which Colombian strains were limited while contemporary country-specific HIV epidemics in this region are characterized by early subdivision into discrete sub-epidemics [22, 23]. Recent phylogeographic studies suggested that South American subtype B epidemic was also seeded from the Caribbean region [23]. However, due to the limited amount of genetic data available from Colombia, little is known about the evolutionary history of HIV-1 in this country.

Here we report the molecular epidemiology of HIV-1 in Colombia based on a large comprehensive collection of virus sequence data (n=610) sampled from seven regions between 2002 and 2007, and we investigate the most probable routes of subtype B introduction and the pathways of HIV migration within Colombia. The identification of the most important foci of HIV transmission in a country with high frequency of internal displacement [24], regional conflict and international migration [25] may be useful for public health purposes.

3.3 MATERIAL AND METHODS

3.3.1 STUDY POPULATION

Nucleotide data was collected from 619 Colombian patients, as part of genotypic resistance tests carried out in the *Centro de Análisis Molecular* in Colombia between 2002 and 2007; this center performed nationwide resistance testing on mainly treated patients during this period. Sequences were obtained by population sequencing using TRUGENE® kit (Siemens Healthcare Diagnostics, Germany). Only one sequence per patient was included. Quality control (QC) was evaluated with the Los Alamos database QC tool (http://www.hiv.lanl.gov/content/sequence/QC/index.html) using previously suggested parameters [26]. Nine sequences were excluded from subsequent analyses because of an excessive

number of stop codons and frameshifts (n=7) or hypermutations (n=2). No socio-demographic information except for city of residence was available.

The final nucleotide dataset consisted of 610 sequences sampled from patients from the following regions: 1) the capital Bogotá (n=275), 2) Cali (n=77), 3) Medellin (n=64), 4) Coffee region (n=42) that included sequences from Armenia, Pereira and Manizales cities, 5) Santander region (n=26) that included sequences from Bucaramanga and Cucuta cities, 6) Caribbean region (n=18) that included sequences from Barranquilla, Cartagena, Santa Marta and Sincelejo cities, 7) Orinoquía and Amazonas (OAM) region (n=6) that included sequences from Villavicencio and Florencia cities (Figure 3.1A). A total of 102 sequences were not assigned to any particular city or region and their most probable location was inferred by using an ambiguous trait code in a phylogeographic approach.

3.3.2 CLASSIFICATION OF HIV-1 DIVERSITY AND RECOMBINATION TESTING

To accurately classify the HIV-1 *pol* sequences, we first used two automated subtyping tools: REGA version 3; available at http://regatools.med.kuleuven.be/typing/v3/hiv/typingtool/ [27] and COMET version 2; available at http://comet.retrovirology.lu/ [28]. Each sequence was classified as either concordant (both tools agreed on the assignment) or discordant (tools had a different assignment). The discordant assignments were subsequently classified using a manual phylogenetic analysis approach [27]. Sequences were screened for phylogenetic signal with TreePuzzle [29] and recombination using SimPlot [30], the Phi-Test [31] implemented in SplitsTree version 4.10 [32] and the genetic algorithm single breakpoint (SBP) method [33] available at www.datamonkey.org.

3.3.3 RELATION OF COLOMBIAN SEQUENCES AND GLOBAL SUBTYPE B DIVERSITY

To investigate the spatial evolutionary dynamics of the HIV-1 subtype B Colombian epidemic in a global context, we first downloaded all other available Colombian sequences (n=22), and complemented our dataset of Colombian sequences with the 10 closest publicly available sequences for each Colombian sequence, identified through a BLASTn search available at http://blast.ncbi.nlm.nih.gov/Blast.cgi. This step was done using an in-house PERL script that retrieves the most similar sequences from a fasta file (script is available upon request). The respective collection year and country of sampling were compiled and all publicly available sequence data were screened for quality control and recombination as described above. This compiled dataset was hereafter named BLASTn-COL dataset (supplementary

tables 3.1 and 3.2). To evaluate whether the country of origin of the strains changed according to the number of closest sequences retrieved by BLASTn, a sensitivity analyses was performed with the 30 and 50 most similar sequences (supplementary table 3.3).

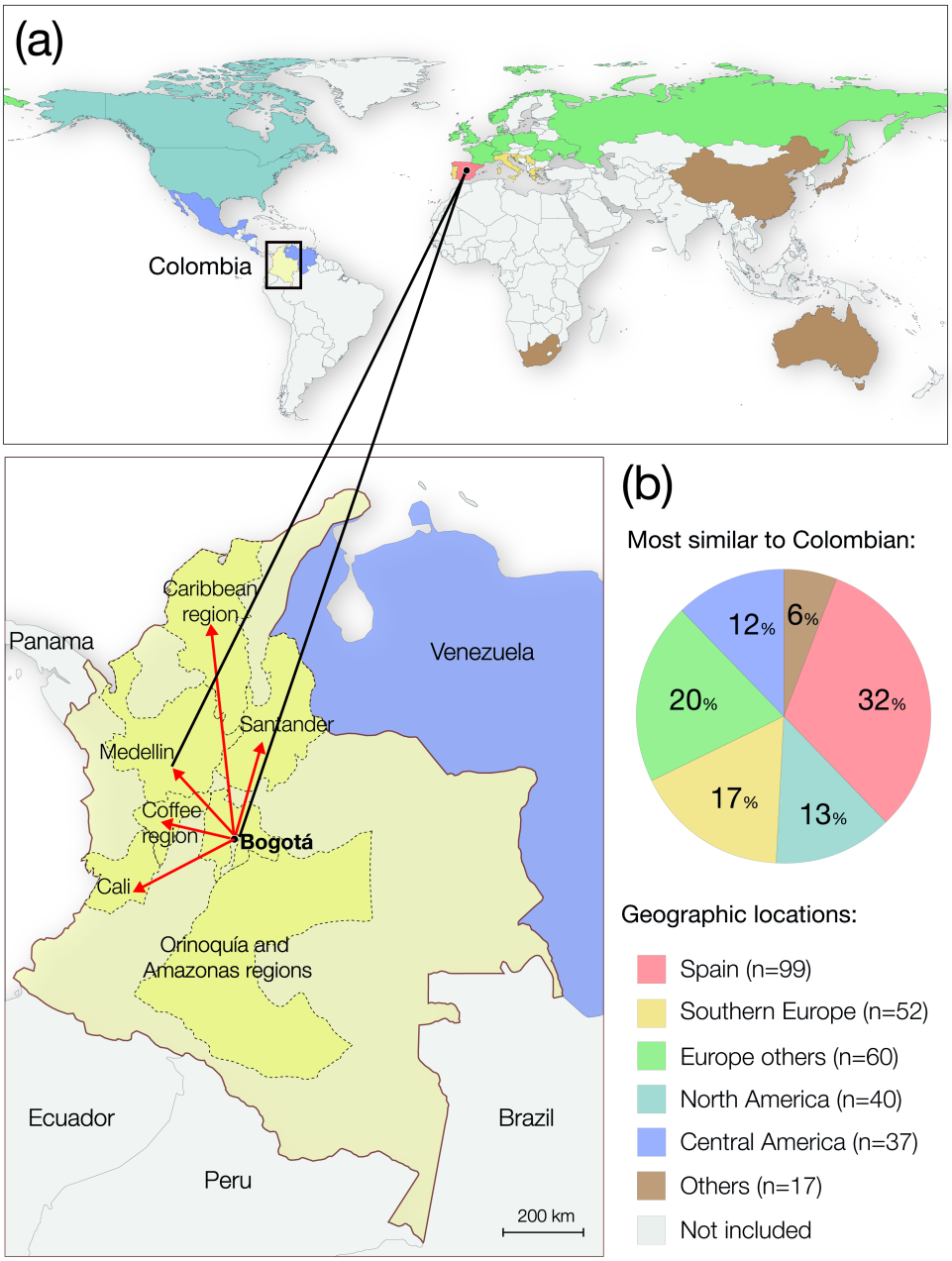


FIGURE 3. 1. SIGNIFICANT EPIDEMIOLOGICAL LINKS OF COLOMBIAN SUBTYPE B WITH OTHER COUNTRIES AND WITHIN COLOMBIA

Figure 3.1: Significant epidemiological links of Colombian Subtype B with other countries and within Colombia. (a) The migration pathways are shown according to the BSSVS analysis in an international context and within Colombia. (b) This analysis was based on the most similar sequences to each Colombian sequence retrieved by BLASTn

To have a complete overview of the Colombian subtype B diversity and its relation with the neighboring countries and worldwide epidemic, we additionally downloaded available *pol* subtype B sequences from Los Alamos database (http://www.hiv.lanl.gov/). Because of the extensive number of sequences, we selected a random dataset of 25 subtype B *pol* sequences per year for each country worldwide. In addition, we included all available sequences from bordering countries such as Brazil, Ecuador, Panama, Peru, and Venezuela (http://www.hiv.lanl.gov/, accessed Sept 2013). Clones, duplicates and multiple sequences per patient were excluded.This compiled dataset was hereafter named LANL-Borders dataset (Supplementary table 3.4). The 2 joined datasets were called BLASTn-COL+ LANL-Borders dataset.

All sequences were aligned using Muscle [34] and minimally edited using SeaView [35]. Thirty-seven codon positions associated with drug resistance mutations were removed from the alignments [36]. A Maximum likelihood (ML) phylogenetic tree was constructed with the BLASTn-COL dataset by using SeaView under the best fitting nucleotide substitution model (GTR+I+Γ) with the approximate likelihood-ratio test based on the Shimodaira-Hasegawa-like (SH) procedure as statistical support. ML trees with the BLASTn-COL+ LANL-Borders datasets were constructed with FastTree under the GTR+Γ nucleotide substitution model with the SH procedure as statistical support [37, 38].

3.3.4 BAYESIAN PHYLOGENETIC ANALYSES

To investigate the spatiotemporal dispersal of subtype B, posterior distributions of phylogenetic trees and ancestral geographic states were estimated in a joint Bayesian Markov chain Monte Carlo (MCMC) framework [39]. First, for the BLASTn-COL dataset we used a GTR+4Γ+I nucleotide substitution model [40], an uncorrelated lognormal molecular clock model [41] and a non-parametric skyline tree prior that avoids assumptions regarding demographic epidemic growth [42]. MCMC analyses were run in triplicate for 250 million steps and trees were sampled every 50,000th state. 10% of the burn-in was discarded and the remaining samples were combined using LogCombiner. From the resulting posterior tree distribution, 3,500 trees were removed as burn-in and the remaining posterior sample was resampled at a frequency of 10% to generate a total of 1,000 posterior empirical trees. BEAGLE library was used in all Bayesian computations to improve run-time [43, 44].

3.3.5 PHYLOGEOGRAPHIC LINKAGE AND COUNTING NUMBER OF TRANSITIONS BETWEEN LOCATIONS

To further improve computational time, spatial diffusion analyses were run on the empirical set of 1,000 trees representative of the posterior distribution. We first identified the most significant non-zero rates of viral dispersal using a Bayesian Stochastic Search Variable Selection (BSSVS) procedure [45], under a discrete asymmetric diffusion model [46]. The statistical significance of the non-zero rates were summarized using a Bayes Factor (BF) test that compares the posterior to the prior odds that a particular rate is required to explain the diffusion process [47]. Second, we quantify the number of transitions for the set of non-zero rates using a robust counting procedure [48-50]. Maximum clade credibility trees (MCC) were summarized using TreeAnnotator and visualized with FigTree v1.4.0 (available at http://tree.bio.ed.ac.uk).

3.3.6 STATISTICAL PHYLOGEOGRAPHIC ANALYSES

To confirm the links obtained by Bayesian phylogenetic analysis at international level, we used a non-parametric model of geographic spread based on parsimony using the BLASTn-COL dataset and a global subtype B dataset as previously described [51]. ML trees were inferred with the GTR+Γ nucleotide substitution model with 250 bootstraps as statistical support using RAxML to perform this analysis [52].

3.3.7 SENSITIVITY TO SPATIAL SAMPLING SCHEME

Three additional datasets comprising a total of 157 Colombian sequences were compiled to investigate the potential effect of sampling bias in estimating the geographic dispersal of HIV-1 in Colombia. Sequences were randomly selected according to a weighted sampling scheme that took into consideration the proportion of HIV-1 in people living in each Colombian region by 2007 [53]. This resulted in datasets comprised by 50 sequences from Bogotá, 29 sequences from Cali, 25 sequences from Medellin, 18 sequences from the Caribbean, 13 sequences from Santander, 13 sequences from the Coffee region, 5 sequences from the OAM region and 4 sequences from unknown regions. Non-Colombian sequences from the BLASTn search were maintained identical (Supplementary table 3.1), resulting in a total of 461 sequences for analysis. Bayesian phylogenetic analyses described above were also performed on these datasets.

3.4 RESULTS

3.4.1 MOLECULAR EPIDEMIOLOGY OF HIV-1 IN COLOMBIA

A total of 610 *pol* nucleotide sequences were available from individuals who enrolled in the genotypic resistance-testing center between 2002 and 2007. All samples were assigned with year of collection and 83.2% (508/610) with city of residence of the patient (Supplementary table 3.1). Most *pol* HIV-1 sequences (alignment length: 807 nt; positions HXB2: 2262-3290) belonged to subtype B (99.8%, 609/610), reflecting a regionally uniform predominance of this clade in Colombia. Only one sequence, sampled in Bogotá in 2003, was classified as sub-subtype F1 (0.2%, 1/610). No evidence of intersubtype recombination was found; however, eight subtype B sequences had discordant results among the three recombination tools because of the presence of unknown fragments with very low phylogenetic signal (average length: 335 nt). Therefore, these sequences were excluded for phylogeography analyses. The remaining sequences had an adequate phylogenetic signal according to TreePuzzle analyses.

3.4.2 PHYLOGENETIC CLUSTERING OF COLOMBIAN VS GLOBAL SUBTYPE B SEQUENCES

The dataset compiled contained the 10 most similar sequences per Colombia strain selected by BLASTn search as described in methods. The countries from where those additional sequences were originating are mentioned in Supplementary table 3.3. The sensitivity analysis did not show differences in the proportion of the countries when the number of similar sequences was increased from 10 to 30 and 50. However, only the dataset with the 10 closest sequences was analyzed because of the additional computational time.

To frame the Colombian subtype B epidemic in a global context, we performed a ML phylogeny using BLASTn-COL dataset (non-Colombian sequences: n=304 and Colombian sequences: n=623; See Figure 3.1B, and supplementary tables 3.1 and 3.2 for details), and the LANL-Borders dataset that included 2968 sequences from neighboring countries and 4963 sequences samples worldwide (Figure 3.2A, Supplementary table 3.4). The ML tree shows that HIV-1B Colombian sequences were intermingled within the global subtype B diversity. In addition, 17.3% Colombian sequences (108/623) fell within 7 clusters with statistical support ranging from 0.846 to 0.992, suggesting only few local transmission clusters among the strains analyzed, and supporting a scenario of multiple independent introductions to Colombia.

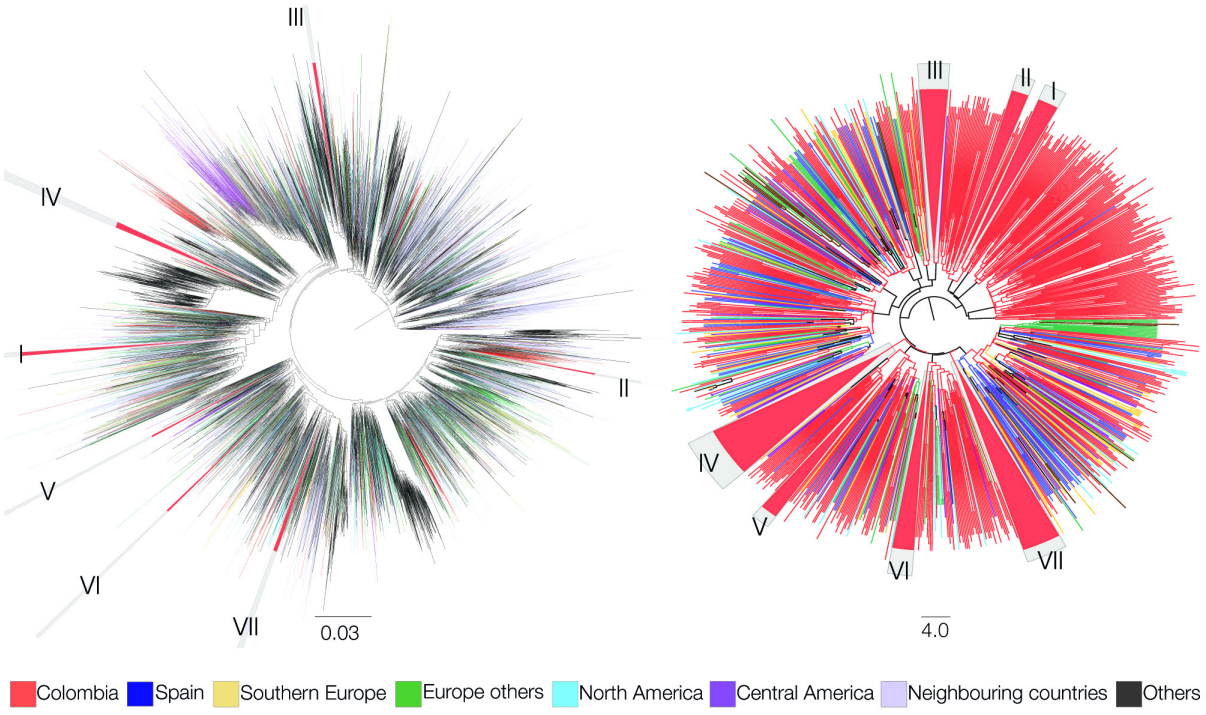


FIGURE 3. 2. PHYLOGENETIC RELATION OF COLOMBIAN AND GLOBAL SUBTYPE HIV-1B DIVERSITY

Figure 3.2: Phylogenetic relation of Colombian and global subtype HIV-1B diversity. Seven clusters consisting of more of 10 Colombian sequences are shown in red. (a) Midpoint-rooted-maximum likelihood tree of the Colombian sequences, the most similar sequences retrieved by BLASTn, neighboring countries and other worldwide sequences (b) Bayesian maximum clade-credibility tree of the Colombian sequences and the most similar sequences retrieved by BLASTn.

3.4.3 TIMING OF HIV-1B INTRODUCTIONS IN COLOMBIA

The Bayesian MCC tree in Figure 3.2B shows six well-supported clusters of HIV-1B with ≥10 Colombian sequences and posterior probability >0.90 [22], using the BLASTn-COL dataset (Table 3.1). Cluster IV of 32 Colombian sequences was found in the ML with 0.992 of statistical support (Figure 3.2A) but with a posterior probability of 0.8 in the MCC tree (Figure 3.2B). All the HIV-1B Colombian clusters were mostly composed of Colombian sequences, except clusters III and V that included also some sequences from Greece and Spain, respectively. These clusters had an MRCA between 1980 and 1985, confirming a scenario of multiple introductions prior to those dates as suggested by the ML analysis.

Clusters	n	TMRCA	95% HPD		Cities, regions or countries involved
I	12	1983,8	1979,9	1987,3	Bogotá (12)
II	10	1980,6	1979,9	1986,6	Amazonas (1), Bogotá (2), Cali (4), Coffee region (1), Medellin (2)
III	17	1982,4	1978,5	1986,8	Bogotá (10), Cali (1), Coffee region (5), Greece (1)
IV	32	1981,1	1977,0	1985,1	Amazonas (1), Bogotá (9), Cali (7), Coffee region (11), Medellin (3), Santander region (1)
V	10	1985,1	1981,6	1988,4	Bogotá (9), Spain (1)
VI	13	1981,1	1977,1	1985,6	Bogotá (6), Cali (1), Coffee region (6)
VII	22	1980,2	1976,3	1984,8	Bogotá (18), Medellin (2) and Santander region (2)

Table 3.1: The mean of the TMRCA for each cluster and the 95 % HPD intervals. The numbers between parentheses are the number of sequences per city, region or country. Abbreviations: TMRCA: The most recent common ancestor, HPD: Highest posterior density.

3.4.4 PHYLOGEOGRAPHIC ANALYSIS OF HIV-1 SUBTYPE B IN COLOMBIA AT INTERNATIONAL LEVEL

To investigate the geographic origins of HIV-1B in Colombia we used both a BSSVS approach that identifies non-zero viral migration rates along with their BF significance and a robust counting procedure that quantifies the number of viral migration events between particular sets of locations. In this case, we focus on estimating the number of migration events between locations linked by significant non-zero rates as estimated by the BSSVS procedure using the BLASTn-COL dataset.

We found a strong BF support for HIV-1B migration between Spain and Bogotá [link “i” in Figure 3.3A, BF>1000; number of Markov jumps (MJ): 89.4; 95% Bayesian credible interval (BCI): 74, 105], as well as between Spain and Medellin [link “l” in Figure 3.3A, BF=447.6, MJ: 20; 95% BCI: 8-29] (Figure 3.1A, 3.3A and Supplementary table 3.5). Another link between Bogotá and Spain was found [BF=40.9; 95% BCI: 0, 10], but the uncertainty interval did not exclude zero (link “r” in Figure 3.3A). Other significant viral pathways obtained using a BSSVS approach for which the uncertainty intervals of the estimated number of migrations did not overlap with zero are shown in supplementary Table 3.5.

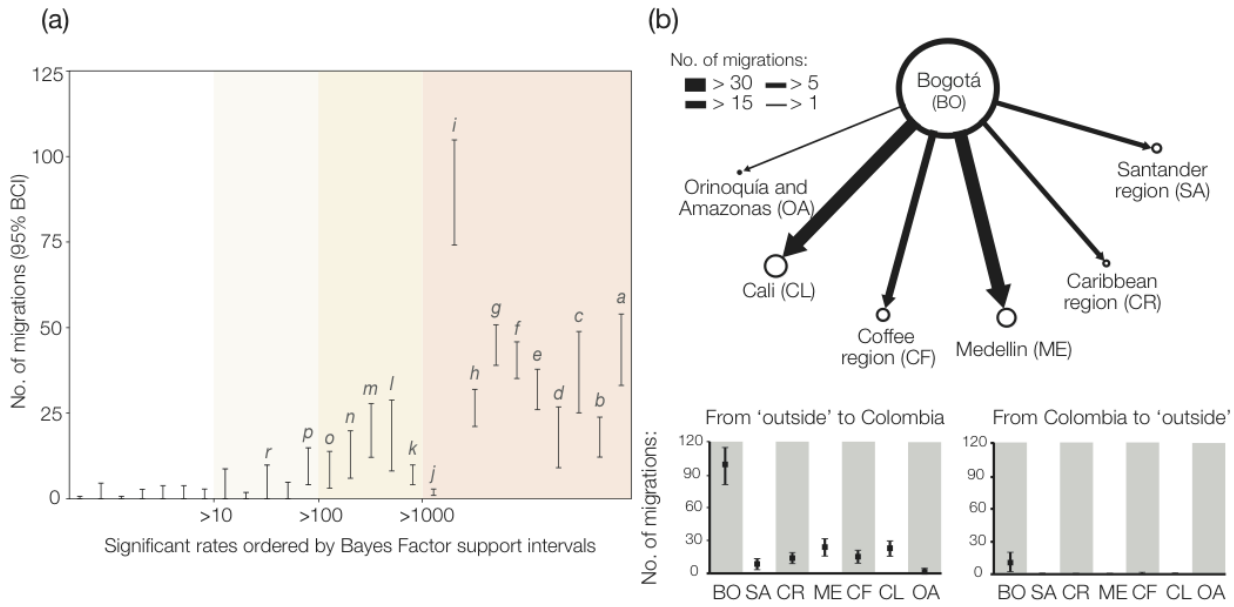


FIGURE 3. 3. BAYESIAN PHYLOGEOGRAPHIC ANALYSES OF HIV-1 SUBTYPE B IN COLOMBIA

Figure 3.3: Bayesian phylogeographic analyses of HIV-1 subtype B in Colombia. (**a**) The number (No) of migrations are shown according to the Bayes Factor and Markov jump analysis. The links with non zero rates are shown between Bogotá, Colombia and Cali (*a*), Bogotá, Colombia and Santander region, Colombia (*b*), Bogotá, Colombia and Medellín, Colombia (*c*), Bogotá, Colombia and Coffee region, Colombia (*d*), Spain and North America (*e*), Spain and West-East-Northern Europe (*f*), Spain and Southern Europe (*g*), Spain and Central-South America (*h*), Spain and Bogotá, Colombia (*i*), West-East-Northern Europe and Others (*j*), Spain and Others (*k*), Spain and Medellín, Colombia (*l*), Spain and Cali, Colombia (*m*), Spain and Coffee region, Colombia (*n*), Bogotá, Colombia and Caribbean region, Colombia (*o*), and Spain and Caribbean region, Colombia (*p*). A link between Bogotá, Colombia and Spain (*r*) was found with a Bayesian credible interval (BCI) overlapping with zero. See also Supplementary table 3.4. (**b**) Number (No) of migrations within Colombia and, between Colombia and other countries (outside) using *k*=8 locations.

To confirm the migration links with Colombia, statistical phylogeographic analysis was performed. A significant exporting migration was found between Colombia and Spain, while no other links were found for other countries with history of migration like neighboring countries, which gives additional robustness to our results with the Bayesian phylogenetic analysis. In an attempt to evaluate the directionality of the migration, ML trees were performed by including the Colombian sequences, the Spanish sequences retrieved from BLASTn, other Spanish sequences from the global dataset and 3 reference sequences of subtype A, or C or D as the out-group. Similar analyses were performed with

the LANL-Borders dataset. In this analyses , the directionality of the migration changed with the outgroup or the dataset such that the direction of migration remained non-conclusive.

3.4.5 PHYLOGEOGRAPHIC ANALYSIS OF HIV-1 SUBTYPE B WITHIN COLOMBIA

At the national level, since sampling bias towards more heavily sampled locations may interfere with spatial root estimation [54], we first used an alternative discretization scheme of the BLASTn-COL dataset (n=927). While the first scheme considered *k*=13 geographic locations (Figure 3.1B and Supplementary table 3.1), the second sampling scheme considered *k*=8 geographic traits and grouped all sequences from abroad Colombia into a single geographic trait, named “outside”. The results for the different discretization schemes were identical, with overlapping number of migrations along with respective uncertainty intervals for viral migration within Colombian locations. In Figure 3.3B, we show the estimated number of migrations (Markov jumps) for each well-supported migration pathway within Colombian locations. Here, we found statistical evidence for viral flow between Bogotá and several other localities. Notably, all Colombian locations were directly linked to the capital Bogotá with significant BF support (Supplementary table 3.5), suggesting that Bogotá has had a central role in HIV-1B Colombian epidemic. Most viral migrations were most likely seeded from Bogotá to Cali [BF>1000, MJ: 44.5; 95% BCI: 33, 54], followed by Bogotá to Medellin [BF>1000, MJ: 37.6; 95% BCI: 25, 49]. In addition, we found around an average of 19 and 17.6 migrations from Bogotá to Santander and to the Coffee region, respectively.

3.4.6 EVALUATION OF THE POTENTIAL EFFECT OF SAMPLING BIAS

To investigate whether the number of sequences per geographic location had any effect on the results, we used three additional data sets (n=461, *k*=13), in which the total number of Colombian sequences (n=157) was randomly down-sampled and proportional to the prevalence in each location. Notably, the BSSVS approach identifies 81.3% (13/16) of the non-zero migration rates as being not only identical between the three down-sampled data sets but also to the larger data BLASTn-COL dataset (Supplementary table 3.5), suggesting that our results are robust to sampling

3.5 DISCUSSION

This large study is the first to describe the molecular epidemiology of the HIV-1 epidemic in Colombia at a nationwide scale, and to reveal its international connectedness. We report the predominance of subtype B in all the seven geographic regions analyzed, despite the migration flow during these years [24, 25]. Our approach takes advantage of the full data available for Colombia even though for some samples the location was unknown, our approach could estimate that the 102 sequences with unknown location sampled in Colombia were most probably from individuals residing mainly in Bogotá (71.6%). Our findings are consistent with previous studies developed in Bogotá, Barranquilla and Medellin [2, 4, 5] and complemented the understanding of the molecular epidemiology in Colombia. These results may be explained by the fact that immigration was low in contrast with emigration until 2007 [25], and the main countries of destination such as Andean region, Spain and United States also had a high prevalence of subtype B infections [4, 6], whereas countries with high prevalence of non B subtypes or BF recombinants such as Brazil and the Southern Cone [6] had limited migration to Colombia [25].

The present study characterized the spatiotemporal spread of subtype B to Colombia. Seven clusters with 17% of Colombian sequences and all other Colombian sequences spread over the tree suggest a scenario of multiple introductions that were observed around late 70s and early 80s. This implies that the Colombian HIV-1B spread occurred after the introduction of HIV in the United States and Central America in the middle 60s and early 70s [20, 22]. The TMRCA estimates agree with the first HIV-1 case reported in the Caribbean region in 1983 [53].

Since our methodology was focused to shed light on the spread at the country level rather than to elucidate the origin of subtype B in the Americas, we used a BLASTn search to determine the countries involved in the Colombian epidemic. Most of the sequences were from Spain, and other European countries rather than neighboring countries with available data. Bayesian analyses showed a spatial root of Colombian subtype B in Spain, but it was not confirmed with ML and statistical phylogeography analyses. Since the main countries involved with migration and tourism to and from Colombia were United States and Venezuela [25, 55], and Spain had limited migration to or from Colombia between the 70s and 80s [56], our finding was unexpected. A study suggested a Caribbean introduction of subtype B to Colombia and Venezuela driven by the booming oil economy in the latter country [23], which also received immigrants from Spain at that time [56]. However, we cannot rule out alternative scenarios given the lack of demographic data of the sequences. Further research and data is required to establish the origins of the subtype B epidemic in Colombia.

Although the link between Colombia and Spain could be related with convenience sampling [54], we diminished this bias using three random datasets according to the HIV-1 prevalence in Colombia, sensitivity analyses and different phylogenetic techniques to evaluate our results in different datasets that represented the global diversity of subtype B. As consequence, we consider our data is robust enough to confirm the link between the Spanish and the Colombian epidemic, but the directionality remains unclear. The intermingling of Colombian and Spanish sequences suggests recent links between the countries. The emigration increased substantially in the 90s until Colombians became the second most important group of Latin American foreigners in Spain in the 2000s [25, 55] and they still remain as an important group of immigrants [57]. Similarly, immigration or tourism to Colombia is mainly from Venezuela, United States, Ecuador and Spain [25, 57]. It is important to note, however, that 72.7% (72/99) of the Spanish sequences included in the BLASTn-COL dataset were isolated from patients originated in Spain from the prospective pan-European SPREAD study [58], whereas only one individual was originated in Colombia and 8 in other Latin American coutries (9.1%) [Vercauteren J. and Struck D., personal communication]. Further studies are necessary to unravel the relation between these countries.

As expected, we showed Bogotá, Cali, and Medellin were the main cities involved in the viral spread in the country because these cities also had the higher number of people living with HIV in 2007 [53]. Although Bogotá was the main sink of the epidemic within an international context, it was also the main exporter of the epidemic within Colombia, whereas Cali and Medellin had a limited role as exporters. These findings differ from the high internal mobility due to violence or socio-economical reasons and internal displacement that happened in the years of this study [24, 25], suggesting that tourism from Bogotá to other Colombian cities could be an important driver of the epidemic [55]. Interestingly, six monophyletic clusters were identified with mainly sequences from Bogotá intermixed with some sequences from other cities or regions, that suggests potential transmission networks among those regions.

It is important to bear in mind the possible limitations of this study. Colombian sequences retrieved from the Center that performs nationwide resistance testing lack epidemiological information that could further shed light on the HIV-1 epidemic [16]. It is however known that heterosexual transmission took the greatest share of infections in the country (57.8%) followed by men who have sex with men or bisexuals (38.4 %) among 33853 cases with known risk transmission until 2007 [53]. Our sample was also not fully representative of some Colombian regions but we diminished this bias by doing analyses

on three weighted datasets and obtained similar results. Other limitation was the use of *pol* sequences rather than full genome sequences, which limited the detection of recombinants.

In conclusion, the HIV epidemic in Colombia is dominated by subtype B over time. Furthermore, we show that multiple introductions happened from the late-1970s onwards, and are still happening today, most likely from Spain. Viral populations in Bogotá, Medellin and Cali play a central role in shaping the HIV-1 epidemic within Colombia. This study provides a framework for future phylodynamic studies targeted at evaluating the dynamics of HIV-1 epidemic at a country-scale, which may provide information for prevention policies.

REFERENCES

[1] UNAIDS-WHO. GLOBAL REPORT: UNAIDS report on the global AIDS epidemic. Geneva; 2013.

[2] Sanchez GI, Bautista CT, Eyzaguirre L, Carrion G, Arias S, Sateren WB, et al. Molecular epidemiology of human immunodeficiency virus-infected individuals in Medellin, Colombia. The American journal of tropical medicine and hygiene. 2006 Apr;74(4):674-7.

[3] Eyzaguirre L, Bautista CT, Ayala C, Acosta J, Negrete M, Sateren WB, et al. First case of HIV Type 1 subtype F among men who have sex with men in Colombia. AIDS research and human retroviruses. 2006 Aug;22(8):808-11.

[4] Montano SM, Sanchez JL, Laguna-Torres A, Cuchi P, Avila MM, Weissenbacher M, et al. Prevalences, genotypes, and risk factors for HIV transmission in South America. Journal of acquired immune deficiency syndromes (1999). 2005 Sep 1;40(1):57-64.

[5] Villarreal JL, Gutierrez J, Palacio L, Penuela M, Hernandez R, Lemay G, et al. Characterization of HIV type 1 envelope sequence among viral isolates circulating in the northern region of Colombia, South America. AIDS research and human retroviruses. 2012 Dec;28(12):1779-83.

[6] Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS (London, England). 2011 Mar 13;25(5):679-89.

[7] Hierholzer J, Montano S, Hoelscher M, Negrete M, Hierholzer M, Avila MM, et al. Molecular Epidemiology of HIV Type 1 in Ecuador, Peru, Bolivia, Uruguay, and Argentina. AIDS research and human retroviruses. 2002 Dec 10;18(18):1339-50.

[8] Alencar CS, Sabino EC, Carvalho SM, Leao SC, Carneiro-Proietti AB, Capuani L, et al. HIV genotypes and primary drug resistance among HIV-seropositive blood donors in Brazil: role of infected blood donors as sentinel populations for molecular surveillance of HIV. J Acquir Immune Defic Syndr. 2013 Jul 1;63(3):387-92.

[9] Pimentel VF, Morgado MG, Bello Bentancor GJ, Guimaraes MD, Castilho E, Veloso VG, et al. Temporal Trends and Molecular Epidemiology of HIV-1 Infection in Rio de Janeiro, Brazil. AIDS research and human retroviruses. 2013 Aug 30.

[10] Bello G, Zanotto PM, Iamarino A, Graf T, Pinto AR, Couto-Fernandez JC, et al. Phylogeographic analysis of HIV-1 subtype C dissemination in Southern Brazil. PloS one. 2012;7(4):e35649.

[11] de Oliveira T, Pillay, D, Gifford, R. J., for the UK Collaborative Group of HIV Drug Resistance. The HIV-1 Subtype C Epidemic in South America Is Linked to the United Kingdom. PloS one. 2010;5(2):e9311.

[12] Fontella R, Soares, M. A., Schrago, C. G. On the origin of HIV-1 subtype C in South America. Aids. 2008;22:2001-11.

[13] Castro E, Moreno M, Deibis L, de Perez G, Salmen S, Berrueta L. Trends of HIV-1 molecular epidemiology in Venezuela: introduction of subtype C and identification of a novel B/C mosaic genome. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology. 2005 Mar;32(3):257-8.

[14] Ahumada-Ruiz S, Casado C, Toala-Gonzalez I, Flores-Figueroa D, Rodriguez-French A, Lopez-Galindez C. High divergence within the major HIV type 1 subtype B epidemic in Panama. AIDS research and human retroviruses. 2008 Nov;24(11):1461-6.

[15] Faria NR, Suchard MA, Rambaut A, Lemey P. Toward a quantitative understanding of viral phylogeography. Current opinion in virology. 2011 Nov;1(5):423-9.

[16] Holmes EC. Evolutionary history and phylogeography of human viruses. Annual review of microbiology. 2008;62:307-28.

[17] Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. Nature reviews Genetics. 2009 Aug;10(8):540-50.

[18] Korber B, Muldoon M, Theiler J, Gao F, Gupta R, Lapedes A, et al. Timing the ancestor of the HIV-1 pandemic strains. Science. 2000 Jun 9;288(5472):1789-96.

[19] Salemi M, Strimmer K, Hall WW, Duffy M, Delaporte E, Mboup S, et al. Dating the common ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes using a new method to uncover clock-like molecular evolution. Faseb J. 2001 Feb;15(2):276-8.

[20] Gilbert MT, Rambaut A, Wlasiuk G, Spira TJ, Pitchenik AE, Worobey M. The emergence of HIV/AIDS in the Americas and beyond. Proceedings of the National Academy of Sciences of the United States of America. 2007 Nov 20;104(47):18566-70.

[21] Robbins KE, Lemey P, Pybus OG, Jaffe HW, Youngpairoj AS, Brown TM, et al. U.S. Human immunodeficiency virus type 1 epidemic: date of origin, population history, and characterization of early strains. Journal of virology. 2003 Jun;77(11):6359-66.

[22] Murillo W, Veras N, Prosperi M, de Rivera IL, Paz-Bailey G, Morales-Miranda S, et al. A Single Early Introduction of HIV-1 Subtype B into Central America Accounts for Most Current Cases. Journal of virology. 2013 Jul;87(13):7463-70.

[23] Junqueira DM, de Medeiros RM, Matte MC, Araujo LA, Chies JA, Ashton-Prolla P, et al. Reviewing the history of HIV-1: spread of subtype B in the Americas. PLoS One. 2011;6(11):e27489.

[24] IDMC. Colombia: Rate of new displacement highest in two decades. Geneva, Switzerland: IDMC- Internal Displacement Monitoring Centre; 2008 17 October.

[25] OIM. Perfil Migratorio de Colombia. Bogotá, Colombia; 2010.

[26] Rhee SY, Kantor R, Katzenstein DA, Camacho R, Morris L, Sirivichayakul S, et al. HIV-1 pol mutation frequency by subtype and treatment experience: extension of the HIVseq program to seven non-B subtypes. AIDS (London, England). 2006 Mar 21;20(5):643-51.

[27] Pineda-Pena AC, Faria NR, Imbrechts S, Libin P, Abecasis AB, Deforche K, 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. Infect Genet Evol. 2013 May 7.

[28] Struck D, Perez-Bercoff D, Devaux C, Schmit JC. COMET: A Novel approach to HIV-1 subtype prediction. 8th European HIV Drug Resistance Workshop; 2010 17- 19 March; Sorrento, Italy; 2010.

[29] Schmidt HA, Strimmer K, Vingron M, von Haeseler A. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics (Oxford, England). 2002 Mar;18(3):502-4.

[30] Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. Journal of virology. 1999 Jan;73(1):152-60.
[31] Bruen TC, Philippe H, Bryant D. A simple and robust statistical test for detecting the presence of recombination. Genetics. 2006 Apr;172(4):2665-81.
[32] Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. Molecular biology and evolution. 2006 Feb;23(2):254-67.
[33] Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. Automated phylogenetic detection of recombination using a genetic algorithm. Molecular biology and evolution. 2006 Oct;23(10):1891-901.
[34] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research. 2004;32(5):1792-7.
[35] Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular biology and evolution. 2010 Feb;27(2):221-4.
[36] Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS One. 2009;4(3):e4724.
[37] Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One. 2010;5(3):e9490.
[38] Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Systematic biology. 2006 Aug;55(4):539-52.
[39] Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution. 2012 Aug;29(8):1969-73.
[40] Yang Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. Journal of molecular evolution. 1994 Sep;39(3):306-14.
[41] Drummond AJ, Ho SY, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. PLoS biology. 2006 May;4(5):e88.
[42] Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular biology and evolution. 2005 May;22(5):1185-92.
[43] Suchard MA, Rambaut A. Many-core algorithms for statistical phylogenetics. Bioinformatics (Oxford, England). 2009 Jun 1;25(11):1370-6.
[44] Ayres DL, Darling A, Zwickl DJ, Beerli P, Holder MT, Lewis PO, et al. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. Systematic biology. 2012 Jan;61(1):170-3.
[45] Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLoS computational biology. 2009 Sep;5(9):e1000520.
[46] Edwards CJ, Suchard MA, Lemey P, Welch JJ, Barnes I, Fulton TL, et al. Ancient hybridization and an Irish origin for the modern polar bear matriline. Curr Biol. 2011 Aug 9;21(15):1251-8.
[47] Bielejec F, Rambaut A, Suchard MA, Lemey P. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics (Oxford, England). 2011 Oct 15;27(20):2910-2.
[48] Minin VN, Bloomquist EW, Suchard MA. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Molecular biology and evolution. 2008 Jul;25(7):1459-71.
[49] Talbi C, Lemey P, Suchard MA, Abdelatif E, Elharrak M, Nourlil J, et al. Phylodynamics and human-mediated dispersal of a zoonotic virus. PLoS pathogens. 2010;6(10):e1001166.
[50] O'Brien JD, Minin VN, Suchard MA. Learning to count: robust estimates for labeled distances between molecular sequences. Molecular biology and evolution. 2009 Apr;26(4):801-14.
[51] Paraskevis D, Pybus O, Magiorkinis G, Hatzakis A, Wensing AM, van de Vijver DA, et al. Tracing the HIV-1 subtype B mobility in Europe: a phylogeographic approach. Retrovirology. 2009;6:49.
[52] Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics (Oxford, England). 2006 Nov 1;22(21):2688-90.
[53] Ministerio de la Protección Social. Informe Nacional sobre los progresos realizados en la aplicación del UNGASS. Bogotá, Colombia; 2008.
[54] Faria NR, Hodges-Mameletzis I, Silva JC, Rodes B, Erasmus S, Paolucci S, et al. The phylogeographic footprint of colonial history in the global dispersal of HIV-2 group A. The Journal of general virology. 2012 Jan 4.
[55] Departamento Administrativo Nacional de Estadistica (DANE). Analisis de los Resultados de la Encuesta Turismo a Hogares. Bogota, Colombia: Departamento Administrativo Nacional de Estadistica (DANE) and Ministerio de Comercio Industria y Turismo,; 2003.
[56] Instituto Nacional de Estadística (INe). Anuario Estadístico de España de 1971 a 1980 y de 1981 a 1990. In: Instituto Nacional de Estadística (INe), ed. *Totales Nacionales: Demografía-Turismo*: available at www.ine.es 1970-1990.
[57] OIM. Perfil Migratorio de Colombia. Bogotá, Colombia; 2012.
[58] Vercauteren J, Wensing AM, van de Vijver DA, Albert J, Balotta C, Hamouda O, et al. Transmission of drug-resistant HIV-1 is stabilizing in Europe. The Journal of infectious diseases. 2009 Nov 15;200(10):1503-8.

CHAPTER 4

HIV-1 TRANSMITTED DRUG RESISTANCE IN LATIN AMERICA AND CARIBBEAN: WHAT DO WE KNOW?

Published in:

Andrea-Clemencia Pineda-Peña, Diana-Carolina Bello, Otto Sussmann, Anne-Mieke Vandamme, Jurgen Vercauteren, Kristel Van Laethem, Arley Gómez-López. HIV-1 Transmitted Drug Resistance in Latin America and Caribbean: What do we know? AIDS Reviews. 2012; 14 (4): 256–67

4.1 SUMMARY

Background: Latin America and the Caribbean countries have increased the scaling-up of antiretroviral treatment in the last years. The increase of transmitted drug resistance has been feared due to the worrisome indicators associated with the emergence of drug resistance and monitored by the WHO. Consequently, our aim was to review all relevant studies on transmitted drug resistance in Latin America and the Caribbean countries, to analyze its levels, to identify the frequency of transmitted drug resistance mutations, and to put these results in the context of the local Latin American and Caribbean countries settings.

Methodology: A systematic search of Spanish, Portuguese, and English literature was performed in databases and international conferences for the period June 1999 to May 2011. In addition, sequences were downloaded from the Los Alamos and Stanford databases and the transmitted drug resistance was reanalyzed according to the WHO Surveillance Drug Resistance Mutation list 2009.

Results and Conclusions: In total, 50 articles, 27 abstracts, and 1,922 patients were included. The resistance varied geographically, but most of the countries have reached the WHO threshold of 5% of resistance. According to the sequences available in public databases, the overall prevalence in Latin America and the Caribbean countries for the period 1996-2009 was 7.7% and by region it was 4.3% for the Caribbean, 3.9% for Mexico, 9.4% for Brazil, 10.5% for the Andean region and 4.9% for the Southern Cone. For the last four investigated years (2006-2009), the information was restricted to Brazilian and Venezuelan studies and revealed an overall transmitted drug resistance of 10%. Throughout the study period, limited information was available for the Caribbean and Central American countries. These findings support the need for developing comprehensive surveys of transmitted drug resistance in these regions.

4.2 INTRODUCTION

The horizontal transmission of HIV-1 strain that already displayed the RT mutation T215F/Y mutation was first described in a homosexual man in 1993 [1]. As it was feared that TDR could negatively impact the success of first-line ART, TDR was extensively studied in resource-rich countries. Retrospective studies provided evidence that infection with drug-resistant HIV-1 could also occur in other risk groups and indicated increasing TDR levels and more extensive mutational patterns. As these studies were often small and based upon convenience sampling, prospective and representative population-based studies were set up in Europe, USA and Canada to monitor the extent and dynamics of TDR [2-5]. The

high rates of TDR (overall prevalence 8-9%) resulted into the recommendation of standard genotypic drug resistance testing to guide ART selection in therapy-naïve patients [6, 7]. Recently, a large European multi-cohort study confirmed the importance of baseline drug resistance testing as patients with TDR and resistance to at least one prescribed drug were three times as likely to virologically fail within one year [8].

As access to ART has been implemented and scaled-up in Latin-America and Caribbean (LAC) since more than a decade, concerns towards the emergence and spread of TDR strains in LAC exist. In general, the prevalence of TDR depends mainly on the time since ART implementation, the scale of ART coverage and the overall virological success of the prescribed regimens. As the majority of HIV-1 patients in LAC started immediately on HAART, the rise in TDR was anticipated to be less extensive compared to resource-rich countries where sequential mono- and bitherapies were initially prescribed. Nevertheless, other factors unique to the particular setting of LAC could negatively influence TDR rates. For instance, the wide-spread use of standard first-line regimens based on low genetic barrier NNRTI in the absence of drug resistance testing, together with limited access to second-line therapies, could promote emergence of drug resistance and its subsequent spread to new HIV-1 infections.LAC is a heterogeneous region that differs extensively in geographical, social and economic characteristics that influence the HIV epidemiology and clinical care facilities. The HIV-1 burden varies considerably between and within the 27 countries belonging to LAC (Figure 4.1). In 2009, 1.7 million people were living with HIV-1 in LAC, of which one third lived in Brazil, and HIV-1 prevalence values ranged between 0.1% in Cuba and 3.1% in Bahamas [9]. Brazil and Argentina took a leading role in the area by providing ART since 1991-1992 and by setting up a national program to provide free access to HAART, clinical care and laboratory monitoring (Figure 4.2) [10, 11]. Currently, Cuba, Chile, and Nicaragua achieve global access, and Argentina, Brazil, Dominican Republic, Mexico and Uruguay are near to universal-coverage levels. However, in other LAC countries ART scale-up has been rather slow and only reached ART coverage between 20% in Bolivia and 66% in Paraguay with still limited access to laboratory testing and second-line regimens [12].

The scope of this review was to collect all relevant studies on TDR in LAC, to analyze TDR levels, to identify the frequency of TDR mutations and to put these results in the context of the local LAC settings.

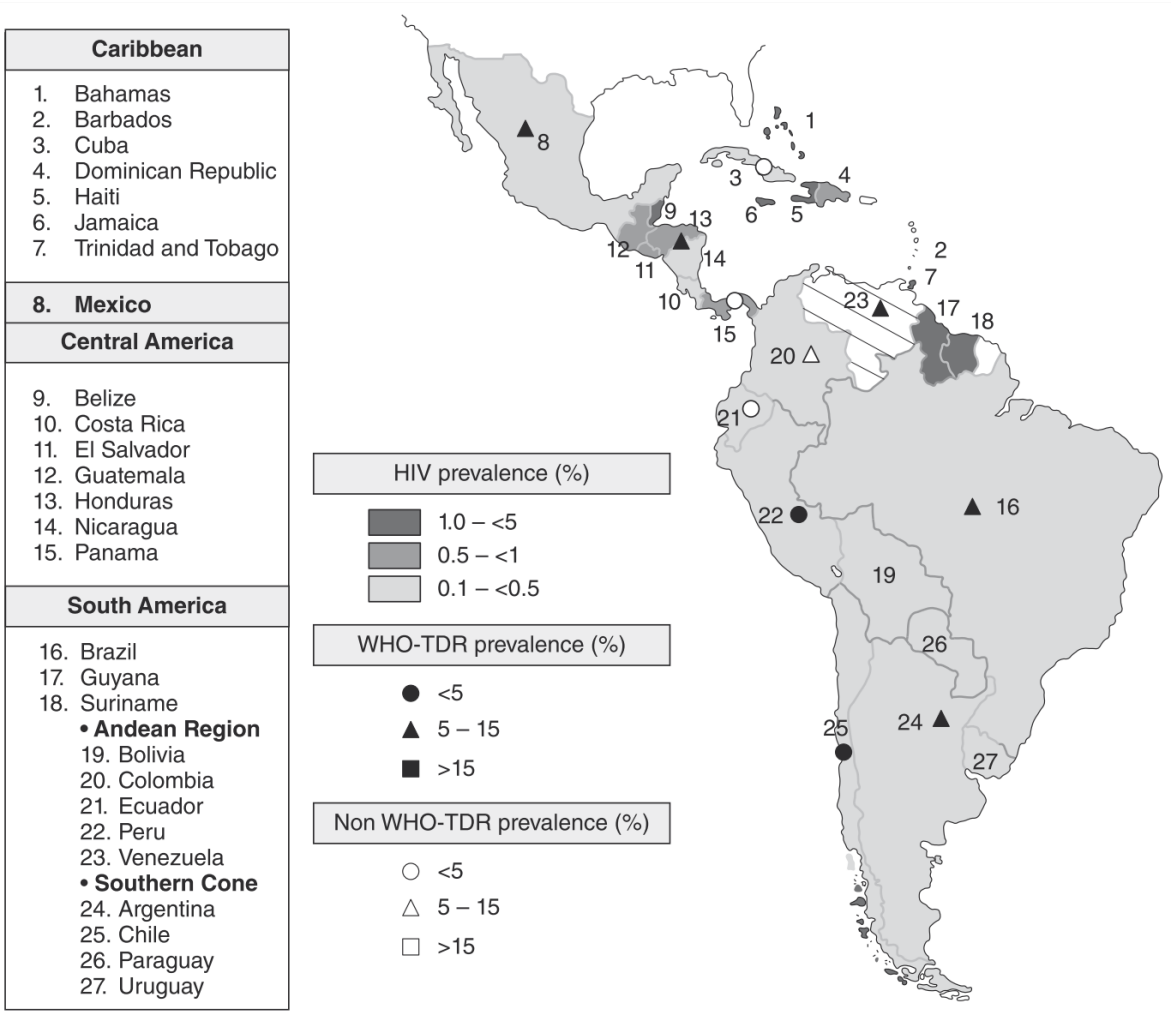


FIGURE 4. 1. HIV AND TDR PREVALENCE IN LATIN AMERICA AND CARIBBEAN

Figure 4.1: HIV and TDR prevalence in Latin America and Caribbean. This region is divided geographically in Caribbean (7 countries), Mexico, Central America (7 countries) and South America (12 countries). The latter region is subdivided based upon geographical, social and economic characteristics: Andean region and Southern Cone. The HIV prevalence according to the UNAIDS report is shown in a grayscale [9]. Although, the overall HIV prevalence in Venezuela (hatched) was 0.75% in 2004, the country was not visualized in a grayscale as the prevalence in 2007 varied between 1.62 % in metropolitan areas and 0.08% in more rural areas [13]. TDR prevalence values from studies that followed the eligibility criteria of WHO surveys are represented in black figures, otherwise they are represented in white. Adapted from UNAIDS report, 2010.

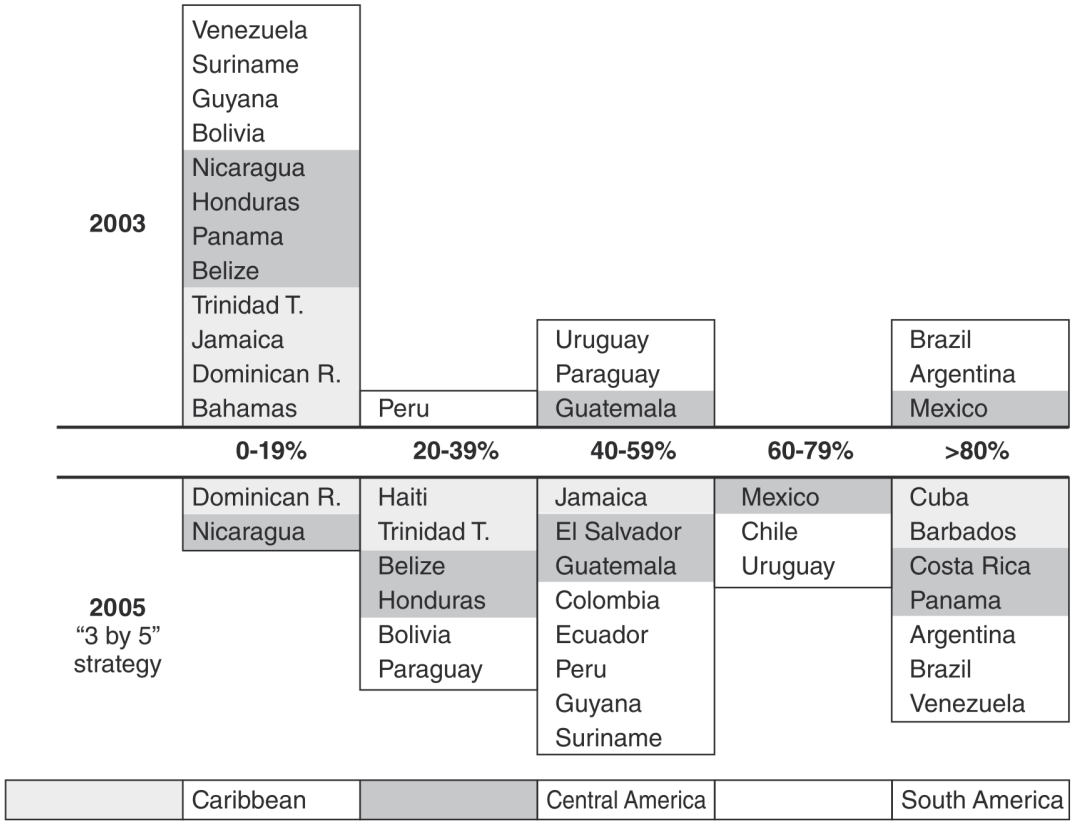


FIGURE 4. 2. SCALING-UP OF ART

Figure 4.2: Scaling-up of ART. The scaling-up of ART in 2003 and after the "3 by 5" strategy launched by WHO and UNAIDS. "3 by 5" stands for treatment for three million HIV/AIDS patients in low-middle income countries by 2005 (see details in the UNAIDS report "3 by 5" and beyond) [14]. The regions are shown in a grayscale. Progress in access to treatment is shown as a shift to the right for a country, comparing 2003 (upper panel) with 2005 (lower panel). Countries are scaled as percentage (%) of those diagnosed and eligible for treatment that actually received treatment. The scale is shown in the middle panel. Data were not available for Barbados, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador and Haiti in 2003. Adapted from UNAIDS[14]. Abbreviations: Trinidad T: Trinidad and Tobago, Dominican R: Dominican Republic.

4.3 MATERIALS AND METHODS

4.3.1 SYSTEMATIC SEARCHING

A systematic strategy (Figure 4.3) was used to search for Spanish, Portuguese and English literature published in the period between June 1999 and May 2011 and available within the following databases: MEDLINE via PubMed, SciELO, BIREME, IBECS, Cochrane, and LILACS. Additionally, abstracts from Conference on Retroviruses and Opportunistic Infections (CROI), International AIDS Conference, IAS

Conference on HIV Pathogenesis and Treatment, Meetings of the Infectious Diseases Society of America, European AIDS Clinical Society, International Congress on Drug Therapy in HIV Infection, and the International Drug Resistance Workshop were identified. The combinations of the following MeSH (Medline descriptor) or DeCS (Lilacs descriptor) terms were used: "HIV" or "HIV-1" and “Drug Resistance" or "Anti-HIV Agents" or “Anti-Retroviral Agents” and "Latin America", "Caribbean", "Central America", "South America", or each individual LAC country (Figure 4.1). For the respective international conferences "HIV" and “Resistance" and the name of the LAC country were used as search items. The data were independently extracted and reviewed by two researchers (Pineda AC, Bello DC). Only studies in which naïve patients were clearly defined as having no exposure to any ARV were included. Duplicates or preliminary versions of studies with similar data were excluded. Resistance surveys in pregnant women and in children and studies investigating MTCT were also excluded due to the high probability of including treated patients. Studies based upon point mutation assays (e.g. Line Probe Assay LiPA) were not included as the data collection was restricted to specific mutations and sensitivity and specificity characteristics of these assays substantially differ from sequencing technology.

The following data were extracted from the articles: region, sample size, data collection period, type of infection, mode of transmission, gender, age, CDC classification, CD4 count, subtype, TDR, technique and algorithm to determine drug resistance.

4.3.2 ANALYSIS OF RESISTANCE

TDR was reanalyzed according to the SDRM list 2009 [15]. The sequences from the publications that were found through the systematic search were downloaded from Los Alamos database (www.hiv.lanl.gov accessed in July 2011) and Stanford database (http://hivdb.stanford.edu/index.html accessed in August 2011) [13]. The RegaDB sequence alignment tool (available http://jose.med.kuleuven.be/) was used to determine amino acid mutations. If sequences were not available, TDR was recalculated based upon the mutations reported in the original figures and tables and excluding mutations that were not part of the SDRM list 2009.

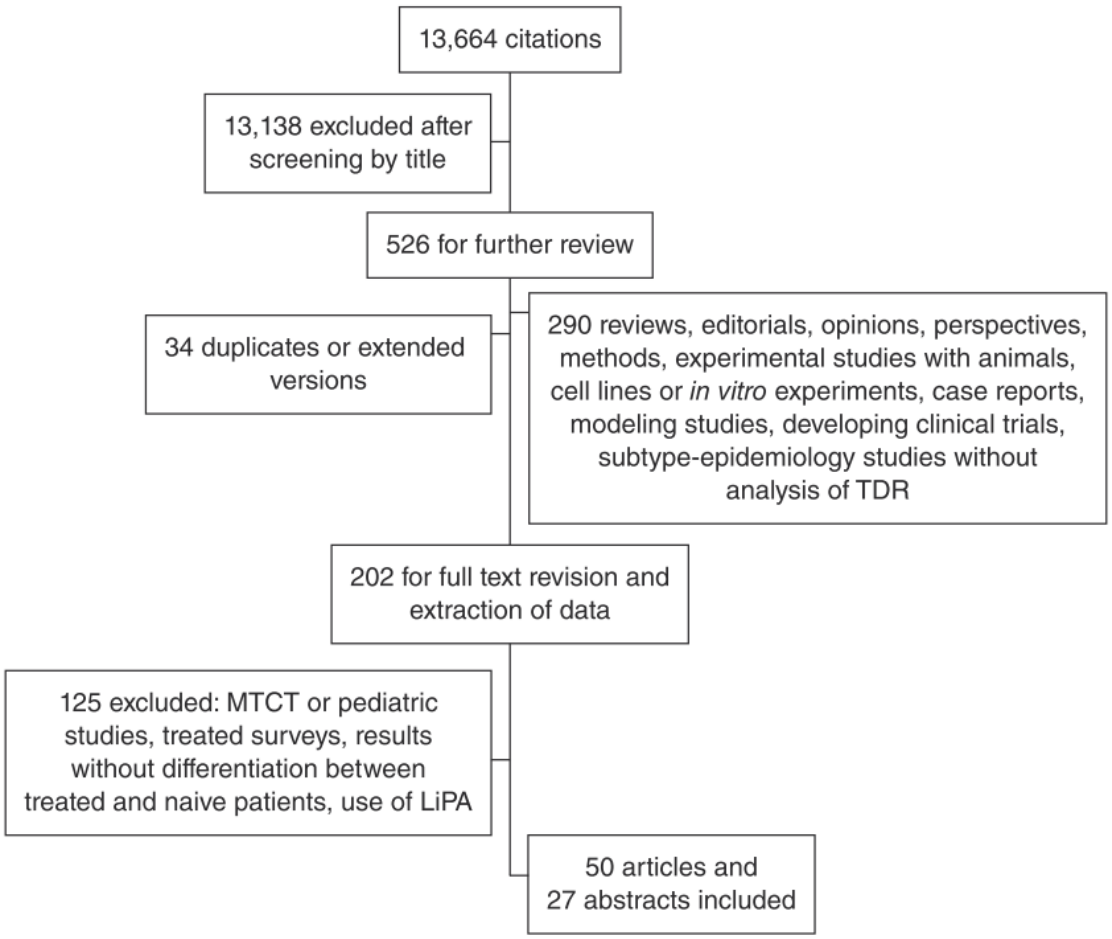


FIGURE 4. 3. SELECTION OF STUDIES

Figure 4.3: Selection of studies. Abbreviations: TDR: transmitted drug resistance, MTCT: Mother-to-child transmission, LiPA: Line Probe Assay.

Both databases were also searched for additional HIV-1 sequences from Latin America and Caribbean. The inclusion criteria were therapy naïve patient (according to database and article referenced) and availability of RT and PR sequences for each patient. The exclusion criteria were duplicated sequences, clones, unknown treatment history, evidence of hypermutation and short sequences (minimum requirements: RT positions 41 to 219 positions and PR positions 30 to 90). Dubious data were not extracted. The viral subtypes were evaluated by REGA HIV-1 subtyping tool [16].

The analysis of the data was done with the statistical software R version 2.13.1. Prevalence values were calculated with a 95% Wilson score confidence interval on the basis of a binomial distribution. Categorical data were compared with the chi-square test. Logistic regression analysis was used to examine TDR time trends in LAC, Brazil and Argentina, likewise the trends of NRTI, NNRTI and PI resistance and of 41L and 103N mutations.

4.4 RESULTS

4.4.1 TDR SURVEILLANCE AND EARLY WARNING INDICATORS (EWIs)

Recently, the fear for TDR in developing countries reappeared due to the worrying reports of EWIs in Africa, Caribbean and Central America [17-20]. EWIs were developed by WHO as indicators for the prevention of drug resistance. They include ART prescribing practices, lost-to-follow-up, retention of first-line therapies, viral suppression and percentage of discontinued drug supplies [17, 20]. These indicators alert earlier about the potential of increasing TDR rates than prospective surveys.

Few studies have evaluated EWIs in LAC. Most of the studies were restricted to Brazil or to the seven AIDS programs from the CCASAnet Cohort [21, 22]. Lost-to-follow-up was 6% overall, but large differences were observed between the different AIDS health care institutions (between 0.6% in Honduras to 17% in Argentina) [21]. In El Salvador, Guatemala, Honduras and Nicaragua only 7 of 13 sites accomplished the goal of ≤20% lost-to-follow-up and 38 of 55 sites in the Caribbean [17, 18]. Low numbers of retention in first-line therapy were also observed in LAC, ranging from 16% in Honduras to 36% in Peru. Adverse events were the main cause in 14% of the cases, predominantly linked to zidovudine (hematological toxicity in 7%) and nevirapine (skin rash in 3%) [21, 22].

To standardize TDR surveillance studies, the WHO has developed minimum recommendations for developing countries to enable comparison. Briefly, the criteria are a minimum of 47 samples from small geographic areas within countries where ART has been available for more than 3 years to at least 20% of the eligible HIV-1 population, HIV-1 infection should have been confirmed by laboratory criteria, participants should not have received any antiviral drug, and if female they should have been younger than 25 years at their first pregnancy. The inclusion of individuals with less than 3 years of HIV-1 infection or/and laboratory evidence of seroconversion or recent infection are required. If the information is accessible, demographical data like transmission risk group, clinical data like non-AIDS stage, or laboratory data like CD4 > 500 cells/ml, no previous positive HIV-1 test are desirable [23, 24]. In addition, the analysis of resistance should be done with the SDRM list, which is updated periodically and excludes highly polymorphic positions [15]. Nevertheless, some SDRM do occur naturally in therapy-naive individuals, which might inflate the estimates of TDR. The working group decided to retain these mutations within the list as stricter inclusion criteria would results into the omission of important drug resistance mutations and these particular mutations might have been the result of unreported therapy-exposure or truly TDR within the therapy-naive dataset [15].

4.4.2 EPIDEMIOLOGY OF TDR IN LATIN AMERICA AND CARIBBEAN

Due to the low achievement of EWIs in several LAC countries, available publications on TDR were reviewed using a systematic search strategy (Figure 4.3). Fifty articles and 27 abstracts were retained (Supplementary Tables 4.1, 4.2 and 4.3). Approximately half of the studies originated from Brazil (53.2%, 41/77). Ten studies were from Mexico (13%, 10/77), three from Caribbean (3.9%, 3/77), four from Central America (5.2%, 4/77), twelve from the Southern Cone (15.6%, 12/77) and seven (9.1%, 7/77) from the Andean region.

Heterosexual contacts and MSM were the most frequent transmission risk factors. Six studies were performed in specific patient populations, such as blood donors [25-27], female sex workers [28], intravenous drug users [29] and a combination of the latter two [30].

Eleven studies used the Serological Testing Algorithm for Recent HIV Seroconversion (STARHS) [31] strategy to differentiate between recent or chronic infection [27, 32-39]. In five other studies the distinction between recent and chronic infection was based upon previous negative serology test results 9-18 months before diagnosis date [40-44].

Although plasma was widely used as source material, Peripheral Blood Mononuclear Cells (PBMC) and Dried Blood Samples (DBS) were also used in 9 and 2 studies, respectively. In most of the studies *in-house* methodologies were used to generate sequences, followed by Trugene® (seven) and ViroSeq® (six) (Supplementary Table 4.1). The genotypic resistance interpretation algorithm HIVdb was widely used for TDR analysis. Twenty-one studies used the International AIDS Society list and 19 studies the SDRM list. The version of the respective algorithms and lists were often not mentioned in the original publication.

The TDR levels as reported within the original publications are displayed within supplementary tables 4.2 and 4.3. In summary, the prevalence of TDR ranged from 0 to 7.4% in the Caribbean for 1996-2003, from 2.8 to 18% in Mexico for 2001-2010, from 0 to 11.6% in Central America for 2002-2007, from 0 to 41% in Brazil for 1994-2010, from 3 to 11% in the Andean Region for 1998-2007 and from 2.5 to 18.8% in the Southern Cone for 1997-2009.

Due to the heterogeneity of mutation lists used within the original studies, TDR was reanalyzed using SDRM list 2009 and resulted into TDR prevalence values between 0 and 8.3% (Caribbean), 2.8 and

6.8% (Mexico), 0 and 7.5% (Central America), 1.3 and 20.3% (Brazil), 0 and 13% (Andean region), and from 1.6 to 12% (Southern Cone). For comparison, values reported in 14 studies that followed the WHO surveillance criteria for TDR monitoring were also displayed in supplementary tables 4.1 and 4.2 [25, 27, 32, 34-36, 38, 39, 41, 42, 44-49].

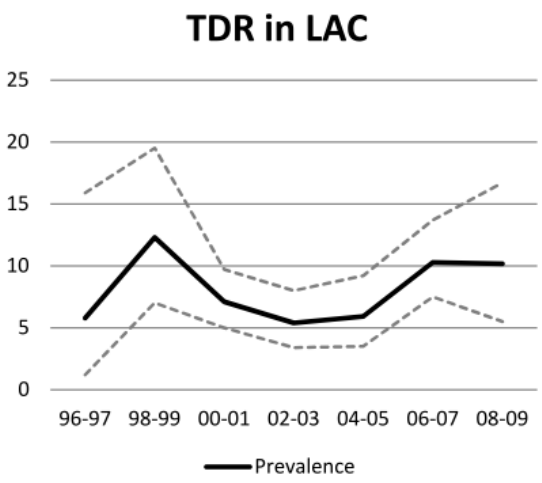
For ten studies, the presence of TDR in each sample could be directly linked to the respective transmission risk factor [28, 34-36, 40, 44, 47, 50-52] and combining this data did not reveal any difference in TDR rate between heterosexuals (37/522, 7.1%) and MSM (39/757, 5.2%). One study that solely sampled in blood donors from São Paulo (Brazil) showed a TDR of 5.6% in 1998-2002 (19/341) [25], whereas another study sampling in a broader geographical area revealed only a TDR of 1.3% in 2000-2004 (1/74)[27]. The majority of patients included in studies where differentiation between recent and chronic infection was made, were chronically infected and therefore differences in TDR between both groups could not be investigated. However, two studies did report a difference [25, 47]. For three studies that included at least 50 recent infections, TDR values were 4% (3/74) in Chile [41], 7.7% (4/52) in Argentina [44], and 8.1% (17/210) in Brazil [42].

To obtain a better understanding of the reported TDR rates, public databases were searched and 1922 relevant HIV-1 sequences were retained for analysis (Figure 4.4A). Based upon automated subtyping [16], HIV-1 subtype B was the most frequent subtype (65.7%, 1262/1922), followed by subtype C (14.4%, 276/1922), BF recombinants (8.7%, 168/1922) and subtype F1 (4%, 77/1922) The latter two mainly originated from Brazil and Argentina (BF: 99.4%, 167/168; F1: 98.7%, 76/77). The remaining genetic variants were CRFs such as CRF31_BC and CRF29_BF in Brazil, CRF_18cpx and CRF19_cpx in Cuba and BD, BC and CF1 recombinants in Brazil and the Southern Cone.

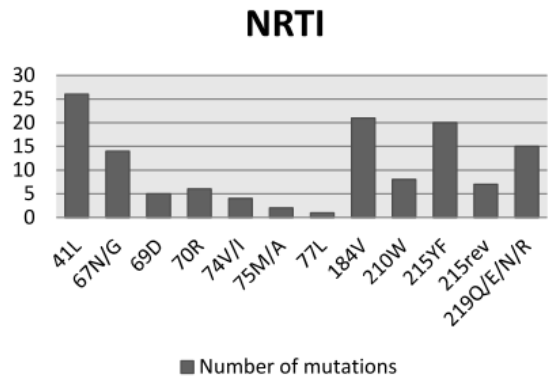
A

Country	Year							Total
	96-97	98-99	00-01	02-03	04-05	06-07	08-09	
BR	-	10/66	23/244	6/68	14/213	33/337	13/128	99/1056
AR	1/28	-	1/47	13/274	4/59	-	-	19/408
VE	-	2/31	-	0/12	-	8/62	-	10/105
CH	-	-	5/79	-	-	-	-	5/79
TT	-	-	1/61	0/2	-	-	-	1/63
MX	-	-	-	2/40	0/11	-	-	2/51
CO	-	-	5/37	-	-	-	-	5/37
CU	-	3/25	-	-	-	-	-	3/25
BB	2/24	-	-	-	-	-	-	2/24
DO	-	-	-	1/12	0/5	-	-	1/17
HT	-	-	-	-	0/16	-	-	0/16
Other	-	-	1/40	-	0/1	-	-	1/41
Total	3/52	15/122	36/508	22/408	18/305	41/399	13/128	148/1922

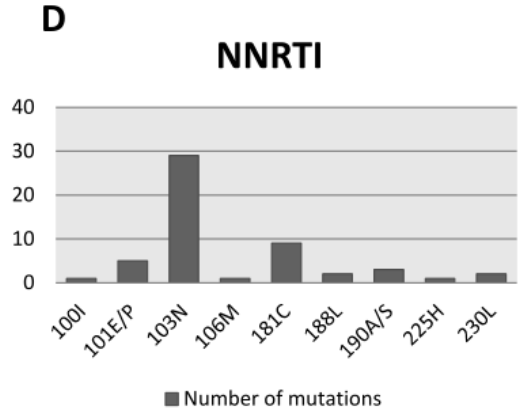
B



C



D



E

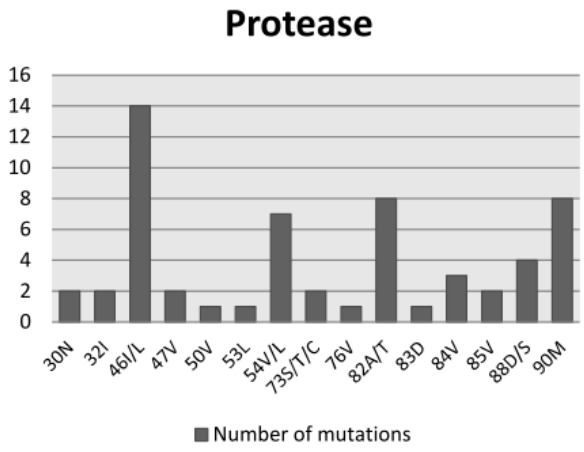


FIGURE 4. 4. PREVALENCE OF TDR AND MUTATIONS IN LAC

Figure 4.4: Prevalence of TDR and mutations in LAC. A) TDR prevalence in LAC. The number of sequences with evidence of TDR (numerator) and the total number of sequences included (denominator) from Los Alamos and Stanford database divided according to country and sampling

period B) and TDR prevalence (%). B) The prevalence of TDR according to sampling period (black line) and the 95% Confidence Intervals (gray dashes) C) The number of sequences with SDRM mutations against Nucleoside Reverse Transcriptase Inhibitors (NRTI), D) against Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTI) and E) against Protease Inhibitors (PI). Abbreviations: BR: Brazil, AR: Argentina, VE: Venezuela, CH: Chile, TT: Trinidad and Tobago, MX: Mexico, CO: Colombia, CU: Cuba, BB: Barbados, DO: Dominican Republic, HT: Haiti, Other: Include other Caribbean islands.

In LAC overall TDR prevalence was 7.7% (148/1922, 95% CI: 6.54-8.98) for the period from 1996 to 2009. The TDR by region was 4.3% in Caribbean for 1996-2005 (8/186, 95% CI: 1.87-8.29), 3.9% in Mexico for 2002-2005 (2/51, 95% CI: 1.87-8.29), 9.4% in Brazil for 1998-2009 (99/1056, 95% CI: 7.68-11.3), 10.5% in the Andean Region for 1999-2007 (15/142, 95% CI: 6.03-16.82) and 4.9% in the Southern Cone for 1997-2005 (24/487, 95%CI: 3.2-7.24). NRTI resistance was 4.4% (84/1922, 95% CI: 3.54-5.38), NNRTI resistance was 2.3% (44/1922, 95% CI: 1.71-3.06] and PI resistance was 2.08% (40/1922, 95% CI: 1.53-2.82]. 5.4% (8/148) and 6.8% (10/148) of the TDR cases had dual class resistance towards NRTI+NNRTI and NRTI+PI, respectively. Triple class resistance was only detected in 0.7% (1/148) of the TDR cases. The overall TDR time trend was not significant, although two TDR peaks above 9% were observed for the periods 1998-1999 and 2006-2007 (Figure 4.4B). The separate time trend analysis for the countries that contributed the majority of samples (Argentina and Brazil) were also not significant, neither the trends for the individual drug classes (NRTI, NNRTI and PI).

The 103N mutation was the most frequently observed mutation (29/148, 19.6%) followed by 215Y/F/rev (27/148, 18.2%) and 41L (26/148, 17.6%) (Figure 4.4C- 4E). Time trends analysis for the individual mutations 41L and 103N did not reveal any significant changes. The frequency of the most prevalent TDR mutations such as 41L, 184V, 103N, 46I/L, 82A/T and 90M was compared with frequencies in the following studies in US, Canada and Europe [2, 3, 5], but no significant differences were observed.

For 447 sequences information was available on whether it concerned a recent or a chronic infection. However, the TDR rate did not differ significantly between recent (4/52, 7.7%) and chronic (41/395, 10.4%) infections. The prevalence of TDR did differ between sequences obtained from PBMC or plasma (1712 sequences included). TDR was significantly higher in PBMC than plasma (33/285, 11.6%, versus 98/1427, 6.8%). However, the exclusion of PBMC samples did not impact significantly the overall TDR (7%, 95% CI: 5.83-8.37). Other factors were not evaluated because of lack of information.

CARIBBEAN AND CENTRAL AMERICA

According to the UNAIDS-WHO report, only a minimum ART coverage was installed in the Caribbean Region in 2001 [14]. However, coverage increased with extra funding to e.g. 20% in Haiti and Dominican Republic in 2005 (Figure 4.2) [14]. As a result, TDR ranged from 0% in several islands in 2000 to 3.6% in Cuba in 2003 [53-56]. In Cuba, there was already limited access to mono- and bitherapy before 2000 which could explain the level of TDR observed in 2003 and its frequency of the NRTI drug class [54]. The overall TDR was estimated to be 4.3% and this was based upon the reanalysis using the WHO TDR surveillance list of the collected sequences that were available in public databases and that were sampled in the time period 2000-2005 [15].

From Central America only fragmented data was available for two countries of this region. Indeed, only 4 publications and no sequences were retained after our search in public databases. A Panamanian study showed no TDR mutations in 2005, but NRTI and NNRTI TDR increased in a later report [57, 58]. However, in the latter study the exact sampling years were unspecified. Reports that studied the TDR rate in Honduras showed a prevalence of 9.2% in 2002-2003 and a prevalence of 7% in 2004-2007 [47, 59]. In Honduras and Panama, TDR was mainly restricted to NRTI and NNRTI drug classes, explained by the first-line therapy options in these countries [47, 58].

MEXICO

Reanalysis of Mexican sequences revealed TDR prevalence of 3.9% in the period 2002-2005. This result is substantially lower than the 16% reported by Escoto *et al* [60], which might be partly explained by the algorithm used to interpret drug resistance as it also included polymorphisms. However, the TDR prevalence of 16% also differed from values obtained in other studies [30, 61], suggesting that the heterogeneity in ART availability between the different regions in Mexico might also impact TDR rates although further research is warranted [61]. A recent comprehensive WHO-survey determined that Mexico has reached 6.8% of TDR for the period 2005-2010 [48]. This overall TDR remained stable at the national level. However, increasing trends for NNRTI and PI TDR were observed. The frequency of mutations was 4.2% for NRTI (mainly type TAM I pathway), 1.9% for NNRTI (mostly 103N) and 1.8% for PI (90M) [48].

ANDEAN REGION

In the Andean Region the TDR prevalence was 10.5% according to sequences sampled in the years 1999, 2001, 2003 and 2007. However, these sequences were mainly collected in Venezuela [49, 62, 63]. A WHO-survey performed in Venezuela reported at least 7% TDR according to the ANRS

algorithm and this value raised to 12.9% when the SDRM list was used for interpretation [49]. In this study, drug resistance to all three drug classes was observed. The high TDR rates may be a consequence of the early introduction of ART as part of an AIDS National Program that was initiated in 2000 and was mainly focused in the capital where the surveys were performed. Additionally, one report mentioned that patients receiving ART and displaying poor adherence or engagement with the health care system were an important source of HIV-1 transmission and TDR in Venezuela [64].

In that era, free access to ART was also guaranteed in other countries belonging to that region (Figure 4.2). However, scaling-up moved slowly and e.g. in Colombia and Peru scale-up reached only respectively 34% and 57% in 2010 according to UNAIDS [12]. One Colombian survey showed a TDR of 5.8% according to HIVdb algorithm. The sequences from this study were not available for reanalysis with SDRM 2009, but other Colombian sequences that could be accessed did not reveal significant differences in TDR [65]. Two studies in Peru and Ecuador revealed low TDR rates of 3.3% and 4.3% in 2002-2003 [35, 66]. However, it is worrying that all drug classes were affected in these countries. It is unknown whether polymorphisms might have affected these values as data was unavailable for reassessment with SDRM 2009.

BRAZIL

The overall TDR rate in Brazil was investigated by reanalyzing 1056 sequences that covered the period of 1998-2009 and this resulted into a value of 9.4%. Before 1998, only few small-sized studies were performed [29, 67]. Barreto *et al* [25] reported a TDR prevalence of 5.6% in 1998-2002 in a group of 341 blood donors, using a WHO-survey approach. Noteworthy, this study was able to include 16% recent infections and they displayed a TDR rate of 12%, indicating high transmission rates of drug resistant strains in São Paulo City at that particular time point [25].

Several studies performed between 2000 and 2004, one WHO-survey [27] and one large nationwide study that sampled in 2001[68], showed that the TDR was below 5% [38, 52, 69-72]. This apparently decrease of TDR could be related to the predominant inclusion of chronically infected individuals. However, other reports showed that TDR was still above 5% [73-75]. These discordances are presumably due to the heterogeneity of the included studies and the geographical variation of TDR.

Since 2005, the TDR rates remained above 5%, with 7% in Curitiba and 8% in Porto Alegre and Santa Catarina [33, 39, 68, 73]. Nationwide studies reported 5.2% TDR in 2007 [76], 8.1% in 2008 [42], and 12.3% in 2008-2010 [46]. The apparent nationwide increase of TDR in Brazil could not be confirmed in

our analysis. This might be explained by the large variations in TDR between and within Brazilian regions that were included in our dataset.

Since 2007, TDR per drug class remained stable with overall values ranging between 5-15%, NNRTI between 3-6%, NRTI between 2-6% and PI between 1-3% [32, 42, 50, 51, 76]. The last WHO-survey showed resistance against NRTI in 7.6%, followed by 4% against NNRTI and PI [46].

SOUTHERN CONE

As 84% of the Southern Cone sequences originated from Argentina, general conclusions cannot be made for this region. Nevertheless, their reanalysis using SDRM 2009 resulted into a TDR prevalence of 4.9% for the years in between 1996 and 2005. In Argentina, the number of studies and their respective sample size remained limited before the period 2000-2003 [28, 77]. However, since 2003 several nationwide surveillance studies have been developed that reported TDR rates that ranged from 3.9% in 2003-2005 to 8.4% in 2006-2008 [34, 36, 44, 78]. The last reported WHO-survey showed a substantial level of resistance towards NNRTI (5.6%), NRTI (4.2%) and PI (3%) [36]. Another study that was performed in recently infected individuals between 2004 and 2009 also supported the 8% of TDR with predominance of NNRTI resistance [79]. Almost all Argentinean studies documented TDR against the three common drug classes. In Chile, TDR was less extensive with an overall TDR of 6.2% in 2000-2005 that slightly decreased to 4% in 2006-2008 [41, 80]. In this country, TDR mutations were mainly restricted to NRTI and NNRTI.

4.5 DISCUSSION

Our results showed a TDR of 7.7% in 1922 therapy-naive HIV-1 patients. However, the majority of sequences were obtained from Brazil, Argentina, Mexico and Chile, countries that correspond to half of the people living with HIV-1 in LAC. This percentage is higher than for other developing countries [81-83] but similar to the 6.4% reported by the WATCH study that included seven studies since 2000-2006 [25, 44, 62, 67, 69, 77, 81, 84], and similar to TDR rates in Europe, Canada and USA [2-5]. Although no significant trend in TDR was observed in our analysis, two peaks were observed in 1998-1999 and in 2006-2007. These peaks overlapped with a high inclusion of sequences from Brazil and Venezuela, countries characterized with the highest TDR rates in LAC.

Although we tried to standardize the analysis by implementing the SDRM 2009 list for interpretation of drug resistance, we acknowledge the limitations of this study. The included studies differed e.g. in the

intrinsic characteristics of the analyzed population (age, risk factor of transmission, geographical and socioeconomic factors), design of the study (method of sampling, type of study, criteria or exclusion criteria), clinical characteristics (recent or chronic infection, immunological status, viral load), drug resistance testing (sample type, sequencing method, sequence quality, algorithm used for the analysis of mutations), ART-related factors in the treated population (mono- or bitherapy history, HAART scale-up, delay of start therapy, availability of new compounds, pharmacogenetics, toxicity, adherence problems, drug failure) and viral factors (subtype, fitness) [85-87]. For example, in one Brazilian study 20% TDR was reported in PBMC samples [37]. However, PBMC testing is more apt to detecting drug resistance mutations than plasma, and the intrinsic regional characteristics of the population might impede the generalization of the obtained results [37, 88].

Although MSM is one of the main drivers of the LAC epidemic and they are 33.3 times more likely to be HIV-1 positive [89], we did not find differences in the TDR rates between heterosexuals and MSM. More studies are necessary to understand the trends of TDR, especially in MSM because the risk of imprisonment in some countries of the Caribbean and the stigma in LAC could hinder the accurate evaluation of this population, especially when TDR prevalence has been reported to range between 3.3% in Peru for 2002 to 17% in Brazil for 2010 [35, 89, 90].

In regard to distribution of mutations, since NNRTI have been used widely in LAC as the first line therapy consequently, the amino acid change in position 103 was the most frequent in the sequences analyzed and in the studies included (see figure 4.4C). The amino acid change 215 and 41 were frequently found in our study. However, 184V was more frequently detected in our study than in European or North-American reports, even though the difference was not statistically significant [2-5]. In general, similar patterns have been found in other developing countries but, in the resource-rich settings NNRTI resistance seems to be stabilizing, probably due to the changes in therapy prescriptions [4, 82]. Overall, we did not find trends in particular drug resistance patterns in the analyzed sequences obtained from LAC.

4.6 CONCLUSIONS AND RECOMMENDATIONS

The overall TDR prevalence in LAC was 7.7% for the period 1996-2009 and reached 10% in the last 4 investigated years when data was solely restricted to Brazil and Venezuela. Additionally, large geographical differences have been found. Although lack of information for Caribbean and Central

America prevents us to draw general conclusions, the reported TDR in Cuba and other Caribbean islands is still under the 5% WHO threshold. However, Honduras, Mexico and South American countries have passed the 5% threshold with the exception of Chile. The limited information available for most of the countries should encourage the development of surveys for drug resistance in newly infected people.

REFERENCES

[1] Erice A, Mayers DL, Strike DG, Sannerud KJ, McCutchan FE, Henry K, et al. Brief report: primary infection with zidovudine-resistant human immunodeficiency virus type 1. The New England journal of medicine. 1993 Apr 22;328(16):1163-5.

[2] Jayaraman GC, Archibald CP, Kim J, Rekart ML, Singh AE, Harmen S, et al. A population-based approach to determine the prevalence of transmitted drug-resistant HIV among recent versus established HIV infections: results from the Canadian HIV strain and drug resistance surveillance program. Journal of acquired immune deficiency syndromes (1999). 2006 May;42(1):86-90.

[3] Weinstock HS, Zaidi I, Heneine W, Bennett D, Garcia-Lerma JG, Douglas JM, Jr., et al. The epidemiology of antiretroviral drug resistance among drug-naive HIV-1-infected persons in 10 US cities. The Journal of infectious diseases. 2004 Jun 15;189(12):2174-80.

[4] Vercauteren J, Wensing AM, van de Vijver DA, Albert J, Balotta C, Hamouda O, et al. Transmission of drug-resistant HIV-1 is stabilizing in Europe. The Journal of infectious diseases. 2009 Nov 15;200(10):1503-8.

[5] SPREAD. Transmission of drug-resistant HIV-1 in Europe remains limited to single classes. AIDS (London, England). 2008 Mar 12;22(5):625-35.

[6] Vandamme AM, Camacho RJ, Ceccherini-Silberstein F, de Luca A, Palmisano L, Paraskevis D, et al. European recommendations for the clinical use of HIV drug resistance testing: 2011 update. AIDS reviews. 2011 Apr-Jun;13(2):77-108.

[7] Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents.Department of Health and Human Services.: Available at http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf. 2011:1-167.

[8] Wittkop L, Gunthard HF, de Wolf F, Dunn D, Cozzi-Lepri A, de Luca A, et al. Effect of transmitted drug resistance on virological and immunological response to initial combination antiretroviral therapy for HIV (EuroCoord-CHAIN joint project): a European multicohort study. The Lancet infectious diseases. 2011 May;11(5):363-71.

[9] UNAIDS-WHO. GLOBAL REPORT: UNAIDS Report on the global AIDS epidemic. Geneva; 2010.

[10] Chequer P, Cuchi P, Mazin R, Garcia Calleja JM. Access to antiretroviral treatment in Latin American countries and the Caribbean. AIDS (London, England). 2002 Dec;16 Suppl 3:S50-7.

[11] Teixeira PR, Vitoria MA, Barcarolo J. Antiretroviral treatment in resource-poor settings: the Brazilian experience. AIDS (London, England). 2004 Jun;18 Suppl 3:S5-7.

[12] UNAIDS-WHO. GLOBAL HIV/AIDS RESPONSE. Epidemic update and health sector progress towards Universal Access. Geneva, Switzerland; 2011.

[13] ONUSIDA-Republica Bolivariana de Venezuela. Informe Anual 2009. Fortaleciendo la Respuesta Nacional y Conjunta al VIH y sida. Caracas, Venezuela; 2009.

[14] UNAIDS-WHO. Progress on Global Access to HIV Antiretroviral Therapy: A Report on "3 by 5" and Beyond. Geneva, Switzerland: WHO; 2006.

[15] Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS One. 2009;4(3):e4724.

[16] de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, et al. An automated genotyping system for analysis of HIV-1 and other microbial sequences. Bioinformatics (Oxford, England). 2005 Oct 1;21(19):3797-800.

[17] Ravasi G, Pérez-Rosales MD, Mejía Cruz VM, Zepeda K, Pérez ME, Aguilar R, et al. WHO HIV drug resistance early warning indicators as treatment program monitoring and evaluation tool in Central American countries. *6th IAS Conference on HIV Pathogenesis, Treatment and Prevention: Abstract no TUAD0105* Rome, Italy 2011.

[18] Jack N, Ravasi G, Schrooten W, Del Riego A, Hirnschall G. Implementing HIV drug resistance early warning indicators in the Caribbean. *XVIII International AIDS Conference: Abstract no THPE0354*. Vienna, Austria 2010.

[19] Jordan MR, Kelley K, Saadani-Hassani A, Obeng-Aduasare Y, Bertagnolio S. Monitoring ART Clinic and Program Performance Using World Health Organization HIV Drug Resistance Early Warning Indicators. 18th Conference on Retroviruses and Opportunistic Infections; 2011; Boston, United States; 2011.

[20] Bennett DE, Jordan MR, Bertagnolio S, Hong SY, Ravasi G, McMahon JH, et al. HIV drug resistance early warning indicators in cohorts of individuals starting antiretroviral therapy between 2004 and 2009: World Health Organization global report from 50 countries. Clin Infect Dis. 2012 May;54 Suppl 4:S280-9.

[21] Tuboi SH, Schechter M, McGowan CC, Cesar C, Krolewiecki A, Cahn P, et al. Mortality during the first year of potent antiretroviral therapy in HIV-1-infected patients in 7 sites throughout Latin America and the Caribbean. Journal of acquired immune deficiency syndromes (1999). 2009 Aug 15;51(5):615-23.

[22] Crabtree-Ramírez B, Caro-Vega Y, Shepherd BE, Wehbe F, Cesar C, Cortes C, et al. Cross-Sectional Analysis of Late HAART Initiation in Latin America and the Caribbean: Late Testers and Late Presenters. PLoS One. 2011;6(5):e20272.

[23] Bennett DE, Bertagnolio S, Sutherland D, Gilks CF. The World Health Organization's global strategy for prevention and assessment of HIV drug resistance. Antiviral therapy. 2008;13 Suppl 2:1-13.

[24] Bennett DE, Myatt M, Bertagnolio S, Sutherland D, Gilks CF. Recommendations for surveillance of transmitted HIV drug resistance in countries scaling up antiretroviral treatment. Antiviral therapy. 2008;13 Suppl 2:25-36.

[25] Barreto CC, Nishyia A, Araujo LV, Ferreira JE, Busch MP, Sabino EC. Trends in antiretroviral drug resistance and clade distributions among HIV-1--infected blood donors in Sao Paulo, Brazil. Journal of acquired immune deficiency syndromes (1999). 2006 Mar;41(3):338-41.

[26] Brindeiro R, Vanderborght B, Caride E, Correa L, Oravec RM, Berro O, et al. Sequence diversity of the reverse transcriptase of human immunodeficiency virus type 1 from untreated Brazilian individuals. Antimicrobial agents and chemotherapy. 1999 Jul;43(7):1674-80.

[27] Sa-Ferreira JA, Brindeiro PA, Chequer-Fernandez S, Tanuri A, Morgado MG. Human immunodeficiency virus-1 subtypes and antiretroviral drug resistance profiles among drug-naive Brazilian blood donors. Transfusion. 2007 Jan;47(1):97-102.

[28] Pando MA, Eyzaguirre LM, Carrion G, Montano SM, Sanchez JL, Carr JK, et al. High genetic variability of HIV-1 in female sex workers from Argentina. Retrovirology. 2007;4:58.

[29] Maia Teixeira SL, Bastos FI, Hacker MA, Guimaraes ML, Morgado MG. Trends in drug resistance mutations in antiretroviral-naive intravenous drug users of Rio de Janeiro. Journal of medical virology. 2006 Jun;78(6):764-9.

[30] Eyzaguirre L, Brouwer KC, Nadai Y, Patterson TL, Ramos R, Firestone Cruz M, et al. First molecular surveillance report of HIV type 1 in injecting drug users and female sex workers along the U.S.-Mexico border. AIDS research and human retroviruses. 2007 Feb;23(2):331-4.

[31] Janssen RS, Satten GA, Stramer SL, Rawal BD, O'Brien TR, Weiblen BJ, et al. New testing strategy to detect early HIV-1 infection for use in incidence estimates and for clinical and prevention purposes. Jama. 1998 Jul 1;280(1):42-8.

[32] Arruda E, Simoes L, Sucupira C, Medeiros M, Arruda E, Diaz RS, et al. Short communication: intermediate prevalence of HIV type 1 primary antiretroviral resistance in Ceara State, Northeast Brazil. AIDS research and human retroviruses. 2011 Feb;27(2):153-6.

[33] Brigido LF, Nunes CC, Oliveira CM, Knoll RK, Ferreira JL, Freitas CA, et al. HIV type 1 subtype C and CB Pol recombinants prevail at the cities with the highest AIDS prevalence rate in Brazil. AIDS research and human retroviruses. 2007 Dec;23(12):1579-86.

[34] Dilernia DA, Lourtau L, Gomez AM, Ebenrstejin J, Toibaro JJ, Bautista CT, et al. Drug-resistance surveillance among newly HIV-1 diagnosed individuals in Buenos Aires, Argentina. AIDS (London, England). 2007 Jun 19;21(10):1355-60.

[35] Lama JR, Sanchez J, Suarez L, Caballero P, Laguna A, Sanchez JL, et al. Linking HIV and antiretroviral drug resistance surveillance in Peru: a model for a third-generation HIV sentinel surveillance. Journal of acquired immune deficiency syndromes (1999). 2006 Aug 1;42(4):501-5.

[36] Pando MA, Gomez-Carrillo M, Vignoles M, Rubio AE, dos Ramos Farias MS, Vila M, et al. Incidence of HIV type 1 infection, antiretroviral drug resistance, and molecular characterization in newly diagnosed individuals in Argentina: A Global Fund Project. AIDS research and human retroviruses. 2011 Jan;27(1):17-23.

[37] Sucupira MC, Caseiro MM, Alves K, Tescarollo G, Janini LM, Sabino EC, et al. High levels of primary antiretroviral resistance genotypic mutations and B/F recombinants in Santos, Brazil. AIDS patient care and STDs. 2007 Feb;21(2):116-28.

[38] Varella RB, Ferreira SB, de Castro MB, Zalis MG, Tavares MD. Human immunodeficiency virus type 1 protease and reverse transcriptase mutation patterns among treatment-naive patients in different stages of infection in Rio de Janeiro, Brazil. Journal of medical virology. 2007 Aug;79(8):1033-9.

[39] Ferreira JL, Thomaz M, Rodrigues R, Harrad D, Oliveira CM, Oliveira CA, et al. Molecular characterisation of newly identified HIV-1 infections in Curitiba, Brazil: preponderance of clade C among males with recent infections. Memorias do Instituto Oswaldo Cruz. 2008 Dec;103(8):800-8.

[40] Acevedo W, Gallardo AM, Galaz J, Afani A, Cortes E. [Detection of primary antiretroviral resistance in Chilean patients recently infected with human immunodeficiency virus (HIV)]. Revista medica de Chile. 2007 Nov;135(11):1406-13.

[41] Afani A, Beltran C, Maria Gallardo A, Roessler P, Acevedo W, Vasquez P. [Prevalence of primary antiretroviral resistance among HIV infected patients in Chile]. Revista medica de Chile. 2010 Jun;138(6):669-76.

[42] Inocencio LA, Pereira AA, Sucupira MC, Fernandez JC, Jorge CP, Souza DF, et al. Brazilian Network for HIV Drug Resistance Surveillance: a survey of individuals recently diagnosed with HIV. Journal of the International AIDS Society. 2009;12(1):20.

[43] Kijak GH, Pampuro S, Gomez Carrillo M, Avila MM, Villahermosa ML, Cuevas MT, et al. Resistance Profiles to Antiretroviral Drugs in HIV-1-Naive Patients. 7th Conference on Retroviruses and Opportunistic Infections 2000; San Francisco, United States; 2000.

[44] Petroni A, Deluchi G, Pryluka D, Rotryng F, Bortolozzi R, Lopardo G, et al. Update on primary HIV-1 resistance in Argentina: emergence of mutations conferring high-level resistance to nonnucleoside reverse transcriptase inhibitors in drug-naive patients. Journal of acquired immune deficiency syndromes (1999). 2006 Aug 1;42(4):506-10.

[45] Arreseigor TS, Valle MM, Rimoldi IV. Primary drug resistance in newly diagnosed HIV-1 infected individuals in Buenos Aires province. *XVIII International AIDS Conference: Abstract no CDC0464*. Vienna, Austria 2010.

[46] De Moraes Soares CMP, Vergara TRC, Sucupira MC, Brites C, Brito JDU, Caseiro MM, et al. Prevalence of Transmitted HIV-1 Antiretroviral Resistance among Patients Initiating ART in Brazil: A Surveillance Using Dried Blood Spots. *18th Conference on Retroviruses and Opportunistic Infections*. Boston, United States 2011.

[47] Murillo W, Paz-Bailey G, Morales S, Monterroso E, Paredes M, Dobbs T, et al. Transmitted drug resistance and type of infection in newly diagnosed HIV-1 individuals in Honduras. J Clin Virol. 2010 Dec;49(4):239-44.

[48] Avila-Rios S, Garcia-Morales C, Garrido-Rodríguez D, Ormsby CE, Andrade-Villanueva J, Gonzales-Hernandez L, et al. National Prevalence and Trends of HIV Transmitted Drug Resistance in Mexico. *18th Conference on Retroviruses and Opportunistic Infections* Boston, United States 2011.

[49] Castillo J, Comegna M, Quijada W, Jauvin V, Pinson P, Masquelier B, et al. Surveillance of HIV type 1 drug resistance among naive patients from Venezuela. AIDS research and human retroviruses. 2009 Dec;25(12):1329-33.

[50] Cardoso LP, Queiroz BB, Stefani MM. HIV-1 pol phylogenetic diversity and antiretroviral resistance mutations in treatment naive patients from Central West Brazil. J Clin Virol. 2009 Oct;46(2):134-9.

[51] Carvalho BC, Cardoso LP, Damasceno S, Stefani MM. Moderate Prevalence of Transmitted Drug Resistance and Interiorization of HIV Type 1 Subtype C in the Inland North State of Tocantins, Brazil. AIDS research and human retroviruses. 2011;27(00):[Epub ahead of print].

[52] Medeiros MS, Arruda EA, Guerrant RL, Brown C, Hammarskjold ML, Rekosh D, et al. Genotype testing and antiretroviral resistance profiles from HIV-1 patients experiencing therapeutic failure in northeast Brazil. Braz J Infect Dis. 2007 Aug;11(4):390-4.

[53] Gittens MV, Roth WW, Roach T, Stringer HG, Jr., Pieniazek D, Bond VC, et al. The molecular epidemiology and drug resistance determination of HIV type 1 subtype B infection in Barbados. AIDS research and human retroviruses. 2003 Apr;19(4):313-9.

[54] Perez L, Alvarez LP, Carmona R, Aragones C, Delgado E, Thomson MM, et al. Genotypic resistance to antiretroviral drugs in patients infected with several HIV type 1 genetic forms in Cuba. AIDS research and human retroviruses. 2007 Mar;23(3):407-14.

[55] Ruibal-Brunet IJ, Cuevas MT, Diaz-Torres H, Villahermosa ML, Noa-Romero E, Vazquez de Parga E, et al. Genotypic resistance mutations to antiretroviral drugs in HIV-1 B and non-B subtypes from Cuba. Revista panamericana de salud publica = Pan American journal of public health. 2001 Sep;10(3):174-80.

[56] Vaughan HE, Cane P, Pillay D, Tedder RS. Characterization of HIV type 1 clades in the Caribbean using pol gene sequences. AIDS research and human retroviruses. 2003 Oct;19(10):929-32.

[57] Ahumada-Ruiz S, Flores-Figueroa D, Toala-Gonzalez I, Thomson MM. Analysis of HIV-1 pol sequences from Panama: identification of phylogenetic clusters within subtype B and detection of antiretroviral drug resistance mutations. Infect Genet Evol. 2009 Sep;9(5):933-40.

[58] Castillo J, Mendoza Y, Martinez A, Artega G, Samaniego R, Estripeaut D, et al. Prevalence of transmitted antiretroviral drug resistance in recently HIV infected individuals and newborns from Panama. . *XVIII International AIDS Conference: Abstract no THPE0358* Vienna, Austria 2010.

[59] Lloyd B, O'Connell RJ, Michael NL, Aviles R, Palou E, Hernandez R, et al. Prevalence of resistance mutations in HIV-1-Infected Hondurans at the beginning of the National Antiretroviral Therapy Program. AIDS research and human retroviruses. 2008 Apr;24(4):529-35.

[60] Escoto-Delgadillo M, Vazquez-Valls E, Ramirez-Rodriguez M, Corona-Nakamura A, Amaya-Tapia G, Quintero-Perez N, et al. Drug-resistance mutations in antiretroviral-naive patients with established HIV-1 infection in Mexico. HIV medicine. 2005 Nov;6(6):403-9.

[61] Valle-Bahena OM, Ramos-Jimenez J, Ortiz-Lopez R, Revol A, Lugo-Trampe A, Barrera-Saldana HA, et al. Frequency of protease and reverse transcriptase drug resistance mutations in naive HIV-infected patients. Archives of medical research. 2006 Nov;37(8):1022-7.

[62] Bouchard M, Masquelier B, Moreno M, Deibis L, De Perez GE, Fleury H, et al. HIV type 1 drug resistance among naive patients from Venezuela. AIDS research and human retroviruses. 2007 Mar;23(3):482-5.

[63] Delgado E, Leon-Ponte M, Villahermosa ML, Cuevas MT, Deibis L, Echeverria G, et al. Analysis of HIV type 1 protease and reverse transcriptase sequences from Venezuela for drug resistance-associated mutations and subtype classification: a UNAIDS study. AIDS research and human retroviruses. 2001 May 20;17(8):753-8.

[64] Hernandez M, Matos D, Rodrigues R, Cedeño C, Rivera A. Resumen del estudio sobre perfil del paciente con VIH y adherencia al tratamiento antirretroviral en Venezuela. Venezuela: Ministerio del Poder Popular para la Salud; 2008.

[65] DiazGranados CA, Mantilla M, Lenis W. Antiretroviral drug resistance in HIV-infected patients in Colombia. Int J Infect Dis. 2010 Apr;14(4):e298-303.

[66] Guanira J, Lama JR, Montoya O, Segura P, Ramos E, Ganoza C, et al. HIV drug resistance in the Andean Region: a look after the universal access to HAART in the Andean region. 18th International HIV Drug Resistance Workshop Basic Principles & Clinical Implications; 2009; Fort Myers, United States: Antivir Ther; 14 Suppl 1; Abstract no. 151 2009. p. A174.

[67] Dumans AT, Soares MA, Pieniazek D, Kalish ML, De Vroey V, Hertogs K, et al. Prevalence of protease and reverse transcriptase drug resistance mutations over time in drug-naive human immunodeficiency virus type 1-positive individuals in Rio de Janeiro, Brazil. Antimicrobial agents and chemotherapy. 2002 Sep;46(9):3075-9.

[68] Brindeiro RM, Diaz RS, Sabino EC, Morgado MG, Pires IL, Brigido L, et al. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. AIDS (London, England). 2003 May 2;17(7):1063-9.

[69] Soares MA, Brindeiro RM, Tanuri A. Primary HIV-1 drug resistance in Brazil. AIDS (London, England). 2004 Jun;18 Suppl 3:S9-13.

[70] Soares EA, Santos RP, Pellegrini JA, Sprinz E, Tanuri A, Soares MA. Epidemiologic and molecular characterization of human immunodeficiency virus type 1 in southern Brazil. Journal of acquired immune deficiency syndromes (1999). 2003 Dec 15;34(5):520-6.

[71] Eyer-Silva WA, Morgado MG. A genotyping study of human immunodeficiency virus type-1 drug resistance in a small Brazilian municipality. Memorias do Instituto Oswaldo Cruz. 2005 Dec;100(8):869-73.

[72] Lacerda HR, Medeiros LB, Cavalcanti AM, Ximenes RA, Albuquerque Mde F. Comparison of the epidemiology, profile of mutations, and clinical response to antiretrovirals among subtypes B and F of the human immunodeficiency virus type 1. Memorias do Instituto Oswaldo Cruz. 2007 Sep;102(6):693-9.

[73] Gonsalez CR, Alcalde R, Nishiya A, Barreto CC, Silva FE, de Almeida A, et al. Drug resistance among chronic HIV-1-infected patients naive for use of anti-retroviral therapy in Sao Paulo city. Virus research. 2007 Nov;129(2):87-90.

[74] Pires IL, Soares MA, Speranza FA, Ishii SK, Vieira MC, Gouvea MI, et al. Prevalence of human immunodeficiency virus drug resistance mutations and subtypes in drug-naive, infected individuals in the army health service of Rio de Janeiro, Brazil. Journal of clinical microbiology. 2004 Jan;42(1):426-30.

[75] Rodrigues R, Scherer LC, Oliveira CM, Franco HM, Sperhacke RD, Ferreira JL, et al. Low prevalence of primary antiretroviral resistance mutations and predominance of HIV-1 clade C at polymerase gene in newly diagnosed individuals from south Brazil. Virus research. 2006 Mar;116(1-2):201-7.

[76] Sprinz E, Netto EM, Patelli M, Lima JS, Furtado JJ, da Eira M, et al. Primary antiretroviral drug resistance among HIV type 1-infected individuals in Brazil. AIDS research and human retroviruses. 2009 Sep;25(9):861-7.

[77] Kijak GH, Pampuro SE, Avila MM, Zala C, Cahn P, Wainberg MA, et al. Resistance profiles to antiretroviral drugs in HIV-1 drug-naive patients in Argentina. Antiviral therapy. 2001 Mar;6(1):71-7.

[78] Rubio AE, Pando MA, Dilernia DA, Trinchero J, Rovner NI, Salomon H. Primary resistance in newly diagnosed HIV-1 infected drug-naïve individuals from Buenos Aires, Argentina: a two-period comparison. 18th International HIV Drug Resistance Workshop Basic Principles & Clinical Implications; 2009; Fort Myers, Florida: Antivir Ther; 14 Suppl 1; Abstract no. 152; 2009. p. A175.

[79] Petroni A, Deluchi G, Lopardo G, Pryluka D, Biscayart C, Rotryng F, et al. Update on transmission of HIV resistance to drug-naive patients in Argentina. International HIV & Hepatitis Virus Drug Resistance Workshop and Curative Strategies; 2010; Dubrovnik, Croatia: Antivir Ther; 15 Suppl 2;Abstract no. 141; 2010. p. A179.

[80] Rios M, Delgado E, Perez-Alvarez L, Fernandez J, Galvez P, de Parga EV, et al. Antiretroviral drug resistance and phylogenetic diversity of HIV-1 in Chile. Journal of medical virology. 2007 Jun;79(6):647-56.

[81] Frentz D, Van de Vijver D, Wensing AM, Bowles E, Boucher CA. Worldwide comparison of mutational patterns in studies on transmission of drug resistant HIV-1 across the World. Update on WATCH. *6th European HIV Drug Resistance Workshop*. Budapest, Hungary 2008.

[82] Hamers RL, Derdelinckx I, van Vugt M, Stevens W, Rinke de Wit TF, Schuurman R. The status of HIV-1 resistance to antiretroviral drugs in sub-Saharan Africa. Antiviral therapy. 2008;13(5):625-39.

[83] Shekelle P, Maglione M, Geotz MB, Wagner G, Wang Z, Hilton L, et al. Antiretroviral (ARV) drug resistance in the developing world. Evidence report/technology assessment. 2007 Sep(156):1-74.

[84] Afani A, Ayala M, Meyer A, Cabrera R, Acevedo W. [Primary resistance to antiretroviral therapy in patients with HIV/AIDS in Chile]. Revista medica de Chile. 2005 Mar;133(3):295-301.

[85] Masquelier B, Bhaskaran K, Pillay D, Gifford R, Balestre E, Jorgensen LB, et al. Prevalence of transmitted HIV-1 drug resistance and the role of resistance algorithms: data from seroconverters in the CASCADE collaboration from 1987 to 2003. Journal of acquired immune deficiency syndromes (1999). 2005 Dec 15;40(5):505-11.

[86] Booth CL, Geretti AM. Prevalence and determinants of transmitted antiretroviral drug resistance in HIV-1 infection. J Antimicrob Chemother. 2007 Jun;59(6):1047-56.

[87] Little SJ, Smith DM. HIV treatment decisions and transmitted drug resistance. Clin Infect Dis. 2005 Jul 15;41(2):233-5.

[88] Bon I, Gibellini D, Borderi M, Alessandrini F, Vitone F, Schiavone P, et al. Genotypic resistance in plasma and peripheral blood lymphocytes in a group of naive HIV-1 patients. J Clin Virol. 2007 Apr;38(4):313-20.

[89] Beyrer C, Baral SD, Walker D, Wirtz AL, Johns B, Sifakis F. The expanding epidemics of HIV type 1 among men who have sex with men in low- and middle-income countries: diversity and consistency. Epidemiologic reviews. 2010 Apr;32(1):137-51.

[90] Bermudez EH, Kerr LR, Pinho AA, Mello MB, de Brito AM, Dourado MIC, et al. Antiretroviral drug resistance in men who have sex with men in Brazil. *XVIII International AIDS Conference: Abstract no FRLBC103*. Vienna, Austria 2010.

CHAPTER 5

TRENDS AND PREDICTORS OF TDR AND CLUSTERS WITH TDR IN A LOCAL BELGIAN HIV-1 EPIDEMIC

Published in:

Andrea-Clemencia Pineda-Peña, Yoeri Schrooten, Lore Vinken, Fossie Ferreira, Guangdi Li, Nídia Sequeira Trovão, Ricardo Khouri, Inge Derdelinckx, Paul De Munter, Claudia Kücherer, Leondios G. Kostrikis, Claus Nielsen, Kirsi Littsola, Annemarie Wensing, Maja Stanojevic, Roger Paredes, Claudia Balotta, Jan Albert, Charles Boucher, Arley Gomez-Lopez, Eric Van Wijngaerden, Marc Van Ranst, Jurgen Vercauteren, Anne-Mieke Vandamme, Kristel Van Laethem. Trends and predictors of transmitted drug resistance (TDR) and clusters with TDR in a local Belgian HIV-1 epidemic. Plos One 2014; 9(7): e101738

5.1 SUMMARY

Background: We aimed to study epidemic trends and predictors for transmitted drug resistance (TDR) in our region, its clinical impact and its association with transmission clusters.

Methodology: We included 778 patients from the AIDS Reference Center in Leuven (Belgium) diagnosed from 1998 to 2012. Resistance testing was performed using population-based sequencing and TDR was estimated using the WHO-2009 surveillance list. Phylogenetic analysis was performed using maximum likelihood and Bayesian techniques.

Results: The cohort was predominantly Belgian (58.4%), MSM (42.8%), and chronically infected (86.5%). The overall TDR prevalence was 9.6% (95% confidence interval (CI): 7.7-11.9), 6.5% (CI: 5.0-8.5) for NRTI, 2.2% (CI: 1.4-3.5) for NNRTI, and 2.2% (CI: 1.4-3.5) for protease inhibitors. A significant parabolic trend of NNRTI-TDR was found (p=0.019). Factors significantly associated with TDR in univariate analysis were male gender, Belgian origin, MSM, recent infection, transmission clusters and subtype B, while multivariate and Bayesian network analysis singled out subtype B as the most predictive factor of TDR. Subtype B was related with transmission clusters with TDR that included 42.6% of the TDR patients. Thanks to resistance testing, 83% of the patients with TDR who started therapy had undetectable viral load whereas half of the patients would likely have received a suboptimal therapy without this test.

Conclusion: TDR remained stable and a NNRTI up-and-down trend was observed. While the presence of clusters with TDR is worrying, we could not identify an independent, non-sequence based predictor for TDR or transmission clusters with TDR that could help with guidelines or public health measures.

5.2 INTRODUCTION

In recent years, the number of newly diagnosed HIV-1 patients increased in Belgium [1] with a rate of 10.7 per 100,000 population in 2011, one of the highest rates in Europe [2]. Studies carried out in Europe and America highlighted the important role of transmission networks in the spread of TDR [3-7]. TDR is a clinical and public health issue because it can compromise the response to ART at the individual and population level [8]. Three nationwide studies were performed previously in Belgium and reported a TDR prevalence of 29% (67/231; 95% CI: 23.5– 35.2) between 1995 and 1998 [9], 7.2% (6/83; 95% CI: 3.4–14.9) in 2000 [10] and 9.5% (27/285, 95% CI: 6.6–13.4) between 2003 and 2006 [11]. However, due to differences in methodology and the lack of a recent study, no up-to-date information is yet available on TDR trends in Belgium. Nevertheless, recent reports revealed the rapid

onward transmission of an HIV-1 strain with K103N mutation [12] and the involvement of transmission clusters (TCs) in approximately half of patients with TDR [4] in a local HIV epidemic in Belgium.

Because other studies consistently showed regional differences between the drivers of the HIV-1 epidemic [13, 14], this study aimed to characterize the temporal trend in TDR, the factors associated with TDR including TCs and the clinical impact of TDR for a period of 15 years in a regional epidemic, serviced by the Leuven University Hospitals. The data included socio-demographic, clinical and virological variables.

5.3 MATERIAL AND METHODS

5.3.1 ETHICS STATEMENT

The research was conducted according to the Declaration of Helsinki. Only patients for whom written informed consent was obtained were included in this study, except patients enrolled in care after 2009. In 2009, UZ Leuven implemented a generic “opt out” system. Patients, who logged an objection to use their medical data for research purposes, were not included in this study. The protocol and this consent procedure were approved by the Ethical Committee UZ Leuven (reference ML-8627, approval B322201316521 S52637).

5.3.2 STUDY POPULATION

We analysed data from the cohort of the AIDS Reference Centre (ARC) in Leuven, the capital of the province of Flemish Brabant (Belgium). The ARC in Leuven has been collecting information since 1997 on treated HIV-1 patients and since 1999, also for naive HIV-1 patients, including epidemiological, clinical and virological data, related with the routine patient healthcare services. The prospective clinical use of baseline genotypic drug resistance testing was implemented in 1999 and stored plasma samples from before 1999 were available to retrospectively perform drug resistance testing upon clinician’s request. Therefore HIV-1 sequences for drug naive patients were either prospectively or retrospectively obtained from a sample taken at diagnosis, except for 135 patients for whom a later pre-therapy sample was used. The inclusion criteria for the analysis of TDR in the present study were newly HIV-1 diagnosed between January 1998 and December 2012, availability of a nucleotide sequence before antiviral therapy initiation and age older than 18 years, and this cohort was called the Leuven newly-diagnosed (ND) cohort for the purpose of this study. The only exclusion criterion used was documented vertical transmission. Recent infections were defined using clinical and laboratory information such as

p24 ELISA, HIV-specific antibody ELISA, and Inno-Lia profile. Patients with the following criteria were classified as recently infected: Fiebig stages I-V [15] or no more than 6 months difference between the last seronegative and first seropositive HIV-1 test [11], CD4 count >200 cells/µl and absence of AIDS-defining conditions [16].

5.3.3 DRUG RESISTANCE TESTING

Drug resistance testing was performed using population-based Sanger sequencing of the *pol* gene fragment encoding PR (amino acids 1 to 99) and 5'-prime end of RT (amino acids 1 to 320). Sequences were obtained using the ViroSeq HIV-1 Genotyping System version 2 (Celera Diagnostics, Alameda, CA) or with an in-house method upon failure of the commercial test [17]. Sequences with associated information are available through Euresist (http://www.euresist.org).

TDR mutations were defined according to the 2009 list of surveillance drug resistance mutations from the World Health Organization [18]. Therefore, the nucleotide sequences were submitted to the Calibrated Population Resistance tool version 6.0 (http://cpr.stanford.edu/cpr.cgi). The clinical impact of genotypic drug resistance on first line therapy was evaluated using Rega algorithm [19] version 9.1.0 (available at http://rega.kuleuven.be/cev/avd/software/rega-algorithm).

5.3.4 HIV-1 SUBTYPING

HIV-1 subtypes and CRFs were determined using two HIV-1 subtyping tools, namely Rega version 3 (http://www.bioafrica.net/typing-v3/hiv) and COMET version 0.3 (http://comet.retrovirology.lu/) [20-22]. Sequences with discordant results were analyzed using manual phylogenetic analysis as was explained previously [22]. Briefly, ML phylogenetic trees under the GTR+Γ nucleotide substitution model were built with RAxML [23] and recombination was verified using SimPlot [24].

5.3.5 TRANSMISSION CLUSTER ANALYSES

To investigate the factors associated with TDR and onward transmission of TDR, cluster analyses were performed on the Leuven ND cohort and four additional datasets as controls: (i) all other *pol* sequences from the ARC in Leuven, including treated HIV-1 patients, HIV-1 patients younger than 18 years old and HIV-1 patients with vertical mode of transmission, (ii) HIV-1 *pol* sequences obtained with the search term "Belgium" as sampling country in the Los Alamos HIV sequence database (retrieved from http://www.hiv.lanl.gov date in April 2013), (iii) HIV-1 *pol* sequences from the collaborative study SPREAD that enrolled patients with newly diagnosed HIV-1 infection from 22 European countries

including Belgium between 2002 and 2008 (see details of the study in [25, 26]), and (iv) the 30 most similar sequences to the Leuven ND cohort (retrieved by Basic Local Alignment Search Tool (BLAST) from http://blast.ncbi.nlm.nih.gov/Blast.cgi). The quality control of the sequences was performed using the tool available at http://www.hiv.lanl.gov/content/sequence/QC/ and the criteria previously described [27]. Separate datasets were constructed according to subtype. As an out-group, two or three reference sequences of subtype D or B (retrieved from http://www.hiv.lanl.gov) were included for B and non-B subtypes, respectively. Sequences were aligned with Muscle as accessory application in the program Mega version 5 [28, 29]. Duplicates were removed and the positions encoding surveillance drug resistance mutations were excluded [18], which resulted in an average final length of 950 nucleotides. In the resulting dataset, we had 755 sequences for subtype A, 4225 for subtype B, 1036 for subtype C, 202 for subtype F, 665 for subtype G, 440 for CRF01_AE and 677 for CFR02_AG to perform phylogenetic analyses on. Subtypes with frequencies less than 1% were not included in TCs analyses.

A ML tree was inferred with the nucleotide substitution GTR+Γ model and 1000 bootstrap replicates in RAxML [23]. TCs, including pairs (two individuals) and larger clusters (≥ 3 individuals), were identified by using Cluster Picker (retrieved from http://hiv.bio.ed.ac.uk/software.html) [30] with a genetic distance less or equal than 0.06 substitutions per site and bootstrap support ≥98%. A sensitivity analysis was performed to evaluate the effect of other genetic distances (0.015, 0.030 and 0.045) [31-34].

The robustness of the identified TCs was evaluated using Bayesian phylogenetic analysis. TCs and the closest control sequences together with two reference sequences as an out-group were selected and trees were constructed with BEAST v1.7.5 [35] using a lognormal relaxed molecular clock with the SRD06 model of nucleotide substitutions [36] and a Bayesian skyline coalescent prior. The analyses were run in triplicate for 100 million states and trees were sampled every 10000th states. MCC were summarized using TreeAnnotator after 10% of the burn-in was discarded and visualized with FigTree v.1.4 (available at http://tree.bio.ed.ac.uk). The TCs with a Bayesian posterior probability of 1 were considered robust enough and included in the analysis.

Finally, we defined TCs with TDR as any pair or cluster with more than 3 patients that included at least one patient with TDR from the Leuven ND cohort. The TCs with TDR with more than 3 individuals with similar TDR mutation profile can be suggestive of onward transmission of TDR, they are specifically indicated as TCs with TDR-OT.

5.3.6 STATISTICAL ANALYSIS

Prevalence of TDR and TDR mutations were calculated with a 95% Wilson score confidence interval (95% CI) on the basis of a binomial distribution, and their trend was calculated by logistic regression analysis. Socio-demographic, virological and clinical variables that were significantly associated with TDR or with TCs with TDR were evaluated in the Leuven cohort. Analyses were performed on patients involved in TCs from the Leuven ND cohort, and from Leuven ND cohort with the other four control datasets. Categorical data were compared using the Chi-square test, the Fisher's exact test or regression techniques as appropriate. The t-test or Mann–Whitney U test was used to compare continuous data. The statistical significance was set at $p<0.05$ two-sided. All data were analysed using the statistical R software version 2.13.1.

Those factors that were found to be significantly associated with TDR or TCs with TDR in univariate analysis were included in a Bayesian network analysis. This is a probabilistic model that describes statistical conditional dependencies between multiple variables and was performed using the B-course software adapted by Deforche et al [37]. In this analysis, the arcs were scored based on the stability of the conditional dependency assessed with 100 non-parametric bootstrap replicates. The arcs with bootstrap over 75% were considered and depicted in the consensus network.

5.4 RESULTS

5.4.1 GENERAL CHARACTERISTICS OF THE POPULATION

778 of the 795 patients who were newly diagnosed with an HIV-1 infection and who received a baseline genotypic drug resistance test between January 1998 and December 2012 at University Hospitals Leuven were included in the analysis, they are referred to as the Leuven ND cohort. Two patients were excluded because their risk group was vertical transmission. For 15 patients, the baseline nucleotide sequence did not fulfill the preset quality criteria: 14 sequences did not have the gene fragments encoding PR or RT, and one sequence was excluded due to the presence of more than four stop codons and indels. The included HIV-1 patients were between 18 and 78 years old and were predominantly male (73.7%), of Belgian origin (58.4%), chronically infected (86.5%) with CDC stage 1 or 2 (67.0%) (Table 5.1). Patients originating from Belgium were more frequently diagnosed with a recent infection and displayed higher viral loads and CD4 counts ($p<0.001$). Of all included HIV-1 patients originated from Belgium, 66.7% reported MSM or bisexual contacts as risk factor, whereas

22.5% reported heterosexual contacts. In contrast, HIV-1 patients originating from Sub-Saharan countries reported infection through heterosexual contacts predominantly (79.3%). HIV-1 patients from Sub-Saharan countries were more likely to be co-infected with hepatitis B than patients from Belgium (60.9% vs. 34.8%; OR: 4.49, 95% CI 1.75–12.15, $p < 0.001$). There were 36 HIV-1 therapy-naive patients who did not receive a baseline drug resistance test in this period. This group included more patients of non-Belgian origin (75.0%) and with CD4 count above 500 cells/mL (42.4%, 14/33).

The demographic characteristics of the Leuven ND cohort were compared to the general HIV-1 population in Belgium, as reported by the Belgian Scientific Institute of Public Health (information until 2011) (available at www.wiv-isp.be) [1]. The Leuven ND cohort contained more men (73.5% vs. 61.0%, $p <0.0002$) and Belgians (58.4% vs. 40.6%, $p <0.0002$) and more MSM (55.9% vs. 42.5%, $p <0.0002$). National data only covered gender and country of origin from 1998 to 2011, and transmission risk from 2005 to 2011.

Characteristics at time of sampling	Total		TDR		Univariate		Multivariate	
	n	%	n	%	OR (95% CI)	*p*	OR (95% CI)	*p*
Patients	778	100	75	100				
Male	573	73.7	67	89.3	3.25 (1.52-7.99)	<0.001*		
Pregnant women	15	1.9	1	1.3				
Age in years at enrolment, Mean (SD)	37.5 (±10.6)		39.6 (±12.4)					
<25	69	8.9	6	8.0				
25-34	287	36.9	23	30.7				
35-44	241	31.0	23	30.7				
45-54	121	15.6	12	16.0				
>55	60	7.7	11	14.7				
Country or region of origin								
Belgium	454	58.4	55	73.3	2.09 (1.20-3.77)	0.006*		
Western Europe (except Belgium)	23	3.0	4	5.3				
High-prevalent regions†					0.23 (0.08-0.52)	<0.001*		
Sub-Saharan Africa	198	25.4	7	9.3	0.27 (0.10-0.61)	<0.001*		
Other	24	3.1	0	0				
Other	75	9.6	9	12.0				
Unknown	4	0.5	0	0				
Risk of transmission								
MSM	333	42.8	47	62.7	2.44 (1.46-4.15)	<0.001*		
Heterosexual (high-prevalent country)	179	23.0	5	6.7				
Heterosexual (non-endemic)	121	15.6	13	17.3				
Bisexual	32	4.1	5	6.7				
IVDU	14	1.8	1	1.3				
Unknown	75	9.6	1	1.3				
Other	24	3.1	3	4.0				
Type of infection								
Chronic	673	86.5	58	77.3				

Characteristics at time of sampling	Total		TDR		Univariate		Multivariate	
	n	%	n	%	OR (95% CI)	*p*	OR (95% CI)	*p*
Recent	105	13.5	17	22.7	2.04 (1.06-3.75)	0.02*		
CDC stage§								
1 and 2	521	67.0	55	73.3				
3	246	31.6	19	25.3				
Unknown	11	1.4	1	1.3				
CD4 cell count, median (IQR) ll	335	(163-493)	365	(236-505)				
<200 cells/mm^3	228	29.3	16	21.3				
200-349 cells/mm^3	174	22.4	18	24.0				
350-499 cells/mm^3	181	23.3	21	28.0				
≥500	182	23.4	19	25.3				
Unknown	13	1.7	1	1.3				
HIV-RNA load, median (IQR), log copies/ml ¶	4.76	(4.12-5.31)	4.76	(4.16-5.25)				
Co-infection								
Hepatitis B	23	3.0	4	5.3				
Negative	412	53.0	36	48.0				
Unknown	343	44.1	35	46.7				
Hepatitis C	406	52.2	1	1.3				
Negative	18	2.3	40	53.3				
Unknown	354	45.5	34	45.3				
Subtype								
B	406	52.2	59	78.7	3.77 (2.09-7.17)	<0.001*	3.04 (1.45-6.35)	0.003
Part of transmission cluster	226	29.0	32	42.6	1.95 (1.15-3.25)	0.010*		

Table 5.1: Characteristics of the Leuven ND cohort and factors associated with TDR. * Fisher's test or logistic regression techniques were used. †High prevalent countries were defined as HIV-prevalence over 1% in adult population (UNAIDS 2012), § CDC stage was defined according to 2008 definitions [38], ǁ CD4 count at diagnosis (median 345, IQR: 158-496) was not statistically different from CD4 count at time of sampling. ¶ Viral load at diagnosis (median 4.76 IQR: 4.04-5.31) was not statistically different from viral load at time of sampling. Abbreviations: %: percentage, CDC: Center for Disease Control and Prevention, CI: Confidence interval, IVDU: Intravenous drug user, IQR: interquartile range, MSM: men who have sex with men, n: number, OR: odds ratio, SD: standard deviation.

5.4.2 SUBTYPES

52.2% of the HIV-1 patients were infected with subtype B, followed by CRF02_AG (11.2%), subtype C (10.3%), subtype A (7.7%), CRF01_AE (6.6%), subtype F (2.8%), subtype G (2.1%) and unique recombinant forms (4.8%). Subtypes D, H, J, CRF09_cpx, CRF12_BF, CRF13_cpx, CRF14_BG, CRF18_cpx, CRF22_01A1, CRF37_cpx, and CRF45_cpx were each found in less than 1%. Of the patients with a subtype B infection, 81.0% were of Belgian origin and 71.9% were MSM. Whereas in patients with non-B infections, 33.6% and 51.6% had a Belgian or sub-Saharan origin, respectively, and 69.9% were infected through heterosexual contacts, followed by bisexual/MSM risk factor (13.2%).

5.4.3 LEVEL AND TRENDS OF TDR

The overall TDR prevalence was 9.6% (75/778; 95% CI 7.7-11.9). The prevalence of TDR against NRTI was 6.5% (51/778; 95% CI 5.0-8.5), against NNRTI was 2.2% (17/778; 95% CI 1.4-3.5), and against PI 2.2% (17/778; 95% CI 1.4-3.5). In recently infected individuals, the prevalence of overall TDR was 16.2% (17/105; 95% CI 10.4-24.4), significantly higher than in patients with chronic or unknown duration of infection (8.6%, 58/673; 95% CI 6.7-11.0). The prevalence of TDR by drug class also varied in recently infected individuals. The prevalence of TDR against NRTI was 12.4% (13/105; 95%CI 7.4-20.0), against NNRTI 1.9% (2/105; 95%CI 0.5-6.7), and against PI 6.7% (7/105; 95%CI 3.3-13.1). Dual resistance was detected in 10 patients (1.3%): 3 displayed TDR against NRTI and NNRTI, 6 against NRTI and PI and one against NNRTI and PI. The latter patient with NNRTI and PI resistance and 4 out of 6 individuals with NRTI and PI resistance were recently infected patients (5/105; 4.8%). No triple class resistance was observed.

The majority of the 75 TDR patients displayed one single mutation (70.7%), mainly related to NRTI (58.5%) and NNRTI resistance (24.5%). The revertants at RT position 215 were the most prevalent (44%), followed by M41L (18.7%), K103N (17.3%), L210W (10.7%), K219Q (10.7%), D67N (6.7%), K219R (5.3%), G190A (4.0%), M184V (2.7%), L74V (1.3%), Y115F (1.3%) and Y181C (1.3%). Within PR, I54VT (10.7%) was the most frequent mutation followed by M46IL (9.3%), N88D (6.7%), V82TS (2.7%), L24I (1.3%), I54T (1.3%) and I85V (1.3%). As the inclusion of PR position 46 within the TDR mutation list has been debated due to its polymorphic nature [39], TDR was recalculated excluding this position. This resulted in an overall TDR of 8.9% (69/778; 95% CI 7.0-11.0) and a PI-TDR of 1.4% (11/778; 95% CI 0.8-2.5).

No significant time trends were found in the overall TDR prevalence, nor in transmitted NRTI and PI resistance (see Figure 5.1A). A parabolic trend was observed for NNRTI-TDR (p=0.019) with a peak in 2008. That corresponded with a peak in occurrence of K103N (p=0.026), the only mutation with a temporal trend. Surprisingly, the parabolic temporal NNRTI-TDR trend was not observed in the recently infected individuals. Instead, a stable temporal trend was observed for overall TDR and individual drug classes in this subset of patients. When the analysis was performed according to region of origin, a significant parabolic trend of NNRTI resistance was only found in patients originating from Belgium (p=0.039).

5.4.4 FACTORS ASSOCIATED WITH TDR

Univariate analysis was performed to identify predictors of TDR (Table 5.1), which were male gender (odds ratio (OR) 3.25, 95% CI 1.52-7.99, p<0.001), Belgian origin (OR 2.09, 95% CI 1.20-3.77, p=0.006), MSM transmission (OR 2.44, 95% CI 1.46-4.15, p<0.001), recent infection (OR 2.04, 95% CI 1.06-3.75, p=0.02), being part of TCs (OR 1.95, 95% CI 1.15-3.25, p=0.010) and infected with subtype B virus (OR 3.77, 95% CI 2.09-7.17, p<0.001). Only the latter remained a significant factor (OR 3.04, 95% CI 1.45-6.35, p=0.003) in multivariate analysis. Since these types of analyses do not display possible interdependencies of the variables and subtype B was most frequently found in MSM originating in Belgium (p <0.001), we evaluated the interdependencies of the variables with a Bayesian network approach. TDR was directly associated only with subtype B, but this subtype was strongly associated with MSM (100% bootstrap support) and with Belgian origin (97% bootstrap support), and to a lesser extent to being part of TCs (78% bootstrap support) (See Figure 5.1B). To verify whether we could find an important predictor of TDR that could be used in guidelines to target a subpopulation of newly diagnosed for preferential drug resistance testing, we repeated the analysis excluding any

information that results from the genotype itself. When subtype B was thus excluded from the analysis, then male gender became directly associated with TDR (64% bootstrap support), and with MSM and Belgian origin (100% bootstrap support), whereas the association between the two latter variables with TCs had lower bootstrap support (31%).

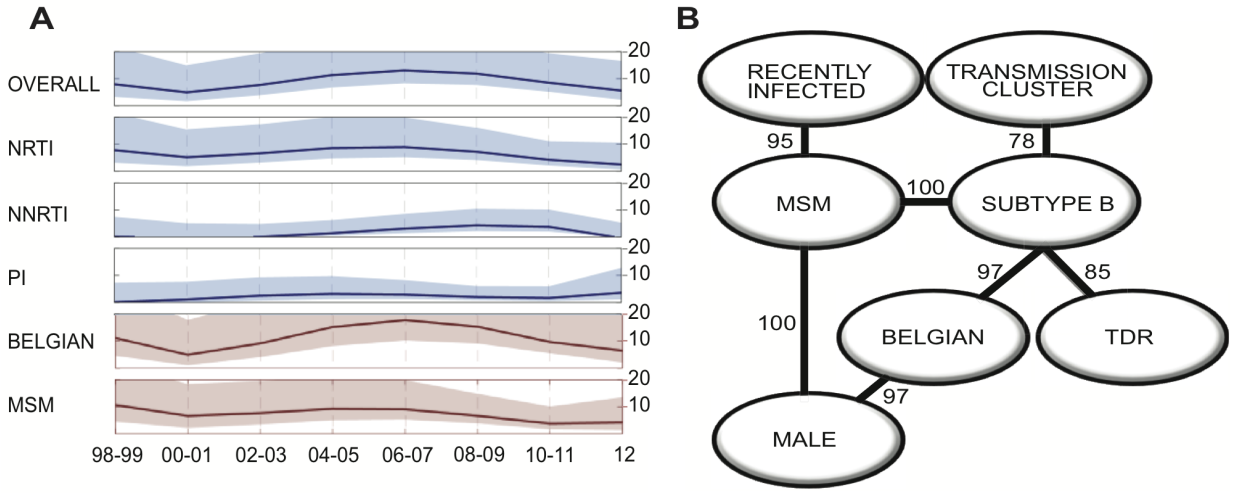


FIGURE 5. 1. TEMPORAL TRENDS AND FACTORS ASSOCIATED WITH TDR

Figure 5.1: Temporal trends and factors associated with TDR. (A) Trends of prevalence of TDR (percentage) and the 95% confidence intervals (light shading) among newly diagnosed HIV-1 patients at ARC Leuven (Belgium) from 1998 to 2012 are shown for the overall-TDR, NRTI-TDR, NNRTI-TDR, PI-TDR in blue, MSM overall-TDR and Belgian overall-TDR in red. **(B)** The significant variables associated with TDR in the univariate analysis were included in the Bayesian network, the number next to the arcs represents the bootstrap support. Abbreviations: NRTI: nucleoside reverse transcriptase inhibitors, NNRTI: non-nucleoside reverse transcriptase inhibitors, MSM: men who have sex with men, PI: protease inhibitors.

5.4.5 TRANSMISSION CLUSTERS

We identified 114 TCs, 16 of which harbored 32 of the 75 TDR patients from our Leuven ND cohort. Five pairs and eight larger clusters of ≥ 3 individuals were found among subtype B infected patients, one cluster of 17 individuals with CRF02_AG, and one pair for each subtypes C and CRF01_AE (see Table 5.2). Six of these 16 TCs with TDR included only a single patient with TDR whereas six clusters were TCs with TDR-OT and included 20 individuals from the Leuven ND cohort (26.7%, 20/75). Singletons were frequently found in patients from the Leuven ND cohort involved in TCs (81.3%,

26/32). Likewise, prevalence against NRTI was the most frequent (81.3%, 26/32), followed by PI (25.0%) and NNRTI (12.5%). Thymidine analogue mutations (TAMs) were predominantly detected in TCs with TDR (81.3%, 26/32), mainly represented by the revertant at position 215 (59.4%, 19/32), followed by the mutations K219QR (18.8%), L210W (15.6%) and M41L (6.3%). Mutations for NNRTI and PI were I54V, N88D (each 15.6%), K103N (12.5%) and M46IL (9.4%).

The characteristics of the Leuven ND cohort patients involved in TCs were evaluated. Patients carrying TDR were significantly more associated with TCs compared to patients without TDR (OR: 1.95, see Table 5.1 and supplementary Table 5.1) and the association remained when considering only larger clusters by excluding pairs (OR: 2.86, 95% CI 1.59-5.01, p=<0.001). Similarly, when including only TCs with TDR-OT, TDR remained significantly associated with TCs (OR: 2.44, 95% CI 1.32-4.36, p=0.002, see supplementary Table 5.1). 32 out of 75 TDR patients were involved in TCs (42.6%, 95% CI 32.1-53.9), while of the 703 patients without TDR, only 194 were found in TCs (27.6%, 95% CI 24.4-31.0). As expected, Leuven ND cohort patients with TDR and involved in TCs were significantly more of Belgian origin (90.6% versus 65.5%; OR: 4.84, 95% CI 1.41-25.78, p=0.005), infected with subtype B (90.6% versus 63.9%; OR: 5.42, 95% CI 1.59-28.85, p=0.001) and characterized with MSM risk factor (78.1% versus 52.6%; OR: 2.71, 95% CI 1.07-7.84, p=0.03) than their counterparts without TDR (see Supplementary Table 5.1). When we focussed on predictors for TCs with TDR by including the data of controls, subtype B remained significantly associated with TCs with TDR (77.4% versus 65.1%, OR: 1.83 95% CI 1.03-3.35, p=0.036). However, TCs of patients with TDR were larger than TCs without TDR (median 3.5 vs. 2 patients per TC, OR: 1.43 95% CI 1.08-1.91, p=0.001). Belgium as sampling country or as country of origin and heterosexual contact were then more frequent in the group of TCs that included solely patients without evidence of TDR. Multivariate analysis did not show any significant factor associated with patients in TCs with TDR versus other TCs. Finally, we also performed separate analyses on TCs with TDR-OT (see supplementary Table 5.1). The same variables remained significantly associated with TCs in the univariate analysis, with the exception of recent infection that became significant. The median of the TCs with TDR-OT was larger than TCs without TDR (5.5 versus 2 patients per TCs, OR: 2.01 (1.29-3.11). Likewise, none of the variables were significant in the multivariate analysis.

Cluster	Patient ID†	Subtype	Risk of transmission	Country of sampling	Country of origin	Country of infection	Type of infection	Year of diagnosis	TDR mutations		
									NRTI	NNRTI	PI
1	AY165241	***B***	Unknown	Sweden	Unknown	Unknown	Naive - Unknown	2000*	-	-	-
	ARCL-1		***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2002***	***M41LM***	-	-
	ESAR-1		Heterosexual - Male	Spain	Pakistan	Spain	Naive - Recent	2003	-	-	-
	FJ481828		MSM	Spain	Unknown	Unknown	Naive - Unknown	2006*	-	-	-
	ESAR-2		MSM	Finland	Russia	Finland	Naive - Chronic	2007	-	-	-
2	EU248399	***B***	Unknown	Belgium	Unknown	Unknown	Naive - Unknown	2003*	-	K103N	-
	JN101707		Unknown	United Kingdom	Unknown	Unknown	Unknown	2004*	-	K103N	-
	ARCL-2		***Heterosexual - Male***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2007***	***-***	***K103N***	***-***
	ARCL-3		***Bisexual - Male***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2009***	***-***	***K103N***	***-***
	JF683797		Unknown	Cyprus	Unknown	Unknown	Naive - Unknown	2009*	-	K103N	-
3	***ARCL-4***	***B***	***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive -Recent§***	***2001***	***L210W, T215S***	***-***	***I54V, N88D***
	EU248438		Unknown	Belgium	Unknown	Unknown	Naive - Unknown	2003*	L210W, T215S	-	I54V, N88D
	ARCL-5		***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Recent***	***2005***	***L210W, T215S***	***-***	***I54V, N88D***
	ARCL-6		***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic§***	***2007***	***L210W, T215S***	***-***	***I54V, N88D***
	ARCL-7		***MSM***	***Belgium***	***Italy***	***Belgium***	***Naive - Recent***	***2010***	***L210W, T215S***	***-***	***I54V, N88D***
	ARCL-8		***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Recent***	***2011***	***L210W, T215S***	***-***	***I54V, N88D***
4	DQ177230	***B***	Unknown	Belgium	Unknown	Unknown	Naive - Unknown	2002*	T215E	-	-
	ARCL-9		***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2005***	***T215E***	***-***	***-***
5	***ARCL-10***	***B***	***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic§***	***2007***	***T215D***	***-***	***-***
	ARCL-11		***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2007***	***T215D***	***-***	***-***
	ARCL-12		***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2007***	***T215D***	***-***	***-***
	ARCL-13		***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2007***	***T215D***	***-***	***-***
	ARCL-14		***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2010***	***-***	***-***	***-***
	ARCL-15		***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2010***	***-***	***-***	***-***
	ARCL-16		***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Recent***	***2011***	***T215D***	***-***	***-***
	ARCL-17		***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Recent***	***2011***	***T215D***	***-***	***-***
	ARCL-18		***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2012***	***T215D***	***-***	***-***
6	DQ206665	***B***	Unknown	Argentina	Unknown	Unknown	Naive - Unknown	2004*	-	K103N, P225H	-

Cluster	Patient ID†	Subtype	Risk of transmission	Country of sampling	Country of origin	Country of infection	Type of infection	Year of diagnosis	TDR mutations		
									NRTI	NNRTI	PI
	JN670104		Unknown	Argentina	Unknown	Unknown	Treated	2005*	M41L, M184V, T215Y	-	D30N, N88D
	ARCL-19		***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2007***	***T215S***	***K103N***	***-***
	ARCL-20		***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Recent***	***2007***	***-***	***K103N***	***-***
7	***ARCL-21***	***B***	***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Recent***	***2009***	***-***	***-***	***M46L***
	ARCL-22		***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2010***	***K219Q***	***-***	***-***
	JQ650683		MSM	Netherlands	Unknown	Unknown	Unknown	Unknown	-	-	M46L
8	***ARCL-23***	***B***	***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Recent***	***2011***	***T215E***	***-***	***-***
	JQ650714		MSM	Netherlands	Unknown	Unknown	Unknown	Unknown	T215E	-	-
9	***ARCL-24***	***B***	***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2006***	***-***	***-***	***-***
	ARCL-25		***Transfusion - Male***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2010***	***K219RK***	***-***	***-***
10	***ARCL-26***	***B***	***MSM***	***Belgium***	***Indonesia***	***Asia***	***Naive - Chronic***	***2001***	***T215D***	***-***	***-***
	ARCL-27		***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2003***	***T215D***	***-***	***-***
11	***ARCL-28***	***B***	***MSM***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***1998***	***T215C***	***-***	***-***
	ARCL-29		***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2002***	***T215S***	***-***	***-***
12	***ARCL-30***	***B***	***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Recent***	***2003***	***-***	***-***	***-***
	ARCL-31		***Bisexual - Male***	***Belgium***	***Belgium***	***Belgium***	***Naive - Recent***	***2009***	***K219R***	***-***	***-***
	ARCL-32		***Bisexual - Male***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2009***	***-***	***-***	***-***
	ARCL-33		***Unknown***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2010***	***-***	***-***	***-***
13		***B***									
	EU817049		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	1998*	K219Q	-	-
	EU817059		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2001*	K219Q	-	-
	EU817062		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2001*	K219R	-	-
	ARCL-34		***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2007***	***K219Q***	***-***	***-***
	ARCL-35		***MSM***	***Belgium***	***Belgium***	***Belgium***	***Naive - Recent***	***2006***	***K219Q***	***-***	***-***
	ESAR-1		MSM/bisexual	Italy	Italy	Unknown	Naive - Chronic	2003	K219Q	-	-
	EU817050		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2003*	K219Q	-	-
	EU817048		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2004*	K219R	-	-

Cluster	Patient ID†	Subtype	Risk of transmission	Country of sampling	Country of origin	Country of infection	Type of infection	Year of diagnosis	TDR mutations		
									NRTI	NNRTI	PI
	EU817061		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2004*	K219Q	-	-
	DQ345509		Unknown	Argentina	Unknown	Unknown	Naive - Recent	2005*	-	-	-
	EU817058		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817060		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817065		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817056		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817051		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817068		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817047		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817055		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817063		Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219QR	-	-
	FJ469703		Unknown	United States	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	ESAR-2		MSM/bisexual	Germany	Germany	Unknown	Naive - Recent	2006	K219Q	-	-
	ESAR-3		MSM/bisexual	Cyprus	Cyprus	Unknown	Naive - Chronic	2007	-	-	-
	ESAR-4		MSM/bisexual	Germany	Germany	Germany	Naive - Recent	2007	-	-	-
	JF683765		Unknown	Cyprus	Unknown	Unknown	Naive - Unknown	2007*	-	-	-
	JF683791		Unknown	Cyprus	Unknown	Unknown	Naive - Unknown	2009*	-	-	-
	JF683808		Unknown	Cyprus	Unknown	Unknown	Naive - Unknown	2009*	-	-	-
14	***ARCL-36***	***C***	***Heterosexual - Female***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2012***	***-***	***-***	***M46LM***
	ARCL-37		***Heterosexual - Male***	***Belgium***	***Ethiopia***	***Unknown***	***Naive - Chronic***	***2012***	***-***	***-***	***-***
15	***ARCL-38***	***CRF01_AE***	***Heterosexual - Female***	***Belgium***	***Thailand***	***Thailand***	***Naive - Chronic***	***2005***	***-***	***-***	***-***
	ARCL-39		***Heterosexual - Male***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2005***	***-***	***-***	***M46IM***
16	JX290261	***CRF02_AG***	Unknown	Russia	Unknown	Unknown	Unknown	2001 or 2002‡	-	-	-

Cluster	Patient ID†	Subtype	Risk of transmission	Country of sampling	Country of origin	Country of infection	Type of infection	Year of diagnosis	TDR mutations		
									NRTI	NNRTI	PI
	AY829204		IVDU - Male	Uzbekistan	Unknown	Unknown	Unknown	2002*	-	-	-
	AY829207		IVDU - Male	Uzbekistan	Unknown	Unknown	Unknown	2002*	-	-	-
	AY829214		IVDU - Male	Uzbekistan	Unknown	Unknown	Unknown	2002*	-	-	-
	HQ449394		Unknown - Female	Russia	Unknown	Unknown	Unknown	2005‡	-	-	-
	DQ465230		MTCT	USA	Unknown	Unknown	Unknown	2006*	-	-	-
	GQ290726		Unknown	South Korea	Unknown	Unknown	Naive - Unknown	2008*	-	-	-
	GQ290743		Unknown	South Korea	Unknown	Unknown	Naive - Unknown	2008*	-	-	-
	HQ412530		Unknown - Male	Russia	Unknown	Unknown	Unknown	2008*	L74I, M184V, K219E	L100I, K101E, Y181C, G190S	-
	HQ115069		Unknown - Male	Ukraine	Unknown	Unknown	Unknown	2009‡	-	-	-
	JX500703		Unknown	Russia	Unknown	Unknown	Unknown	2010*	-	-	-
	ARCL-40		***Heterosexual - Female***	***Belgium***	***Kazakhstan***	***Unknown***	***Naive - Chronic***	***2011***	***K219RK***	***-***	***-***
	JX500697		Unknown	Russia	Unknown	Unknown	Unknown	2011*	-	-	-
	JX500706		Unknown	Russia	Unknown	Unknown	Unknown	2011*	-	-	-
	KC509858		Unknown - Female	Russia	Unknown	Unknown	Unknown	2012*	-	-	-
	KC120872		Unknown	South Korea	Unknown	Unknown	Unknown	Unknown	-	-	-
	KC120881		Unknown	South Korea	Unknown	Unknown	Unknown	Unknown	-	-	-

Table 5.2: Characteristics of transmission clusters containing Leuven patients with TDR. Abbreviations: ARCL Aids Reference Center Leuven, CRF: circulating recombinant form, ESAR: European Society for translational Antiviral Research, IVDU: intravenous drug user, NRTI: nucleoside reverse transcriptase inhibitors, NNRTI: non-nucleoside reverse transcriptase inhibitors, MSM: men who have sex with men, MTCT: mother to child transmission, PI: protease inhibitors,† Patient ID includes patients of the Leuven cohort (bold and italics), ESAR controls and accession numbers of NCBI database * Control sequences have available year of sampling. ‡ Control sequences with year of diagnosis available. § Sequences were also included when the patient was on antiretroviral treatment.

The characteristics of the Leuven TDR patients involved in TCs are shown in Table 5.2. Five pairs included mainly naive MSM originating from Belgium, infected with a subtype B strain carrying one of the 215 revertants, whereas two pairs with subtype C and CRF01_AE strains displaying mutations at PR position 46 included heterosexual Belgians with a foreign partner. Two large subtype B TCs with TDR included 9 or more patients. Cluster number 5 was composed of nine therapy-naive individuals originating from Belgium with MSM as a risk factor and diagnosed between 2007 and 2012. Seven of them displayed a revertant at RT position 215 while two strains had no TDR. Cluster number 13 (Figure 5.2B) included 26 naive patients mainly from United Kingdom and countries of western and southern Europe, western Asia and America. The main mode of transmission was MSM infected with viruses carrying mutations at position 219, except for six individuals who mainly originated from Cyprus and did not display any mutations. The remaining subtype B TCs were composed of three to six individuals. All had at least one individual originating from Belgium and one from another country. For instance, cluster number 1 included one patient originating from Belgium and two patients from Russia and Pakistan but infected in other countries like Finland and Spain. One peculiar subtype B cluster with extensive NRTI and PI resistance (cluster number 3) contained 6 therapy-naive MSM with TDR all with a Belgian connection, either infected in Belgium or originating from Belgium. On the other hand, the largest cluster of non-B subtypes was composed of 17 individuals infected with CRF02_AG and sampled in different countries from Central-eastern Asia and Eastern Europe and characterized by different risks factors including heterosexual orientation, IVDU or vertical transmission. The resistance pattern was also heterogeneous in this cluster. The majority of patients did not display TDR mutations, whereas one naive patient sampled in Belgium and originating from Kazakhstan displayed a mixture of K219RK, and one patient who was tested in Russia was probably treated and displayed high-level resistance against NRTI and NNRTI.

Since the definition of clustering is still a matter of debate, additional analyses were performed to assess the impact of the genetic distance on the identification of TCs with values of 0.015, 0.03 and 0.045. As a result, the number of TCs with TDR decreased to 9, 13 and 17, respectively, including 22.7%, 29.3% and 36.0% of the 75 patients with TDR. The number of TCs that included TDR patients was larger with the genetic distance 0.045 because the cluster with 9 individuals identified as number 5 in Table 5.2 was split in one pair (ARCL-16-TDR and ARCL-18-TDR) and one cluster with the remaining patients. With the stringent criteria of 0.015, TDR was still associated with TCs (OR: 2.56 05% CI 1.32-4.76, p=0.003). Subtype B infection remained as the factor associated with TCs with TDR (p=0.002), whereas sociodemographic factors such as Belgian nationality and MSM contact were not significantly linked.

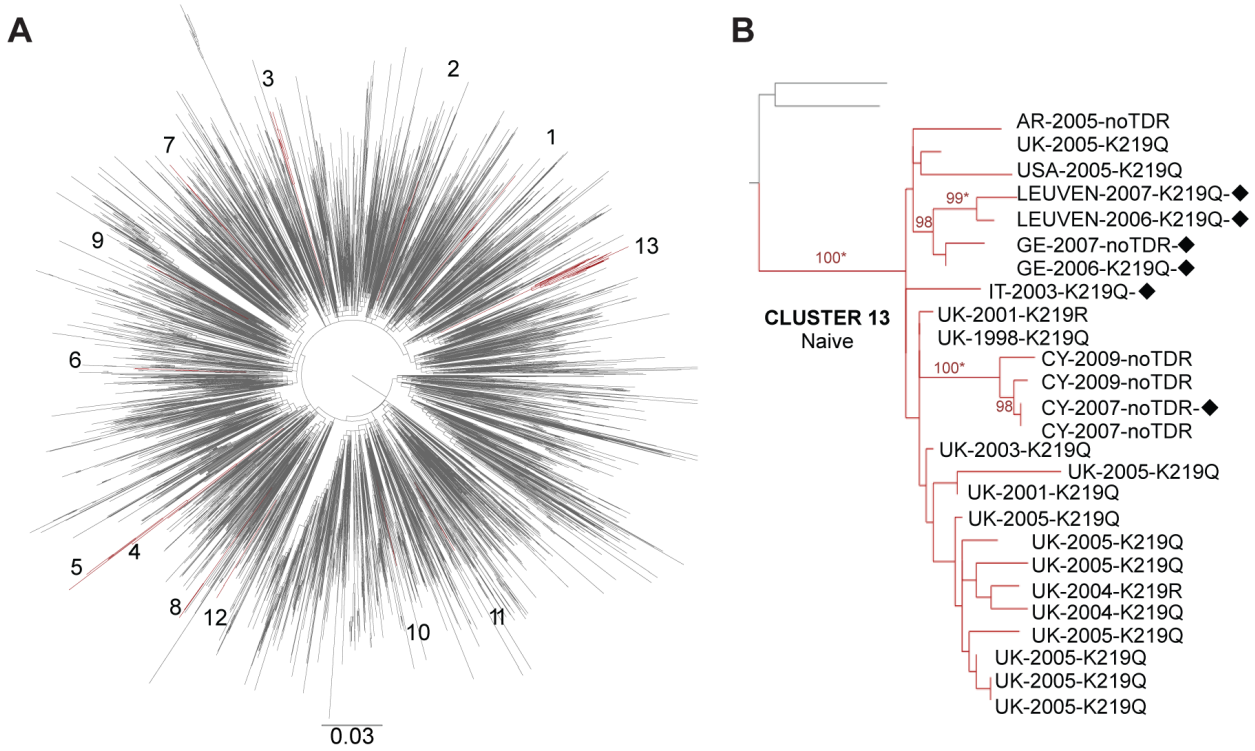


FIGURE 5. 2. EXAMPLES OF SUBTYPE B TRANSMISSION CLUSTERS WITH TDR

Figure 5.2: Examples of subtype B transmission clusters (TCs) with TDR: A maximum likelihood (ML) tree per subtype was constructed, and TCs were confirmed by Bayesian Phylogenetic analyses. **(A)** The ML tree for subtype B (Leuven ND cohort and control sequences) with the TCs colored in dark red. **(B)** The largest TC of subtype B: composed of therapy-naive patients, several nationalities and mutations at RT position 219; bootstrap values above 98% are shown**.** Abbreviations: AR: Argentina, CY: Cyprus, GE: Germany, IT: Italy, UK: United Kingdom, USA: United States of America, black diamond: men who have sex with men, asterisk: posterior distribution equal to 1 in the Bayesian phylogenetic analysis.

5.4.6 POTENTIAL IMPACT OF TDR ON FIRST-LINE REGIMEN

For each patient with TDR, we analysed the genotypic susceptibility score (GSS) of the three regimens most frequently prescribed during the year the patient was diagnosed, and for each patient who in the meantime had started treatment we also estimated the GSS of the first-line therapy received. Rega Algorithm 9.1.0 was used for these analyses, which has a recommended GSS for patients with TDR. This recommendation was introduced in March 2007 and suggests a regimen with GSS of ≥3.5 when TDR is detected, thereby suggesting a triple therapy with a fully active boosted PI which receives a score 1.5, but not with an NNRTI which receives a score of 1 when fully active.

Amongst the 636 patients who started ART, the GSS of the actual prescribed first-line regimen was ≥3.5 in 261 patients (41.0%), 3 in 357 patients (56.1%), and <3 in 18 patients (2.8%). The latter group included five patients without TDR, of which three patients were treated with bi-therapy in 1998, one with bi-therapy including a PI in 2002 and one with a mono PI regimen in 2010. Although the demographic characteristics were similar between the patients with a GSS ≥3 and <3, the latter group more often started ART before 2002 (8/18, p<0.001).

In the group of patients with viruses carrying TDR, 60 started therapy before the end of 2012. The GSS was ≥3.5 in 34 patients (56.7%), equal to 3 in 13 patients (21.7%), and <3 in 13 patients (21.7%). The latter group displayed resistance only to NRTIs (6/13), only to PI (2/13) or dual resistance (5/13).
Sustained undetectable viral load during the first-line therapy was reached in 83.3% (50/60) of the TDR patients, whereas 3.3% of the patients (2/60) had sustained low level viremia without evidence of virological rebound above 500 copies/ml. Four patients (6.7%) had early changes in ART due to toxicity and 1 patient (1.7%) died shortly after therapy initiation. Only 5.1% of the patients (3/60) displayed virological failure and the emergence of major NNRTI resistance-related mutations. Two of these patients displayed evidence of only NRTI TDR mutations at baseline and received a NRTI+NNRTI regimen with a GSS equal to 3. The third patient started a NRTI+NNRTI therapy with a GSS equal to 1 in 2004, two days after the first contact date and two weeks before the drug resistance results were available. In this patient, the therapy was quickly changed after receiving the baseline drug resistance report indicating NRTI and NNRTI TDR mutations and observing no virological response. Subsequent drug resistance testing on a later sample revealed further accumulation of NNRTI resistance.

To appreciate the value of baseline drug resistance testing, the GSS was calculated for the three most frequently prescribed first-line regimens in patients displaying TDR in our cohort per year (see supplementary Table 5.2). Theoretically, 49.3% of TDR patients (37/75) were likely to receive a potential suboptimal regimen (GSS <3) if baseline drug resistance testing had not been available to the treating physician. However, the frequency of these patients is decreasing over time (p=0.002). Up until 2003, the most common first-line regimens included a thymidine analogue with either a NNRTI- or unboosted PI. Thereafter, tenofovir or abacavir were more commonly used as supporting NRTI, but a GSS <3 was still mainly observed for NRTI+NNRTI regimens. A low GSS under a boosted PI based regimen would only have accounted for approximately 15% of the TDR cases per year between 2005 and 2011.

5.5 DISCUSSION

The TDR surveillance in the 778 included patients who were newly diagnosed with HIV-1 at our clinic in Leuven showed a stable overall prevalence of 9.6% between 1998 and 2012. This result is in line with the 9.5% of the latest national survey that included 285 patients who were newly diagnosed in Belgium between 2003 and 2006 [11]. It was also consistent with the stable overall TDR levels of 9.7% in Spain and of 9.0% in France, results from national surveys with a similar design and time frame as our study [34, 40], and with the overall trend of 8.9% in Europe between 2002 and 2007 [26, 41]. However, the overall TDR in our local epidemic was higher than the 5.6% between 2003 and 2010 in Sweden [5] and the 6.5% between 2001 and 2009 in Ghent (Belgium) [4]. These regional differences highlight the importance of studying local epidemics and suggest that TDR prevalence may vary within a single country. Indeed, our findings may not be generalizable to the HIV-1 epidemic in Belgium because our cohort has a higher prevalence of MSM and of individuals originating from Belgium. It should be pointed out, however, that the demographic characteristics data of our study population was more complete than the national database for which nationality and mode of transmission were not available in approximately 25% of the cases. Although our study could have overestimated the level of TDR, due to patients who were unwilling to disclose their ART status, measures were taken to decrease the number of misclassifications. Patient records were exhaustively revised by clinicians and virologists, and individuals with evidence of drug resistance and viral load profiles suggestive of treatment were not considered drug naive.

While subtype B, being MSM, male gender, originating from Belgium, recently infected, and involvement in TCs were all significant predictors for TDR, only subtype B remained significantly associated in the multivariate analysis. Bayesian analyses however, showed the dependency of this factor on being of Belgian origin and MSM. Similarly when subtype B was excluded, male gender became directly associated with TDR with low support but significantly dependent on Belgian origin and MSM. These findings are in agreement with previous reports from other Belgian, European and American studies [4, 5, 11, 40, 42]. In a recent study, the association of subtype with country of sampling, risk group and gender has been interpreted as evidence for highly compartmentalized epidemics in Europe [43]. Therefore, the early introduction of subtype B in European MSM and their broad access to HIV care and ART for decades might explain the single direct association of subtype B with TDR in many resource-rich settings. Nevertheless, recent studies revealed an increasing prevalence of TDR among Sub-Saharan African migrants residing in Spain and Sweden, potentially linked to the increasing drug resistance levels in Africa [5, 44]. However in our cohort, Sub-Saharan

African patients were still associated with less TDR and we did not observe a time trend in those patients (data not shown).

Fluctuations in TDR levels were observed over the entire study period, but the only significant trend was a parabolic trend detected for transmitted NNRTI resistance. The overall NNRTI TDR prevalence was 2.2%, with a maximum of 6.5% in 2008-2009. This parabolic trend, mainly linked to the detection of K103N and to a Belgian origin, was not observed among recently infected patients. Although a parabolic trend with a peak in 2004 was also described in the SPREAD study that included data up until 2005 [26], the same surveillance up to 2007 showed a linear increase over time [41] potentially associated with the frequent use of NNRTI in first-line regimens as the authors suggested. Similarly, the local change of prescribing practices to more potent regimens in later years, use of drug resistance testing and the longer time period analyzed in this study could explain the parabolic trend. In the total cohort, TDR associated with NRTI and PI resistance fluctuated around 6.5% and 2.2%, respectively. Among recently infected patients, NRTI- and PI-TDR levels increased to 12.4% and 6.7% respectively. This increase was not observed for NNRTI resistance, presumably due to the lower impact of NNRTI mutations on viral fitness with consequently a lower likelihood of reversion to wild-type and of a TDR underestimation by population-based Sanger sequencing in chronically infected patients.

Singletons were predominantly detected in our cohort with TAMs as the most commonly observed. Although the use of zidovudine has decreased over the last few years, we did not find any time trend for TAMs. The peak in NNRTI TDR in 2008 was not related to clustered transmission of TDR or migration from other countries as has been suggested in other settings [12, 44]. However, it might have been linked with the enhanced use of NNRTI-containing combinations in the years before. From 2009 onwards, the most commonly prescribed regimen was the potent combination tenofovir + emtricitabine + efavirenz. Up until 2012, no other available NNRTI- or PI/ritonavir-based regimen had proven superior to this regimen with respect to virologic responses.

If resistance testing had not been performed and patients with TDR would have received one of the preferred first-line regimens at that time, approximately half of them would likely have received a regimen in which the virus had lost susceptibility to at least one of the prescribed drugs. Irrespective of the detected TDR by population-based Sanger sequencing, all of them would have had a higher risk of virological failure, as NNRTI-based regimes were commonly prescribed from 2002 onwards [8]. However, 83% of the patients with TDR and who started ART achieved undetectable viral load thanks to the prescription of potent regimens enabled by the availability of drug resistance results. Only 3

patients with baseline resistance to NRTI had virological failure with development of NNRTI resistance after the initiation of a NRTI+NNRTI regimen.

In this cohort, 42.6% of the TDR patients were involved in TCs, which included nine clusters and seven pairs. Because we were interested in observing TCs over a period of 15 years, a genetic distance of 0.06 substitutions per site and a bootstrap support of ≥98% were used to define TCs. These TCs were also confirmed using Bayesian Phylogenetic techniques, indicating that the obtained results were robust [4, 7, 34]. Although, the comparison between studies of transmission networks is difficult due to the differences in sampling, phylogenetic techniques and the lack of a standardized TC definition, our results were in line with a study carried out in Ghent (Belgium) in which 18 out of 33 (55%) TDR patients were involved in pairs or larger clusters [4]. When a more stringent criterion of 0.015 genetic distance was used, the percentage of TDR patients involved in TCs decreased to approximately 23%. In general, some conclusions can be drawn from the TCs analyses. First, none of the factors were significantly associated with TCs with TDR in the multivariate model, although a dependency between subtype B and clustering was found in the Bayesian network and TDR was significantly associated with TCs in the univariate analysis. Therefore we were unable to identify a non-sequence based predictor of being in TCs with TDR, even though the odds were higher for patients who were MSM and originating in Belgium. This result is similar to other studies performed in Europe [4, 34]. Second, TAMs were more frequently found in TCs and this was also observed in the Ghent cohort and other settings [3, 4, 7]. Third, TCs with TDR involved mainly therapy-naive individuals, chronically or recently infected, rather than ART-experienced patients, which could suggest that drug naive people, potentially unaware of their HIV seropositive status, are the main source of TDR instead of patients failing ART [3, 6, 7]. Fourth, 7 out of 16 TCs with TDR involved patients of different nationalities. Although, we were not able to retrace the country of infection in many instances, this may imply that migration plays an important role in the local spread of subtype B as previously described [45], but also of TDR. Fifth, spread of non-B subtypes in the local epidemic was still limited and was related with heterosexuals as has been described in other epidemics [6]. They were also not prone to spread TDR. They often involved Belgians and other nationalities that could imply a limited intermixing of the HIV-1 epidemic between locals and immigrants.

Although we used all the sequences available in public databases and from a collaborative European dataset, for 7 TCs with TDR we did not find evidence that patients other than the ones followed at our clinic were involved in transmission networks. Similarly, 6 TCs with TDR patients from the Leuven ND cohort included only a single patient with TDR, and for these patients, no evidence of onward

transmission of TDR is available. When 3 or more patients in a TC had the same TDR mutation profile, we indicated the cluster as TDR-OT, since this is suggestive of onward transmission of TDR, although we cannot exclude that all these TDR patients received their resistance from a treated patient. The association between TDR and TCs remained, also for those TCs with TDR-OT. Since 27% of the patients from the Leuven ND cohort were involved in those TCs, this could imply an important role of local transmission on the spread of TDR. Nevertheless, we cannot exclude the possibility that the networks might include other intermediary individuals who were not sampled or unaware of their seropositive status, known limitations of phylogenetic analyses.

Singletons and the TAM 215 were predominant in TCs, but the clinical impact on the current first-line therapies remains limited. However, two TCs that involved MSM individuals originating from Belgium with viruses carrying the K103N were detected. The latest diagnosis date was 2009 in these TCs, suggesting that in our local epidemic this mutation was not involved in a recent spread in contrast to a reported outbreak in Namur (Belgium) [12]. Continuous monitoring of the spread of this mutation is required to establish the impact on current practices. On the other hand, two large clusters were detected with the TAM 219 and they involved different nationalities from Europe, Asia and America. One of these TCs contained one patient living in Belgium but originating from Kazakhstan, and control sequences from Uzbekistan that were part of an outbreak of CRF02_AG among IVDU [46]. Our analyses revealed the involvement of other countries and risk groups and the absence of K219R in many of the clustered sequences. The other large cluster included MSM individuals infected with subtype B and control sequences mainly from United Kingdom [47] and other countries in Europe. The majority of strains in this cluster displayed K219Q with only a few strains displaying K219R or no TDR.

5.6 CONCLUSION

In summary, this study showed a stable trend of almost 10% overall TDR between 1998 and 2012 in our Leuven (Belgium) cohort. TDR associated with NNRTI resistance displayed a parabolic trend that overlapped with an up-and-down NNRTI TDR trend in Belgians and with the trend of K103N. Our cohort was mainly composed of chronically infected patients and around 43% of the patients with TDR were involved in transmission networks, suggesting public health policies that target early diagnoses of recently infected patients are needed. Although the main factor related with TDR was subtype B, this variable was dependent on Belgian nationality and MSM mode of transmission. While these variables were also associated with being in TCs with TDR, we were unable to significantly identify a population

that could be targeted for future TDR prevention policies. More local, national and international surveillance studies are needed to confirm the significance and durability of our observations, as changes in TDR levels and patterns are not straightforward to predict due to potential changes in prevention, testing and treatment strategies and changes in other potentially important drivers, such as e.g. behaviour and migration.

REFERENCES

[1] ISP-WIV. Epidemiologie van AIDS en HIV infectie in België. Epidemiologie du SIDA et de l'infection à VIH en Belgique. In: Direction opérationnelle Santé publique et Surveillance, Service Maladies infectieuses au sein de la Population générale, Programme de surveillance VIH/SIDA/IST, eds. 2012.

[2] European Centre for Disease Prevention and Control/WHO Regional Office for Europe. HIV/AIDS surveillance in Europe 2011. Stockholm: European Centre for Disease Prevention and Control; 2012.

[3] Brenner BG, Roger M, Routy JP, Moisi D, Ntemgwa M, Matte C, et al. High rates of forward transmission events after acute/early HIV-1 infection. The Journal of infectious diseases. 2007 Apr 1;195(7):951-9.

[4] Chalmet K, Staelens D, Blot S, Dinakis S, Pelgrom J, Plum J, et al. Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections. BMC infectious diseases. 2010;10:262.

[5] Karlsson A, Bjorkman P, Bratt G, Ekvall H, Gisslen M, Sonnerborg A, et al. Low prevalence of transmitted drug resistance in patients newly diagnosed with HIV-1 infection in Sweden 2003-2010. PLoS One. 2012;7(3):e33484.

[6] Yerly S, von Wyl V, Ledergerber B, Boni J, Schupbach J, Burgisser P, et al. Transmission of HIV-1 drug resistance in Switzerland: a 10-year molecular epidemiology survey. AIDS (London, England). 2007 Oct 18;21(16):2223-9.

[7] Callegaro A, Svicher V, Alteri C, Lo Presti A, Valenti D, Goglio A, et al. Epidemiological network analysis in HIV-1 B infected patients diagnosed in Italy between 2000 and 2008. Infect Genet Evol. 2011 Apr;11(3):624-32.

[8] Wittkop L, Gunthard HF, de Wolf F, Dunn D, Cozzi-Lepri A, de Luca A, et al. Effect of transmitted drug resistance on virological and immunological response to initial combination antiretroviral therapy for HIV (EuroCoord-CHAIN joint project): a European multicohort study. The Lancet infectious diseases. 2011 May;11(5):363-71.

[9] Van Vaerenbergh K, Debaisieux L, De Cabooter N, Declercq C, Desmet K, Fransen K, et al. Prevalence of genotypic resistance among antiretroviral drug-naive HIV-1-infected patients in Belgium. Antiviral therapy. 2001 Mar;6(1):63-70.

[10] Derdelinckx I, Van Laethem K, Maes B, Schrooten Y, De Wit S, Florence E, et al. Current levels of drug resistance among therapy-naive HIV-infected patients have significant impact on treatment response. Journal of acquired immune deficiency syndromes (1999). 2004 Dec 15;37(5):1664-6.

[11] Vercauteren J, Derdelinckx I, Sasse A, Bogaert M, Ceunen H, De Roo A, et al. Prevalence and epidemiology of HIV type 1 drug resistance among newly diagnosed therapy-naive patients in Belgium from 2003 to 2006. AIDS research and human retroviruses. 2008 Mar;24(3):355-62.

[12] Ruelle J, Ingels MG, Jnaoui K, Ausselet N, Vincent A, Sasse A, et al. Transmission Network of an HIV Type 1 Strain with K103N in Young Belgian Patients from Different Risk Groups. AIDS research and human retroviruses. 2013 Oct;29(10):1306-9.

[13] Castro E, Khonkarly M, Ciuffreda D, Burgisser P, Cavassini M, Yerly S, et al. HIV-1 drug resistance transmission networks in southwest Switzerland. AIDS research and human retroviruses. 2010 Nov;26(11):1233-8.

[14] Skoura L, Metallidis S, Buckton AJ, Mbisa JL, Pilalas D, Papadimitriou E, et al. Molecular and epidemiological characterization of HIV-1 infection networks involving transmitted drug resistance mutations in Northern Greece. J Antimicrob Chemother. 2011 Dec;66(12):2831-7.

[15] Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, Peddada L, et al. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS (London, England). 2003 Sep 5;17(13):1871-9.

[16] UNAIDS-WHO. When and how to use assays for recent infection to estimate HIV incidence at a population level. Geneva, Switzerland; 2011.

[17] Maes B, Schrooten Y, Snoeck J, Derdelinckx I, Van Ranst M, Vandamme AM, et al. Performance of ViroSeq HIV-1 Genotyping System in routine practice at a Belgian clinical laboratory. Journal of virological methods. 2004 Jul;119(1):45-9.

[18] Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS One. 2009;4(3):e4724.

[19] Vercauteren J, Beheydt G, Prosperi M, Libin P, Imbrechts S, Camacho R, et al. Clinical evaluation of Rega 8: an updated genotypic interpretation system that significantly predicts HIV-therapy response. PLoS One. 2013;8(4):e61436.

[20] de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, et al. An automated genotyping system for analysis of HIV-1 and other microbial sequences. Bioinformatics (Oxford, England). 2005 Oct 1;21(19):3797-800.

[21] Struck D, Perez-Bercoff D, Devaux C, Schmit JC. COMET: A Novel approach to HIV-1 subtype prediction. 8th European HIV Drug Resistance Workshop; 2010 17- 19 March; Sorrento, Italy; 2010.

[22] Pineda-Pena AC, Faria NR, Imbrechts S, Libin P, Abecasis AB, Deforche K, 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. Infect Genet Evol. 2013 May 7.

[23] Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics (Oxford, England). 2006 Nov 1;22(21):2688-90.

[24] Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. Journal of virology. 1999 Jan;73(1):152-60.

[25] SPREAD. Transmission of drug-resistant HIV-1 in Europe remains limited to single classes. AIDS (London, England). 2008 Mar 12;22(5):625-35.

[26] Vercauteren J, Wensing AM, van de Vijver DA, Albert J, Balotta C, Hamouda O, et al. Transmission of drug-resistant HIV-1 is stabilizing in Europe. The Journal of infectious diseases. 2009 Nov 15;200(10):1503-8.

[27] Rhee SY, Kantor R, Katzenstein DA, Camacho R, Morris L, Sirivichayakul S, et al. HIV-1 pol mutation frequency by subtype and treatment experience: extension of the HIVseq program to seven non-B subtypes. AIDS (London, England). 2006 Mar 21;20(5):643-51.
[28] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research. 2004;32(5):1792-7.
[29] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and evolution. 2011 Oct;28(10):2731-9.
[30] Ragonnet-Cronin M, Hodcroft E, Hue S, Fearnhill E, Delpech V, Brown AJ, et al. Automated analysis of phylogenetic clusters. BMC bioinformatics. 2013;14:317.
[31] Hue S, Clewley JP, Cane PA, Pillay D. HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. AIDS (London, England). 2004 Mar 26;18(5):719-28.
[32] Frentz D, Wensing AM, Albert J, Paraskevis D, Abecasis AB, Hamouda O, et al. Limited cross-border infections in patients newly diagnosed with HIV in Europe. Retrovirology. 2013;10:36.
[33] Kaye M, Chibo D, Birch C. Phylogenetic investigation of transmission pathways of drug-resistant HIV-1 utilizing pol sequences derived from resistance genotyping. Journal of acquired immune deficiency syndromes (1999). 2008 Sep 1;49(1):9-16.
[34] Yebra G, Holguin A, Pillay D, Hue S. Phylogenetic and demographic characterization of HIV-1 transmission in Madrid, Spain. Infect Genet Evol. 2013 Mar;14:232-9.
[35] Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC evolutionary biology. 2007;7:214.
[36] Shapiro B, Rambaut A, Drummond AJ. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Molecular biology and evolution. 2006 Jan;23(1):7-9.
[37] Deforche K, Silander T, Camacho R, Grossman Z, Soares MA, Van Laethem K, et al. Analysis of HIV-1 pol sequences using Bayesian Networks: implications for drug resistance. Bioinformatics (Oxford, England). 2006 Dec 15;22(24):2975-9.
[38] Centers for Disease Control and Prevention. Revised Surveillance Case Definitions for HIV Infection Among Adults, Adolescents, and Children Aged <18 Months and for HIV Infection and AIDS Among Children Aged 18 Months to <13 Years — United States, 2008. In: MMWR, ed. Atlanta 2008:12.
[39] Frentz D, van de Vijver DA, Boucher CA, Albert J. Estimates of HIV transmitted drug resistance can be inflated due to natural sequence polymorphisms. Journal of acquired immune deficiency syndromes (1999). 2011 Dec 15;58(5):e135-7.
[40] Descamps D, Assoumou L, Chaix ML, Chaillon A, Pakianather S, de Rougemont A, et al. National sentinel surveillance of transmitted drug resistance in antiretroviral-naive chronically HIV-infected patients in France over a decade: 2001-2011. J Antimicrob Chemother. 2013 Jun 24;in press.
[41] Frentz D, Van de Vijver DA, Abecasis AB, Albert J, Hamouda O, Jorgensen LB, et al. The Epidemiology of HIV-1 Transmitted Drug Resistance: Increase in transmitted resistance to nonnucleoside reverse transcriptase inhibitors among newly diagnosed HIV-1 infections in Europe. Rotterdam, the Netherlands: Erasmus MC; 2012.
[42] Readhead AC, Gordon DE, Wang Z, Anderson BJ, Brousseau KS, Kouznetsova MA, et al. Transmitted antiretroviral drug resistance in New York State, 2006-2008: results from a new surveillance system. PLoS One. 2012;7(8):e40533.
[43] Abecasis AB, Wensing AM, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DA, et al. HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. Retrovirology. 2013 Jan 14;10(1):7.
[44] Yebra G, de Mulder M, Perez-Elias MJ, Perez-Molina JA, Galan JC, Llenas-Garcia J, et al. Increase of transmitted drug resistance among HIV-infected sub-Saharan Africans residing in Spain in contrast to the native population. PLoS One. 2011;6(10):e26757.
[45] Paraskevis D, Pybus O, Magiorkinis G, Hatzakis A, Wensing AM, van de Vijver DA, et al. Tracing the HIV-1 subtype B mobility in Europe: a phylogeographic approach. Retrovirology. 2009;6:49.
[46] Carr JK, Nadai Y, Eyzaguirre L, Saad MD, Khakimov MM, Yakubov SK, et al. Outbreak of a West African recombinant of HIV-1 in Tashkent, Uzbekistan. Journal of acquired immune deficiency syndromes (1999). 2005 Aug 15;39(5):570-5.
[47] Leigh Brown AJ, Lycett SJ, Weinert L, Hughes GJ, Fearnhill E, Dunn DT. Transmission network parameters estimated from HIV sequences for a nationwide epidemic. The Journal of infectious diseases. 2011 Nov;204(9):1463-9.

CHAPTER 6 GENERAL DISCUSSION

Different prevention and promotion campaigns are essential to decrease the number of new HIV infections. Public policies such as the United Nations Declaration of Commitment on HIV/AIDS in 2001 and the United Nations Political Declaration on HIV/AIDS in 2006 and 2011 have been setup to tackle the pandemic [1]. They include strategies to achieve specific health outcomes, sustainability, goals of coverage, resources, and elimination of HIV-related restrictions, inequalities or stigma in all countries. A first priority for low-middle-income countries is the scaling up of ART, with universal access to antiviral drugs and improvement of the continuum in HIV care [1], whereas the priority for high income countries is to decrease HIV transmission by using several additional strategies such as TasP, PrEP and PEP, supported by the HPTN 052 study that identified a 96% reduction of HIV transmission in serodiscordant couples using TasP [2]. The characterization of local HIV epidemics is needed to evaluate the impact, cost-effectiveness and sustainability of these strategies [3]. Therefore, we attempted to characterize our local epidemic in Belgium, a high-income country, and in Colombia, an upper-middle-income country [4]. In this chapter, we will discuss in more detail the differences in health care between both countries that triggered different research questions and objectives for these specific settings. Subsequently, we will discuss our main findings and will end with some global and local perspectives.

6.1 A TALE OF TWO LOCAL HIV-1 EPIDEMICS

Belgium had an estimated number of PLHIV between 16,000 and 26,000 for 2012 (prevalence ~0.3%) [1]. According to the Belgian Scientific Institute for Public Health (ISP-WIV), there were 1,227 new diagnosis in 2012 [5], which was one of the highest rates in the European Union, despite ART coverage [6]. The most frequently reported mode of transmission was heterosexual contact, accounting for 53.8% of new infections, followed by MSM for 44.2% [5].

Belgium has a privileged response against the HIV epidemic. There are nine AIDS reference centers where ART and PEP can be provided [7]. In terms of screening, 64-tests/1000 inhabitants/year are performed, however HIV diagnosis still occurred in 42% of the cases at CD4 counts less than 350 cell/mm^3 [7]. In 2011, ART was provided to 81% of the patients engaged in care, and 83% of them displayed undetectable viral load. The lost-to-follow-up was 5.5% between 2008 and 2009 and was mainly related to newly diagnosed patients or patients originating from Sub-Saharan Africa. There are seven AIDS Reference Laboratories where all positive screening tests are confirmed and routine

virological and immunological parameters can be evaluated in addition to genotypic drug resistance testing for naive and treated patients with virological failure. The strengthening of research programs and anti-stigma laws are an important measure in the recently launched national plan [7].

In Colombia, the accumulative number of people living with HIV, AIDS or AIDS-related deaths between 1983 and 2013 was 92,379 according to country report to UNAIDS from the Ministry of Health [8]. There were 8,208 new diagnoses of HIV/AIDS in 2013, mainly men (72%) and between 25 and 35 years old (35%). Sexual contact was the main risk behavior (98.3%) [9], but this information is only available until 2009 and for 44,053 cases, in which 63% were attributed to heterosexual contacts, followed by 34% bisexual or MSM contacts [10]. However, UNAIDS estimated that there were 150,000 PLHIV in 2012 (prevalence ~0.5%), 9,000 new infections and 6,500 AIDS-related deaths, suggesting underreporting in the national statistics [1].

The government of Colombia decreed the management of HIV/AIDS and other sexual transmitted diseases since 1997 [11], and a structured national plan was released in 2004 [12]. Nowadays, the health system guarantees free access to ART, although mainly generic NRTI and NNRTI are provided and brand-PI. From 2006 onwards, the national guidelines recommended ART when CD4 reached ≤ 350 cells/mm^3 with a viral load ≥ 100,000 copies/mL [13]. Recently, the thresholds were redefined in accordance to the current WHO guidelines [14, 15]. Frequent clinical follow-up and immunological and virological testing every six months are also provided. In contrast, genotypic drug resistance testing was initially only recommended at virological failure on second-line regimens [13]. Since 2012, drug resistance testing could be performed at any virological failure defined as inability to maintain viral suppression <50 copies/mL [14]. Prevention strategies have been implemented, such as free testing, free condoms for PLHIV, targeted education to and prevention surveys in vulnerable populations, specialized education of health-care professionals and anti-stigma laws [9, 16]. However, the statistics seem to show a different scenario due to the apparent underreporting of HIV infections by several administrative and health entities or the high internal displacement within the country [17]. For instance, HIV screening is guaranteed by the health system but only HIV screening in pregnant women is reported with the aim to follow-up MTCT. This revealed that the coverage of screening in pregnant women was only 60% in 2010 [9]. Additionally, the CD4 count at diagnosis was in more than 40% of the patients <200 cells/mm^3 [18]. The decentralization of the health system makes it difficult to evaluate how many centers are treating PLHIV or how many laboratories are providing viral load and CD4. Regardless of the ART scale up that started in 2000, it was estimated that still 51% of the population was in need of therapy in 2012 according to the WHO report [1], although the Ministry of Health

reported a coverage of 80% for the same year [9]. With the new recommendations for ART initiation in the WHO 2013 guidelines, the proportion of PLHIV in need of ART increased to 90% in 2012 [1]. The lost-of-follow-up during the first year was 14.2% between 2009- 2011 [18], but the retention in ART was only 53% in 2012, well below the goal of >85% set by the WHO [18, 19].

Consequently, the challenges in these two countries are currently different. Belgium, as other high-income countries, is looking forward to decrease the number of new infections based on several prevention strategies such as the guarantee of free condoms and water-and/or silicon-based lubricants for commercial gay venues or high migrant population areas, education about prevention and reinforcement of the community stakeholders that work with migrants or other vulnerable populations, improvement of the access to syringes to reduce HIV transmission in IVDU, the availability of non-occupational PEP, among other measures. Belgium has been working on a national plan against HIV/AIDS since 2012, in which one of the priorities is to support research to improve and extent existing data and to monitor if prevention strategies and care are improving the quality of life of PLHIV or the people who are at risk of HIV transmission. Other priorities include research on the feasibility of PrEP and TasP [7].

In contrast, low and middle-income countries still face the scaling-up of ART [1], where 28.6 million of people are in need of ART [15]. Therefore, the challenges are somehow different such as the elimination of MTCT, the decrease of Tuberculosis-related deaths among PLHIV, mobilization of financial resources for AIDS response, reinforcement of HIV health integration, elimination of gender inequality or stigma, and abolition of punitive laws [1]. Although one of the aims is to reduce HIV transmission by 50% by 2015, prevention strategies are generally associated with effective measures such as male circumcision, decrease of STD or condom use. TasP or PrEP are not yet recommended due to the lack of information about cost-effectiveness and the necessity to prioritize access to ART for the ones in need [20, 21]. Currently, research mainly focuses on monitoring heath care and prevention strategies. However, there is still limited information about the main drivers of the epidemic in Colombia. This information could guide the decision makers in defining the next steps in the national plan and in the research agenda.

6.2 MOLECULAR EPIDEMIOLOGY RESEARCH IN COLOMBIA AND BELGIUM

6.2.1 PERFORMANCE OF SUBTYPING TOOLS

In *chapter 2,* subtyping tools were evaluated for their performance in the classification of viral strains, an important step in molecular epidemiology research and a necessary pre-processing step in the study of viral transmission networks or viral spread. The REGA subtyping tool was designed based on the phylogenetic definition of subtypes [22] by including NJ methods and bootscanning methods [23-25]. However, REGAv2 presented several limitations such as the high number of unassigned sequences and the inclusion of CRFs only up to CRF14_BG, when more than 58 CRFs had been identified until 2013. Therefore, REGAv3 was developed with the aim to decrease the number of unassigned sequences by including reference sequences for the more recently identified CRFs and by introducing some tags to alert the user for possible recombinants, such as the assignments "-like or potential recombinant". In addition, other co-authors designed a more fashionable interface to show the summary of a given dataset in terms of frequency of subtypes, characteristics of any sequence (length, support, position in the genome), the phylogenetic statistics (average bootstrap support of the NJ trees, the average bootscan support) and the exact recombination breakpoints according to the bootscan analysis (available at http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/).

We compared REGAv3 with the most commonly used subtyping tools that are based on similarity, statistical or phylogenetic methods. We took two large datasets from a public database with full-length genomes and a clinical database from a country with high subtype variability with the aim to retrieve the highest number of non-B subtypes, which is a frequent caveat of this type of studies [26, 27]. Even though we used those datasets, we only obtained representative number of sequences for pure subtypes such as A, B, C, D, F, G and CRFs such as CRF01_AE and CRF02_AG. These subtypes are responsible for 89% of the infections in the global HIV-1 epidemic [28]. Subtype C and subtype B were identified with a sensitivity of more than 96% by Stanford, phylogenetic and statistical-based tools. These subtypes are commonly found in Oceania, India, East-Southern Africa and the Western world, respectively [28]. This suggests that subtype classifications would give similar results, independent from the subtyping tool used in the study of an epidemic. Although these two subtypes together do represent approximately 60% of the HIV-1 pandemic, several local epidemics are often characterized with completely different genetic forms. Indeed, several reports have shown the increase in non-B subtypes and recombinants because of migration and mixing of the local strains [29-34]. Since global traveling is

increasing, it is not surprising that some HIV subtypes may disappear or new subtypes may arise in a local epidemic [35]. In this respect, the subtyping tools performed differently in the identification of non-B subtypes. In general, subtyping tools like COMET, jpHMM, REGAv3 and SCUEAL (in alphabetical order) had sensitivity higher than 90% for other pure subtypes, whereas only COMET and REGAv3 had comparable sensitivity values for some CRFs such as CRF01_AE that is prevalent in South and East Asia and CRF02_AG that is prevalent in West Africa [28].

The election of a subtyping tool depends on its performance, the technical facilities at the lab and the goal of the specific analysis e.g. evaluation of recombination, surveillance purposes and clinical follow-up. In *chapter 2*, the performance and operational characteristics of automated subtyping tools were studied and summarized. We also provided some recommendations to guide people who need to classify HIV strains but are not familiar with the specific technicalities, such as the use of at least two subtyping tools for surveillance purposes due to the variable performance of subtyping tools for non-B and non-C subtypes. It is important to note this scheme has not been validated, but our results suggest this would be a more correct approach for the evaluation of viral diversity than to only rely on the assignment of one subtyping tool. For instance, we found 98% of concordance between COMET, REGAv3 and the *MPhy* in the LANL dataset (data no shown), these tools being the ones with a high sensitivity for most subtypes. In addition, we used this strategy in subsequent chapters of this thesis. In *Chapter 5,* 719 out of 778 (92.4%) sequences analyzed were concordant between COMET and REGAv3 (data no shown) in a cohort characterized by a high number of non-B subtypes [36]. The discordant sequences were mainly URFs that accounted for 4.8% of the total population.

6.2.2 VIRAL DIVERSITY AND MOBILITY OF SUBTYPE B

In *chapter 3*, we studied the viral diversity and spread of subtypes in Colombia because nationwide studies were not yet available and it was unknown if other non-B subtypes had become predominant in recent years given the increase of recombinants BF or other non-B in neighboring countries such as Brazil or Venezuela (see *chapter4*) [37-40]. Two previous studies included 115 patients from Medellin using *pol* sequences and 237 patients from Bogotá using Heteroduplex mobility assay, both performed between 2001 and 2002 [41, 42]. A later report included 35 patients from Barranquilla using *env* C2V3C3 region sequences (Figure 1.7) [43]. All studies revealed the predominance of subtype B. Also in our study, 99.8% of the sequences were subtype B despite the high migration and internal displacement in Colombia [44, 45]. Possible explanations are that infections are mainly transmitted

locally and migration occurs between countries where subtype B has been frequent such as United States, Spain and Venezuela [28, 45].

In this respect, Belgium, as well Austria and Luxembourg, were the main sinks for subtype B in Europe [46]. On the other hand, little was yet known about the dispersal of subtype B in Colombia. Therefore, we tried to investigate this in chapter 3 and we showed that the dispersal of subtype B occurred by multiple introductions in the early 1980's. This is later than the estimates for neighboring countries such as Central America or Brazil [47, 48], but it is consistent with the first AIDS case in Colombia [49]. Multiple introductions were observed between Spain and Bogotá, the capital of Colombia, which also acted as the main hub of the epidemic. Despite the high displacement in the country, Bogotá exported most of the infections to other regions. This suggests that tourism did play a role in the spread of HIV within the country. However, we lacked the transmission group and other important demographic information that could have informed us about the dynamics of the epidemic. Another possible bias in our sample was the inclusion of mainly treated patients who were in follow-up, whereas internally displaced people could have been underestimated in this analysis due to their irregular follow-up. Further research should be done to investigate to what extent the role of internal displacement or tourism is.

In *chapter 5*, the viral diversity within Belgium was also evaluated to a limited extent. Previous studies found 52% subtype B infections, followed by approximately 11% subtype A and C infections among 281 patients at the AIDS Reference Centers from Brussels and Leuven, and an increase of non-B subtypes from zero to 57% from 1983 to 2001 [36]. From 2003 to 2006, all the eight Belgian AIDS Reference Centers collaborated in the Pan-European SPREAD study, that detected the following proportions of subtypes: B (58%), CRF02_AG (11%), A (10%), C or URFs (8%), G (3%) and others (2%) (Figure 1.7). This study also suggested based upon the survey questionnaire that half of the non-B subtypes originated from Sub-Saharan Africa; however trends and more detailed phylogenetic analyses were not done [50]. In our study, we also found that half of the infections were caused by subtype B, followed by CRF02_AG and subtype C with approximately 10%. We also checked for trends of the most frequent subtypes, and this revealed that the number of subtype B infections was stable over time, in contrast to other countries in Europe such as France, Sweden, and UK [29, 30, 34]. These countries displayed an increase in non-B subtypes mainly related to immigrants originating from high-prevalent countries. This pattern seems to be expanding to other countries like Germany and UK, where MSM seem to play a role in the spread of non-B subtypes [32, 51]. In our Leuven cohort, 81% of the patients infected with subtype B originated from Belgium, and the remaining originated from Northern, Southern and Western

Europe (9.6%) whereas 2% subtype B infections originated from high prevalence countries (data no shown). Likewise, 51.6% of the individuals infected with non-B subtypes originated from sub-Saharan Africa, followed by 33.6% individuals originated from Belgium and 6.2% from Southeast Asia. Although, these findings could suggest the intermixing of the epidemic as is occurring in other settings [33, 52, 53], more detailed analyses will be required to address this matter.

We also observed an association of subtype B with TDR that could be related with the predominance of subtype B in Western Europe, a region with a history of suboptimal ART exposure [50]. It is important to bear in mind that there is no evidence that subtypes could influence the frequency of TDR, nor that the methodology used could impact the measurement as the SDRM list has already excluded the different polymorphisms present in non-B subtypes [54]. Therefore, to better understand the identification of subtype B as the main factor associated with TDR, our analysis was complemented with Bayesian network analysis to evaluate dependencies. We found that male gender was the main factor associated with TDR with low support, and gender was related to MSM and Belgian origin with 100 bootstrap support. The association of subtype and gender was also described in a pan-European study, where gender was associated with subtype via risk factor [31].

6.2.3 TRANSMITTED DRUG RESISTANCE

In *chapter 4*, we aimed to find any available data on the level and patterns of TDR in Colombia for the evaluation of the efficacy of standard first-line regimens used in the country, as patients do not receive baseline drug resistance testing [14]. One survey showed 5.8% of resistance in 2006 based upon the Stanford algorithm [55], and we found similar values in our systematic review where we interpreted all available sequences in databases at that time using the WHO SDRM 2009 list [56]. Recently, another survey in Cali showed 6.6% of resistance between 2008 and 2010 using the same list [57]. According to these results, a strict monitoring of TDR should be performed soon in this city. However, baseline drug resistance testing cannot yet be recommended because the sampling in all surveys performed up until now is not representative for the entire country. Further work is required to establish if this measure would be cost-effective and feasible given the actual statistics about monitoring and people in need of ART [1, 9, 18].

In low-middle-income countries the scaling-up of ART is still a priority, but the surveillance of TDR should be done if ART has been available for more than 3 years and TDR testing should be considered according to the detected TDR prevalence [58]. In general, routine resistance testing is not available for

naive patients, and most of the national guidelines in LAC only recommend it when virological failure occurs [59]. However, the key questions are still how well PLHIV are monitored and how the TDR results of surveillance studies reflect the actual TDR prevalence and trends in the last years. The EWIs were developed to answer the first question but unfortunately, the results were not encouraging. The low follow-up and retention in care indicators suggested a potential risk for increased drug resistance in sub-Saharan Africa, Caribbean and Central America [60-62]. In Latin America, there were conflicting results due to differences in the design of the studies, such as the use of different lists for the evaluation of TDR and inclusion criteria's. In addition, limited information on EWIs was available in South America as Colombia was the only country for which all relevant data was available [63]. In our systematic review, we found overall 7.7% TDR in Latin America and a stable trend [54]. It should be pointed out that this trend was in agreement two other systematic reviews that were released in the same time period [64, 65]. The strength of our analysis was the exclusive focus on all countries in Latin America and Caribbean, the stringent recalculation of the prevalence values based on sequences and tables, and the highlight of studies that were performed according to WHO criteria [58]. Using the 5% threshold recommended by WHO, more extensive TDR surveillance should be performed in Mexico, Brazil, and Venezuela [37, 66-68]. In some regions within Brazil and Venezuela, there was even an increase in TDR above 10%, suggesting the need for routine baseline resistance testing. New publications confirmed the increase of TDR in Brazil, but for Venezuela no recent data is available that could confirm the increase [69, 70]. Nevertheless, the new Brazilian guidelines interpreted the surveys with caution due to the heterogeneity in the methodology used in the past studies and therefore need to perform more representative studies for the entire country [71]. Meanwhile, the indications for baseline resistance testing are pregnancy and newly infected individuals with a partner under ART. Contrary to expectations, the strategy "test and treat" or TasP has been implemented, which is a big step for an upper-middle income country [71].

In *chapter 5*, we studied the Leuven cohort, where baseline drug resistance testing for naive patients was already a routine procedure at the end of the 90's, and where the retrospective evaluation of left-over plasma samples for viral load were used to guide clinicians who wanted to start therapy at a later date. We found a stable TDR trend in the overall, NRTI and PI resistance. This was in agreement to previous reports on Europe and Belgium [50, 72, 73]. This could be related to the prescribing practices of more potent ART regimens, individualization of treatment using resistance testing, and the strict follow-up of viral load and adherence. However, an up-and-down trend of NNRTI was observed with a peak in 2008. This down trend was not observed in the SPREAD study that included newly diagnosed patients from 22 European countries until 2010, and in which a peak in NNRTI TDR was observed in

similar period for which the cause has not been investigated [72, 74]. Our analyses excluded that immigrants were responsible of this peak. It coincided with a similar peak in the prevalence of K103N in MSM originating from Belgium and not any TDR trend in Sub-Saharan immigrants was observed. Our analysis also excluded the contribution of possible local transmission clusters to the NRTI trend. In this analysis, we used sequences of the entire HIV population followed at the ARC in Leuven and several controls from Belgium and other European countries. However, it is important to note that individuals unaware of their serostatus or sampled in other regions could have been part of an unidentified cluster, which is a common limitation in cluster analyses [75, 76]. Nevertheless, the ongoing transmission of the K103N mutation was observed in some patients in Gent [77] and the region of Namur, the capital of the province Namur and Wallonia, where a cluster of 29 individuals mainly naive, MSM and originated from Belgium was observed. 22 cases of them had viruses carrying the K103N mutation, but none of our sequences was part of this cluster and none of our clusters had that size [78]. These findings highlight the different dynamics within the HIV epidemic, even within countries.

The search for target populations, in which focused routine resistance testing or target prevention strategies could be performed, has drawn the attention in recent reports [79-87]. Given the complete data of the AIDS Reference Center, we performed an exhaustive analysis of transmission clusters and associated factors. We found that 42% of the TDR patients were involved in transmission clusters and mainly associated with subtype B (13 clusters), 7 of them with at least one individual originated from or sampled in other countries. The spread of TDR among non-B subtypes was limited and mainly circumscribed to pairs (3 pairs), all of them included at least one individual originating from or sampled in other country. As expected, subtype B was the factor associated with transmission clusters, but Bayesian analyses showed the dependency of this variable with MSM risk factor and Belgian origin. Therefore, the Belgian MSM population seems to play a central role in the spread of TDR, but the non-significance of these variables in the multivariate and Bayesian analyses limit our conclusions. Moreover, these results cannot be extrapolated to the entire Belgian epidemic due to the higher proportion of MSM and Belgian people in our cohort compared to the national cohort. Ultimately, our study showed the importance of baseline drug resistance testing and the retrospective evaluation of the impact on health care. For instance, 83% of the patients with TDR achieved nevertheless undetectable viral load, and none of the treated patients were involved in the local spread of TDR as indicated by the phylogenetic analysis. In addition, if baseline resistance testing had not been performed, half of the TDR patients would have received a suboptimal regimen. These findings support the impact of genotyping and the proposal of some researches to use it in low-middle-income countries [88].

6.3 PERSPECTIVES

6.3.1 PERSPECTIVES IN THE TWO LOCAL EPIDEMICS

This thesis constitutes the first step in the analysis of the two HIV epidemics here described, with the aim to formulate recommendations for public health policies, with TasP being one of the promising strategies thanks to the availability of better and tolerable regimens, the effect of HIV on chronic endothelial inflammation and the impact on the prevention of new infections according to the HPTN 052 trial [2]. We already achieved a complete view on TDR and the contribution of transmission clusters in Leuven, Belgium. However, several questions remain. For instance, how transmission clusters, irrespective of the presence of TDR, contribute to the spread of HIV in Leuven and Belgium: does the epidemic have super-transmitters or is transmission mainly related to small and local clusters, are mainly recently infected or chronically infected individuals involved, are these patients already diagnosed or engaged in care at the time of transmission, are migrants participating in the viral spread, are the infections mainly acquired in Belgium or are they imported, is the migrant epidemic intermixed with the local Belgian epidemic, and what are the demographic and clinical characteristics of these patients. These questions are currently being analyzed in the Leuven cohort using phylogenetic analyses, but this analysis should be enlarged to the country level. For this, the participation of all ARC and ARL are needed and plans are made to hopefully perform this in the near future within the framework of the Belgium Research on AIDS and HIV Consortium (BREACH) (available at http://www.breach-hiv.be/). To complete these studies, the next steps should ideally include behavioral data such as condom use, substance use, serosorting, concurrency, sexual mixing [89]; as well as the engagement in care and the possible barriers to access health care. Following this, mathematical models could be developed based on these parameters to predict how the epidemic is going to evolve and what the impact of each prevention strategy could be [90]. Finally, cost-effectiveness analyses will be needed to evaluate if particular prevention strategies should be implemented at country level [91].

Regarding the Colombian epidemic, some other priorities should be first addressed before similar research can be carried out. The first step is to improve the quantity and quality of the HIV database with the inclusion of as much demographic, clinical and laboratory data and to complete it with tools that help the management of sequences in clinical practice, such as is possible with Rega DB [92]. Then the focus should be on the basic characterization of the epidemic in terms of the health care cascade, such as possible barriers to HIV diagnosis and the characterization of the population with late diagnosis, retention in care, therapy adherence and suppression of viral load. The surveillance of TDR should be

also in the list of priorities according to the WHO recommendations [93] and the result of our systematic review. However, the surveillance of secondary resistance may be useful to evaluate if multiple resistance is disappearing as in high-income countries and might be more readily achievable as resistance testing is more commonly performed in failing patients [94-96]. Finally, behavioral studies can complement the description of the epidemic and evaluate how other strategies such as condom use can be increased to tackle the HIV epidemic.

6.3.2 GENERAL CHALLENGES IN THE DYNAMICS OF THE HIV EPIDEMIC

Although in Europe resistance remains stable and several therapeutic alternatives are available [72, 73, 94-96], resistance still constitutes a threat in low-middle-income countries where around 10 million of people are currently receiving ART [88]. Several reports have shown a concomitant rise in resistance in low income countries [65, 97, 98]. The inability to perform viral load in most of the centers, the difficulties in the integration of monitoring services, the lack of trained health-care personal and limited access and options for first-line therapies among other factors have contributed to the incredible profile of resistance in some African settings. On-going administration of virological failing therapies have impaired options for second line therapies as no other alternatives are yet available in these settings [99]. This problem in low-middle income countries will probably be transferred to high-income countries due to tourism and migration [88]. Several stakeholders have called the attention to public health entities to tackle this situation, but up to now the first measure was the compulsory implementation of viral load in these settings, the continuous surveillance of EWIs and the periodic performance of drug resistance surveys [15, 19, 93]. Therefore, a global challenge is the set-up of complete health-care facilities to monitor therapy and response in low-middle-income countries and not only to prioritize the scaling-up of ART.

The monitoring of the health care cascade in LAC countries constitutes a priority. For instance, diagnosis and linkage to care should be improved given the frequency of late presenters ranging from 56% in Argentina to 91% in Haiti when the definition of CD4 count ≤ 200 cell/mm at baseline or AIDS-defining illnesses prior to or at ART initiation is used [100-102]. Moreover, the low retention in care and other EWIs suggest that the surveillance of resistance should be prioritized [61, 63]. According to our studies, the prevalence of TDR was still stable in LAC and the pattern of mutations was relatively similar to other high-income countries. However, some limited data suggest an increase and therefore, surveillance of TDR should be more profoundly monitored to enable the implementation of adequate

measures if needed. Notorious improvements in TDR surveys have been observed in African and Caribbean countries [20], albeit this improvement was not observed in some Central American and Andean region countries. Finally, studies in vulnerable populations or populations characterized with a high HIV prevalence should be encouraged. For instance, MSM have a 33.3 times higher risk of being infected with HIV [103, 104], vulnerable population such as females [105, 106], seasonal immigrants like the *bateyes* in Dominican Republic (Haitian descendents working in sugar canes); ethnic groups such as the Mayans in Guatemala, other several indigenous or rural communities in the Andean Region and, the *Garifunas* (Nigerian descendents) in Honduras or other countries in Central America are at socio-economical risk [107]. It is worth mentioning that the latter group may have a higher level of TDR than the general population in Honduras, as indicated by a small study [108].

A lot of resources are devoted to study resistance but there are still some improvements to make. In terms of TDR, the SDRM list will be updated soon [109]. In this new version, some controversial mutations like the 46IL (also discussed in our studies) and 85V (was present in 1.3% of the cases in the Leuven cohort and none of the sequences in the LAC analysis) will be excluded because they occur as a polymorphisms, whereas other RT or PI mutations should be included [109]. On the long-term, a similar effort for the assessment of INI TDR should be done as more and more patients in high-income countries will be treated with this new drug class. These drugs have excellent tolerance and with the new data on dolutegravir, this latter drug will be part of many first-line regimens. Additionally, some authors are very optimistic about its use in prevention and ART simplification strategies [110, 111]. In general, the surveillance of TDR and acquired resistance should be continued and collaborative research is still needed to better understand the resistance profile of new drugs. In addition, NGS technologies are also opening a door to evaluate resistance against all drug classes in one test, and to understand viral evolution and drug resistance development in the context of viral protein interactions and to investigate the impact of minority variants on response [112-114].

There are also methodological challenges for the near future. One of them is the rapid and accurate classification of subtypes of large datasets, IN, full-genome and NGS sequences. Automatic subtyping tools require continuous updates given the increasing number of recombinants in high-income countries, the intermixing of variants plus the scaling-up of facilities to perform resistance testing in low and middle-income countries. Special attention should go to their improvement in the correct assignment of non-B and non-C subtypes. Some subtyping tools would require the expansion of the

coverage up to full-length genome, given the increasing performance of drug resistance testing against INI in clinical practice and of next generation sequencing (NGS). Although NGS is not yet an alternative as Sanger sequencing in clinical practice, it is broadly used in studies investigating host-virus interactions, minority variants with drug resistance mutations and vaccines [115]. Another methodological challenge is the clear definition of transmission clusters as disagreements could influence the results and conclusions in prevention studies or clinical trials [116, 117]. In our research, we used several definitions to assess the impact of the thresholds in genetic distance on our results. However, it will be important to reach a consensus in this research field to enable the comparisons of different studies. Two groups have designed software to evaluate transmission clusters and our group is currently working on a similar script [118, 119]. This will facilitate us a flexible framework to compare different definitions of clusters and link the results with clinical data. Additionally, the directionality of transmission is still difficult to assess and is important information. At this moment, at least longitudinal testing or time-scaled phylogenies need to be performed to make some conclusions [117, 120]. Other needs for standardization in definitions to enable the comparison of studies in different settings, are situated at the level of sequencing technology (Sanger sequencing or NGS), the region investigated (full-length genome, *env* or *pol*), the measurement of acute or early infection, sampling density, dealing with missing data [117].

6.4 CONCLUSION

In conclusion, we have described the viral diversity and TDR in the cohorts of Leuven, Belgium and Colombia. The study of viral diversity should preferentially be performed with COMET or REGAv3 automated subtyping tools because their good performance and operational characteristics. However, the analysis of short sequences like PR requires a more extensive phylogenetic analysis. The TDR prevalence and viral diversity between 1998 and 2012 was stable in Leuven, although a parabolic trend for NNRTI resistance was observed. No specific population was significantly associated with TDR or the local spread of TDR, although MSM originating from Belgium do play a central role. The number of clusters with TDR patients and the number of chronically infected individuals involved suggest the need of a national survey to formulate public health policies in prevention of HIV transmission and promoting early diagnosis. On the other hand, LAC displayed also a stable TDR trend but the lack of data for Central America and Andean region emphasizes the need for TDR surveillance in these regions. Although the available data of Colombia showed a TDR prevalence around 6%, a national survey

according to the WHO guidelines should be considered as an imperative step in the research agenda. The viral diversity remained dominated by subtype B until 2007. Additional phylogenetic analyses suggested that Bogotá was the main hub of the epidemic, whereas other large cities such as Medellin and Cali were mainly sinks. These findings suggest that tourism plays a role in the spread of HIV, but further research will be needed to understand what are the populations involved in the spread of HIV at the country level. This thesis shaded some light on the Leuven and Colombian epidemic, and some results could help to formulate effective prevention policies in the future.

REFERENCES

[1] UNAIDS-WHO. GLOBAL REPORT: UNAIDS report on the global AIDS epidemic. Geneva; 2013.

[2] Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 infection with early antiretroviral therapy. The New England journal of medicine. 2011 Aug 11;365(6):493-505.

[3] Cambiano V, O'Connor J, Phillips AN, Rodger A, Lodwick R, Pharris A, et al. Antiretroviral therapy for prevention of HIV transmission: implications for Europe. Euro Surveill. 2013;18(48):20647.

[4] The World Bank. How we classify countries: Country and Lending Groups. Data 2013 July 1 [cited 2014 Feb]; Available from: http://data.worldbank.org/about/country-classifications/country-and-lending-groups

[5] Belgian Scientific Institute for Public Health (ISP-WIV). Epidemiologie du SIDA et de l'infection à VIH en Belgique. In: Direction opérationnelle Santé publique et Surveillance, Service Maladies infectieuses dans la Population générale, Programme de surveillance VIH/SIDA/IST, eds. Bruxelles, Belgique 2014.

[6] Likatavicius G, Van de Laar M. HIV and AIDS in the European Union, 2011. Euro Surveill. 2012;17(48).

[7] Ministry of Social Affairs and Health. Belgian National HIV/AIDS Plan 2014-2019. Belgium 2013.

[8] Ministerio de Salud y Protección Social, Instituto Nacional de Salud. Informe GARPR: Seguimiento de la Declaracion de compromiso sobre el VIH/sida. Bogota, Colombia: Ministerio de la Proteccion Social; 2014.

[9] Ministerio de la Protección Social. Informe UNGASS: Seguimiento de la Declaracion de compromiso sobre el VIH/sida. Bogota, Colombia: Ministerio de la Proteccion Social; 2012.

[10] Ministerio de la Protección Social. Informe UNGASS: Seguimiento de la Declaracion de compromiso sobre el VIH/sida. Bogota, Colombia: Ministerio de la Proteccion Social; 2010.

[11] Ministerio de la Protección Social. Decreto Numero 1543 de 1997: Manejo de la infección por el Virus de Inmunodeficiencia Humana (VIH), el Síndrome de la Inmunodeficiencia Adquirida (SIDA) y las otras Enfermedades de Transmisión Sexual (ETS). In: Ministerio de la Protección Social, Ministerio de Justicia y del Derecho, Ministerio de Educación, eds. República de Colombia 1997:16.

[12] Ministerio de Protección Social, Instituto Nacional de Salud. Plan Intersectorial de Respuesta ante el VIH/Sida, 2004-2007, Colombia. República de Colombia 2004.

[13] Ministerio de la Protección Social. Guía para el manejo de VIH/sida basada en la evidencia Colombia. In: Ministerio de la Protección Social, ed. Republica de Colombia: Available at: http://www.minproteccionsocial.gov.co/VBeContent/library/documents/DocNewsNo16914DocumentNo6250.pdf. 2006.

[14] Ministerio de la Protección Social, Asociación Colombiana de Infectología (ACIN), United Nations Population Fund (UNFPA). Documentos preliminares Guías de practica clínica para la atención integral del VIH/sida para adolescentes, adultos y niños. Republica de Colombia; 2012.

[15] WHO. Consolidated Guidelines on the use of Antiretroviral Drugs for Treating and Preventing HIV infection: Recommendations for a Public Health approach. Geneva, Switzerland: World Health Organization 2013.

[16] Ministerio de Protección Social, UNAIDS. Plan Nacional de Respuesta ante el VIH y el SIDA, Colombia, 2008-2011. República de Colombia 2008.

[17] OPS/OMS, UNAIDS. Análisis de la situación de los sistemas de Monitoreo y Evaluación de la Respuesta al VIH en los países Andinos y Chile; 2008.

[18] PAHO. Antiretroviral Treatment in the Spotlight: A Public Health Analysis in Latin America and the Caribbean. Washington, DC; 2013.

[19] WHO. Meeting Report on assessment of World Health Organization HIV drug resistance Early Warning Indicators: report of the Early Advisory Indicator Panel Meeting. Geneva, Switzerland; 2012.

[20] Bertagnolio S, Perno CF, Vella S, Pillay D. The Impact of HIV Drug Resistance on the Selection of First- and Second-Line ART in Resource-Limited Settings. The Journal of infectious diseases. 2013 Jun 15;207 Suppl 2:S45-8.

[21] Gupta RK, Wainberg MA, Brun-Vezinet F, Gatell JM, Albert J, Sonnerborg A, et al. Oral antiretroviral drugs as public health tools for HIV prevention: global implications for adherence, drug resistance, and the success of HIV treatment programs. The Journal of infectious diseases. 2013 Jun 15;207 Suppl 2:S101-6.

[22] Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, et al. HIV-1 nomenclature proposal. Science (New York, NY. 2000 Apr 7;288(5463):55-6.

[23] Posada D, Crandall KA. Selecting models of nucleotide substitution: an application to human immunodeficiency virus 1 (HIV-1). Molecular biology and evolution. 2001 Jun;18(6):897-906.

[24] Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. Journal of virology. 1999 Jan;73(1):152-60.

[25] Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution. 1987 Jul;4(4):406-25.

[26] Holguin A, Lopez M, Soriano V. Reliability of rapid subtyping tools compared to that of phylogenetic analysis for characterization of human immunodeficiency virus type 1 non-B subtypes and recombinant forms. Journal of clinical microbiology. 2008 Dec;46(12):3896-9.

[27] Yebra G, de Mulder M, Martin L, Perez-Cachafeiro S, Rodriguez C, Labarga P, et al. Sensitivity of seven HIV subtyping tools differs among subtypes/recombinants in the Spanish cohort of naive HIV-infected patients (CoRIS). Antiviral Res. 2010 Jan;89(1):19-25.

[28] Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS (London, England). 2011 Mar 13;25(5):679-89.

[29] Chaix ML, Seng R, Frange P, Tran L, Avettand-Fenoel V, Ghosn J, et al. Increasing HIV-1 non-B subtype primary infections in patients in France and effect of HIV subtypes on virological and immunological responses to combined antiretroviral therapy. Clin Infect Dis. 2013 Mar;56(6):880-7.

[30] The UK Collaborative Group on HIV Drug Resistance. The increasing genetic diversity of HIV-1 in the UK, 2002-2010. AIDS (London, England). 2014 Nov 19;28(5):773-80.
[31] Abecasis AB, Wensing AM, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DA, et al. HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. Retrovirology. 2013 Jan 14;10(1):7.
[32] Somogyi S, Meixenberger K, Bartmeyer B, Jansen K, Scheufele R, Dupke S, et al. Continued increase of the prevalence of HIV-1 non-B subtypes: Update from the German HIV-1 Seroconverter Study. 12th European Workshop on HIV & Hepatitis-Treatment Strategies & Antiviral Drug Resistance; 2014; Barcelona, Spain: Reviews in Antiviral Therapy & Infectious Diseases; 2014.
[33] von Wyl V, Kouyos RD, Yerly S, Boni J, Shah C, Burgisser P, et al. The role of migration and domestic transmission in the spread of HIV-1 non-B subtypes in Switzerland. The Journal of infectious diseases. 2011 Oct 1;204(7):1095-103.
[34] Karlsson A, Bjorkman P, Bratt G, Ekvall H, Gisslen M, Sonnerborg A, et al. Low prevalence of transmitted drug resistance in patients newly diagnosed with HIV-1 infection in Sweden 2003-2010. PLoS One. 2012;7(3):e33484.
[35] Rambaut A, Posada D, Crandall KA, Holmes EC. The causes and consequences of HIV evolution. Nat Rev Genet. 2004 Jan;5(1):52-61.
[36] Snoeck J, Van Laethem K, Hermans P, Van Wijngaerden E, Derdelinckx I, Schrooten Y, et al. Rising prevalence of HIV-1 non-B subtypes in Belgium: 1983-2001. Journal of acquired immune deficiency syndromes (1999). 2004 Mar 1;35(3):279-85.
[37] Castillo J, Comegna M, Quijada W, Jauvin V, Pinson P, Masquelier B, et al. Surveillance of HIV type 1 drug resistance among naive patients from Venezuela. AIDS research and human retroviruses. 2009 Dec;25(12):1329-33.
[38] Guimaraes ML, Velarde-Dunois KG, Segurondo D, Morgado MG. The HIV-1 epidemic in Bolivia is dominated by subtype B and CRF12_BF "family" strains. Virology journal. 2012;9:19.
[39] Brindeiro RM, Diaz RS, Sabino EC, Morgado MG, Pires IL, Brigido L, et al. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. AIDS (London, England). 2003 May 2;17(7):1063-9.
[40] Sprinz E, Netto EM, Patelli M, Lima JS, Furtado JJ, da Eira M, et al. Primary antiretroviral drug resistance among HIV type 1-infected individuals in Brazil. AIDS research and human retroviruses. 2009 Sep;25(9):861-7.
[41] Montano SM, Sanchez JL, Laguna-Torres A, Cuchi P, Avila MM, Weissenbacher M, et al. Prevalences, genotypes, and risk factors for HIV transmission in South America. Journal of acquired immune deficiency syndromes (1999). 2005 Sep 1;40(1):57-64.
[42] Sanchez GI, Bautista CT, Eyzaguirre L, Carrion G, Arias S, Sateren WB, et al. Molecular epidemiology of human immunodeficiency virus-infected individuals in Medellin, Colombia. The American journal of tropical medicine and hygiene. 2006 Apr;74(4):674-7.
[43] Villarreal JL, Gutierrez J, Palacio L, Penuela M, Hernandez R, Lemay G, et al. Characterization of HIV type 1 envelope sequence among viral isolates circulating in the northern region of Colombia, South America. AIDS research and human retroviruses. 2012 Dec;28(12):1779-83.
[44] IDMC. Colombia: Rate of new displacement highest in two decades. Geneva, Switzerland: IDMC- Internal Displacement Monitoring Centre; 2008 17 October.
[45] OIM. Perfil Migratorio de Colombia. Bogotá, Colombia; 2010.
[46] Paraskevis D, Pybus O, Magiorkinis G, Hatzakis A, Wensing AM, van de Vijver DA, et al. Tracing the HIV-1 subtype B mobility in Europe: a phylogeographic approach. Retrovirology. 2009;6:49.
[47] Bello G, Guimaraes ML, Morgado MG. Evolutionary history of HIV-1 subtype B and F infections in Brazil. AIDS (London, England). 2006 Mar 21;20(5):763-8.
[48] Murillo W, Veras N, Prosperi M, de Rivera IL, Paz-Bailey G, Morales-Miranda S, et al. A Single Early Introduction of HIV-1 Subtype B into Central America Accounts for Most Current Cases. Journal of virology. 2013 Jul;87(13):7463-70.
[49] Ministerio de la Protección Social. Informe Nacional sobre los progresos realizados en la aplicación del UNGASS. Bogotá, Colombia; 2008.
[50] Vercauteren J, Derdelinckx I, Sasse A, Bogaert M, Ceunen H, De Roo A, et al. Prevalence and epidemiology of HIV type 1 drug resistance among newly diagnosed therapy-naive patients in Belgium from 2003 to 2006. AIDS research and human retroviruses. 2008 Mar;24(3):355-62.
[51] Fox J, Castro H, Kaye S, McClure M, Weber JN, Fidler S. Epidemiology of non-B clade forms of HIV-1 in men who have sex with men in the UK. AIDS (London, England). 2013 Sep 24;24(15):2397-401.
[52] Xiridou M, van Veen M, Coutinho R, Prins M. Can migrants from high-endemic countries cause new HIV outbreaks among heterosexuals in low-endemic countries? AIDS (London, England). 2010 Aug 24;24(13):2081-8.
[53] Yebra G, Holguin A, Pillay D, Hue S. Phylogenetic and demographic characterization of HIV-1 transmission in Madrid, Spain. Infect Genet Evol. 2013 Mar;14:232-9.
[54] Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS One. 2009;4(3):e4724.
[55] Diazgranados CA. Vigilancia de Resistencia en Pacientes No Expuestos y en Pacientes Expuestos. Bogotá D.C., Colombia 2008.
[56] Pineda-Pena AC, Bello DC, Sussmann O, Vandamme AM, Vercauteren J, van Laethem K, et al. HIV-1 transmitted drug resistance in Latin America and the Caribbean: what do we know? AIDS reviews. 2012 Oct-Dec;14(4):256-67.
[57] Galindo-Orrego P, Mueses-Marin HF, Galindo-Quintero J, Martinez-Cajas JL. [Transmitted HIV Resistance in ART-Naive Patients in Cali, Colombia]. Infectio. 2013;17(1):19-27.
[58] Bennett DE, Bertagnolio S, Sutherland D, Gilks CF. The World Health Organization's global strategy for prevention and assessment of HIV drug resistance. Antiviral therapy. 2008;13 Suppl 2:1-13.
[59] AIDSTAR-One HIV Prevention Knowledge database. National Treatment Guidelines. Focus Areas: Treatment 2013 [cited 2014 Feb]; Available from: http://www.aidstar-one.com/focus_areas/treatment

[60] Bennett DE, Jordan MR, Bertagnolio S, Hong SY, Ravasi G, McMahon JH, et al. HIV drug resistance early warning indicators in cohorts of individuals starting antiretroviral therapy between 2004 and 2009: World Health Organization global report from 50 countries. Clin Infect Dis. 2012 May;54 Suppl 4:S280-9.

[61] Jack N, Ravasi G, Schrooten W, Sutherland D, Ghidinelli M, Del Riego A. Implementing early-warning indicators of HIV drug resistance in the Caribbean. Clin Infect Dis. 2012 May;54 Suppl 4:S290-3.

[62] Ravasi G, Pérez-Rosales MD, Mejía Cruz VM, Zepeda K, Pérez ME, Aguilar R, et al. WHO HIV drug resistance early warning indicators as treatment program monitoring and evaluation tool in Central American countries. *6th IAS Conference on HIV Pathogenesis, Treatment and Prevention: Abstract no TUAD0105* Rome, Italy 2011.

[63] Ravasi G, Jack N, Alonso Gonzalez M, Sued O, Perez-Rosales MD, Gomez B, et al. Progress of implementation of the World Health Organization strategy for HIV drug resistance control in Latin America and the Caribbean. Revista panamericana de salud publica = Pan American journal of public health. 2011 Dec;30(6):657-62.

[64] Frentz D, Boucher CA, van de Vijver DA. Temporal changes in the epidemiology of transmission of drug-resistant HIV-1 across the world. AIDS reviews. 2012 Jan-Mar;14(1):17-27.

[65] Gupta RK, Jordan MR, Sultan BJ, Hill A, Davis DH, Gregson J, et al. Global trends in antiretroviral resistance in treatment-naive individuals with HIV after rollout of antiretroviral treatment in resource-limited settings: a global collaborative study and meta-regression analysis. Lancet. 2012 Oct 6;380(9849):1250-8.

[66] Rangel HR, Garzaro DJ, Torres JR, Castro J, Suarez JA, Naranjo L, et al. Prevalence of antiretroviral drug resistance among treatment-naive and treated HIV-infected patients in Venezuela. Memorias do Instituto Oswaldo Cruz. 2009 May;104(3):522-5.

[67] Avila-Rios S, Garcia-Morales C, Garrido-Rodriguez D, Ormsby CE, Hernandez-Juan R, Andrade-Villanueva J, et al. National prevalence and trends of HIV transmitted drug resistance in Mexico. PLoS One. 2011;6(11):e27812.

[68] De Moraes Soares CMP, Vergara TRC, Sucupira MC, Brites C, Brito JDU, Caseiro MM, et al. Prevalence of Transmitted HIV-1 Antiretroviral Resistance among Patients Initiating ART in Brazil: A Surveillance Using Dried Blood Spots. *18th Conference on Retroviruses and Opportunistic Infections*. Boston, United States 2011.

[69] Alencar CS, Sabino EC, Carvalho SM, Leao SC, Carneiro-Proietti AB, Capuani L, et al. HIV genotypes and primary drug resistance among HIV-seropositive blood donors in Brazil: role of infected blood donors as sentinel populations for molecular surveillance of HIV. Journal of acquired immune deficiency syndromes (1999). 2013 Jul 1;63(3):387-92.

[70] Velasco-de-Castro CA, Grinsztejn B, Veloso VG, Bastos FI, Pilotto JH, Fernandes N, et al. HIV-1 diversity and drug resistance mutations among people seeking HIV diagnosis in voluntary counseling and testing sites in Rio de Janeiro, Brazil. PLoS One. 2014;9(1):e87622.

[71] Ministério da Saúde. Protocolo Clínico e Diretrizes Terapêuticas para Manejo da Infecção pelo HIV em Adultos. In: Departamento de DST AeHV, ed. Brasilia, Brazil: Ministério da Saúde 2013.

[72] Frentz D, Van de Vijver DA, Abecasis AB, Albert J, Hamouda O, Jorgensen LB, et al. The Epidemiology of HIV-1 Transmitted Drug Resistance: Increase in transmitted resistance to nonnucleoside reverse transcriptase inhibitors among newly diagnosed HIV-1 infections in Europe. Rotterdam, the Netherlands: Erasmus MC; 2012.

[73] Vercauteren J, Wensing AM, van de Vijver DA, Albert J, Balotta C, Hamouda O, et al. Transmission of drug-resistant HIV-1 is stabilizing in Europe. The Journal of infectious diseases. 2009 Nov 15;200(10):1503-8.

[74] Hofstra M, Sauvageot N, Albert J, Alexiev I, Garcia F, Struck D, et al. Prevalence of resistance mutations associated with decreased susceptibility to rilpivirine in antiretroviral naive patients from Europe. 12th European Workshop on HIV & Hepatitis; 2014; Barcelona, Spain: Reviews in Antiviral Therapy & Infectious Diseases; 2014.

[75] Abecasis AB, Geretti AM, Albert J, Power L, Weait M, Vandamme AM. Science in court: the myth of HIV fingerprinting. The Lancet infectious diseases. 2011 Feb;11(2):78-9.

[76] Brenner B, Wainberg MA, Roger M. Phylogenetic inferences on HIV-1 transmission: implications for the design of prevention and treatment interventions. AIDS (London, England). 2013 Apr 24;27(7):1045-57.

[77] Chalmet K, Staelens D, Blot S, Dinakis S, Pelgrom J, Plum J, et al. Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections. BMC infectious diseases. 2010;10:262.

[78] Ruelle J, Ingels MG, Jnaoui K, Ausselet N, Vincent A, Sasse A, et al. Transmission Network of an HIV Type 1 Strain with K103N in Young Belgian Patients from Different Risk Groups. AIDS research and human retroviruses. 2013 Oct;29(10):1306-9.

[79] Skoura L, Metallidis S, Buckton AJ, Mbisa JL, Pilalas D, Papadimitriou E, et al. Molecular and epidemiological characterization of HIV-1 infection networks involving transmitted drug resistance mutations in Northern Greece. J Antimicrob Chemother. 2011 Dec;66(12):2831-7.

[80] Audelin AM, Gerstoft J, Obel N, Mathiesen L, Laursen A, Pedersen C, et al. Molecular phylogenetics of transmitted drug resistance in newly diagnosed HIV Type 1 individuals in Denmark: a nation-wide study. AIDS research and human retroviruses. 2011 Dec;27(12):1283-90.

[81] Brenner BG, Roger M, Moisi DD, Oliveira M, Hardy I, Turgel R, et al. Transmission networks of drug resistance acquired in primary/early stage HIV infection. AIDS (London, England). 2008 Nov 30;22(18):2509-15.

[82] Buskin SE, Ellis GM, Pepper GG, Frenkel LM, Pergam SA, Gottlieb GS, et al. Transmission cluster of multiclass highly drug-resistant HIV-1 among 9 men who have sex with men in Seattle/King County, WA, 2005-2007. Journal of acquired immune deficiency syndromes (1999). 2008 Oct 1;49(2):205-11.

[83] Callegaro A, Svicher V, Alteri C, Lo Presti A, Valenti D, Goglio A, et al. Epidemiological network analysis in HIV-1 B infected patients diagnosed in Italy between 2000 and 2008. Infect Genet Evol. 2011 Apr;11(3):624-32.

[84] Castor D, Low A, Evering T, Karmon S, Davis B, Figueroa A, et al. Transmitted drug resistance and phylogenetic relationships among acute and early HIV-1-infected individuals in New York City. Journal of acquired immune deficiency syndromes (1999). 2012 Sep 1;61(1):1-8.

[85] Drescher SM, von Wyl V, Yang WL, Boni J, Yerly S, Shah C, et al. Treatment-naive individuals are the major source of transmitted HIV-1 drug resistance in men who have sex with men in the Swiss HIV Cohort Study. Clin Infect Dis. 2013 Jan;58(2):285-94.

[86] Grgic I, Lepej SZ, Lunar MM, Poljak M, Vince A, Vrakela IB, et al. The prevalence of transmitted drug resistance in newly diagnosed HIV-infected individuals in Croatia: the role of transmission clusters of men who have sex with men carrying the T215S surveillance drug resistance mutation. AIDS research and human retroviruses. 2013 Feb;29(2):329-36.
[87] Yerly S, Junier T, Gayet-Ageron A, Amari EB, von Wyl V, Gunthard HF, et al. The impact of transmission clusters on primary drug resistance in newly diagnosed HIV-1 infection. AIDS (London, England). 2009 Jul 17;23(11):1415-23.
[88] Hamers RL, Kityo C, Lange JM, de Wit TF, Mugyenyi P. Global threat from drug resistant HIV in sub-Saharan Africa. BMJ (Clinical research ed. 2012;344:e4159.
[89] DeGruttola V, Smith DM, Little SJ, Miller V. Developing and evaluating comprehensive HIV infection control strategies: issues and challenges. Clin Infect Dis. 2010 May 15;50 Suppl 3:S102-7.
[90] The HIV Modelling Consortium Treatment as Prevention Editorial Writing Group. HIV treatment as prevention: models, data, and questions--towards evidence-based decision-making. PLoS medicine. 2012;9(7):e1001259.
[91] Meyer-Rath G, Over M. HIV treatment as prevention: modelling the cost of antiretroviral treatment--state of the art and future directions. PLoS medicine. 2012;9(7):e1001247.
[92] Libin P, Beheydt G, Deforche K, Imbrechts S, Ferreira F, Van Laethem K, et al. RegaDB: community-driven data management and analysis for infectious diseases. Bioinformatics (Oxford, England). 2013 Jun 1;29(11):1477-80.
[93] WHO. World Health Organization Protocol for Surveillance of Transmitted HIV Drug Resistance: 2012 update. Geneva, Switzerland; 2012.
[94] Castagliola D, Ledergerber B, Torti C, van Sighem A, Podzamczer D, Mocroft A, et al. Trends in virological and clinical outcomes in individuals with HIV-1 infection and virological failure of drugs from three antiretroviral drug classes: a cohort study. The Lancet infectious diseases. 2012 Feb;12(2):119-27.
[95] De Luca A, Dunn D, Zazzi M, Camacho R, Torti C, Fanti I, et al. Declining prevalence of HIV-1 drug resistance in antiretroviral treatment-exposed individuals in Western Europe. The Journal of infectious diseases. 2013 Apr 15;207(8):1216-20.
[96] Vercauteren J, Theys K, Carvalho AP, Valadas E, Duque LM, Teofilo E, et al. The demise of multidrug-resistant HIV-1: the national time trend in Portugal. J Antimicrob Chemother. 2013 Apr;68(4):911-4.
[97] Hamers RL, Wallis CL, Kityo C, Siwale M, Mandaliya K, Conradie F, et al. HIV-1 drug resistance in antiretroviral-naive individuals in sub-Saharan Africa after rollout of antiretroviral therapy: a multicentre observational study. The Lancet infectious diseases. 2011 Oct;11(10):750-9.
[98] Ndembi N, Hamers RL, Sigaloff KC, Lyagoba F, Magambo B, Nanteza B, et al. Transmitted antiretroviral drug resistance among newly HIV-1 diagnosed young individuals in Kampala. AIDS (London, England). 2011 Apr 24;25(7):905-10.
[99] Hamers RL, Sigaloff KC, Kityo C, Mugyenyi P, de Wit TF. Emerging HIV-1 drug resistance after roll-out of antiretroviral therapy in sub-Saharan Africa. Curr Opin HIV AIDS. 2013 Jan;8(1):19-26.
[100] Tuboi SH, Schechter M, McGowan CC, Cesar C, Krolewiecki A, Cahn P, et al. Mortality during the first year of potent antiretroviral therapy in HIV-1-infected patients in 7 sites throughout Latin America and the Caribbean. Journal of acquired immune deficiency syndromes (1999). 2009 Aug 15;51(5):615-23.
[101] Crabtree-Ramírez B, Caro-Vega Y, Shepherd BE, Wehbe F, Cesar C, Cortes C, et al. Cross-Sectional Analysis of Late HAART Initiation in Latin America and the Caribbean: Late Testers and Late Presenters. PLoS One. 2011;6(5):e20272.
[102] McGowan CC, Cahn P, Gotuzzo E, Padgett D, Pape JW, Wolff M, et al. Cohort Profile: Caribbean, Central and South America Network for HIV research (CCASAnet) collaboration within the International Epidemiologic Databases to Evaluate AIDS (IeDEA) programme. International journal of epidemiology. 2007 Oct;36(5):969-76.
[103] Baral S, Sifakis F, Cleghorn F, Beyrer C. Elevated risk for HIV infection among men who have sex with men in low- and middle-income countries 2000-2006: a systematic review. PLoS medicine. 2007 Dec;4(12):e339.
[104] Beyrer C, Baral SD, van Griensven F, Goodreau SM, Chariyalertsak S, Wirtz AL, et al. Global epidemiology of HIV infection in men who have sex with men. Lancet. 2012 Jul 28;380(9839):367-77.
[105] UNAIDS-WHO-UNICEF. Challenges Posed by the HIV Epidemic in Latin America and the Caribbean. Lima-Peru: WHO; 2009.
[106] Soto-Ramirez LE. HIV/AIDS in Latin America. Science (New York, NY. 2008 Jul 25;321(5888):465.
[107] Cohen J. HIV/AIDS: Latin America & Caribbean. Overview: the overlooked epidemic. Science (New York, NY. 2006 Jul 28;313(5786):468-9.
[108] Murillo W, Paz-Bailey G, Morales S, Monterroso E, Paredes M, Dobbs T, et al. Transmitted drug resistance and type of infection in newly diagnosed HIV-1 individuals in Honduras. J Clin Virol. 2010 Dec;49(4):239-44.
[109] WHO. World Health Organization Annual HIV drug resistance (HIVResNet) Meeting. In: Wainberg MA, De Luca A, editors.; 2014; Boston, USA; 2014.
[110] Colasanti J, Marconi VC, Taiwo B. Antiretroviral reduction: is it time to rethink the unthinkable? AIDS (London, England). 2014 Apr 24;28(7):943-7.
[111] Wainberg MA, Mesplede T, Raffi F. What if HIV were unable to develop resistance against a new therapeutic agent? BMC medicine. 2013;11:249.
[112] Li G, Verheyen J, Rhee SY, Voet A, Vandamme AM, Theys K. Functional conservation of HIV-1 Gag: implications for rational drug design. Retrovirology. 2013;10:126.
[113] Radford AD, Chapman D, Dixon L, Chantrey J, Darby AC, Hall N. Application of next-generation sequencing technologies in virology. The Journal of general virology. 2012 Sep;93(Pt 9):1853-68.
[114] Li JZ, Paredes R, Ribaudo HJ, Svarovskaia ES, Metzner KJ, Kozal MJ, et al. Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis. Jama. 2011 Apr 6;305(13):1327-35.
[115] Luciani F, Bull RA, Lloyd AR. Next generation deep sequencing and vaccine design: today and tomorrow. Trends in biotechnology. 2012 Sep;30(9):443-52.
[116] Hayes R, Ayles H, Beyers N, Sabapathy K, Floyd S, Shanaube K, et al. HPTN 071 (PopART): rationale and design of a cluster-randomised trial of the population impact of an HIV combination prevention intervention including universal testing and treatment - a study protocol for a cluster randomised trial. Trials. 2014;15:57.

[117] Dennis AM. Phylogeny Studies in HIV Prevention Research. Imperial College, London: National Institutes of Health, National Institute of Allergy and Infectious Diseases, Division of AIDS, HIV Prevention Trials Network, Bill & Melinda Gates Foundation; 2012.
[118] Prosperi MC, Ciccozzi M, Fanti I, Saladini F, Pecorari M, Borghi V, et al. A novel methodology for large-scale phylogeny partition. Nature communications. 2011 May;2:321.
[119] Ragonnet-Cronin M, Hodcroft E, Hue S, Fearnhill E, Delpech V, Brown AJ, et al. Automated analysis of phylogenetic clusters. BMC bioinformatics. 2013;14:317.
[120] Fisher M, Pao D, Brown AE, Sudarshi D, Gill ON, Cane P, et al. Determinants of HIV-1 transmission in men who have sex with men: a combined clinical, epidemiological and phylogenetic approach. AIDS (London, England). 2010 Jul 17;24(11):1739-47.

SUPPLEMENTARY MATERIAL

CHAPTER 2: AUTOMATED SUBTYPING OF HIV-1 GENETIC SEQUENCES FOR CLINICAL AND SURVEILLANCE PURPOSES: PERFORMANCE EVALUATION OF THE NEW REGA VERSION 3 AND SEVEN OTHER TOOLS

Supplementary material is available at:
http://www.sciencedirect.com/science/article/pii/S1567134813001810

- Supplementary material table 2.1: Assignments not reproducible by the subtyping tools
- Supplementary material table 2.2: Sensitivity and Specificity of the subtyping tools using the clinical dataset
- Supplementary material table 2.3: Sensitivity and Specificity of the subtyping tools using the LANL dataset only
- Supplementary material table 2.4: Sensitivity and Specificity of the subtyping tools in RT using the LANL dataset only
- Supplementary material Table 2.5: Sensitivity and Specificity of the subtyping tools in PR using the LANL dataset only
- Supplementary material 2.6: Accession numbers of the LANL dataset
- Supplementary material Figure 2.1: ML tree of the genetic diversity of the LANL dataset.
- Supplementary material Figure 2.2: ML tree of the genetic diversity of the clinical dataset
- Supplementary material Figure 2.3: ML tree of the *pol* region which includes G, and CRF14_BG from the Portuguese dataset, LANL dataset and the reference sequences retrieved from Los Alamos database.

CHAPTER 3: MOLECULAR EPIDEMIOLOGY AND PHYLODYNAMICS OF THE HIV-1 EPIDEMIC IN COLOMBIA

Supplementary material Table 3.1: Countries and Regions included in the BLASTn-COL dataset and used to perform Bayesian Phylogenetic analyses. The 10 most similar sequences to the Colombian sequences were retrieved from BLASTn. The countries were grouped in regions to analyze the data. The location of sequences with unknown origin was estimated with Bayesian analyses and the results are in italics. *Additional Colombian sequences that were retrieved from BLASTn. Abbreviations: n: Number of sequences, USA: United States of America, OAM: Orinoquía and Amazonas

	Region	Country	n	(%)
Sequences retrieved by BLASTn	Spain	Spain	99	10.7
	North America	Canada	21	2.3
		USA	19	2.0
	Southern Europe	Croatia	1	0.1
		Greece	25	2.7
		Italy	3	0.3
		Portugal	3	0.3
		Serbia	8	0.9
		Slovenia	12	1.3
	Rest of Europe	Belgium	2	0.2
		Denmark	1	0.1
		France	5	0.5
		Germany	3	0.3
		Ireland	13	1.4
		Luxembourg	4	0.4
		Norway	1	0.1
		Poland	18	1.9
		Romania	1	0.1
		Russia	2	0.2
		Switzerland	5	0.5

Region		Country	n	(%)
		Ukraine	1	0.1
		United Kingdom	4	0.4
	Central and South America	Costa Rica	2	0.2
		Honduras	2	0.2
		Mexico	29	3.1
		Venezuela	3	0.3
	Others	Australia	2	0.2
		China	2	0.2
		Georgia	7	0.8
		Israel	4	0.4
		Japan	1	0.1
		South Africa	1	0.1
Final Colombian dataset	Colombia	Bogotá	269+18*	
		Cali	76	
		Caribbean	18+1*	
		Coffee	42	
		Medellin	63	
		OAM	6	
		Santander	25+3*	
		Unknown	102	
		Bogotá	*73*	
		Cali	*11*	
		Caribbean	*1*	
		Coffee	*6*	
		Medellin	*8*	
		Santander	*3*	
		Total	623	67.2
	Total		**927**	**100.0**

Supplementary material Table 3.2. GenBank accession numbers, sampling locations and year for the dataset retrieved from BLASTn.

GenBank	Country	Year
AB078708	China	2002
AB485638	USA	2009
AB485640	USA	2009
AB640503	Japan	2011
AF301289	USA	2000
AF395350	USA	2001
AF514275	USA	2002
AJ401737	Luxembourg	2000
AJ401788	Luxembourg	2000
AJ401794	Luxembourg	2000
AJ401800	Luxembourg	2000
AJ971093	Slovenia	2005
AY013837	USA	2000
AY016968	Venezuela	2000
AY016973	Venezuela	2000
AY030760	USA	2001
AY064298	Australia	2001
AY260915	CostaRica	2003
AY260931	CostaRica	2003
AY800676	USA	2004
AY801082	USA	2004
AY801206	USA	2004
DQ177224	Belgium	1989
DQ177233	Belgium	1989
DQ369203	Italy	2006
DQ537049	USA	2006
DQ823364	Ukraine	2006
EF368737	Canada	2007
EF368889	Canada	2007
EF369020	Canada	2007
EF369068	Canada	2007
EF549344	USA	2007
EU091157	Canada	2007
EU091158	Canada	2007
EU091164	Canada	2007
EU091165	Canada	2007
EU091167	Canada	2007
EU091171	Canada	2007
EU091173	Canada	2007
EU091174	Canada	2007
EU091175	Canada	2007
EU091176	Canada	2007
EU091178	Canada	2007

GenBank	Country	Year
GQ399577	Poland	2003
GQ399578	Spain	2004
GQ399586	Slovenia	2004
GQ399587	Slovenia	2004
GQ399588	Spain	2004
GQ399595	Spain	2003
GQ399609	Greece	2004
GQ399627	Denmark	2004
GQ399656	Slovenia	2004
GQ399658	Spain	2003
GQ399680	Greece	2004
GQ399683	Ireland	2004
GQ399704	Spain	2003
GQ399730	Spain	2003
GQ399745	Spain	2004
GQ399747	Spain	2004
GQ399748	Spain	2003
GQ399753	Spain	2003
GQ399767	Spain	2004
GQ399772	Ireland	2005
GQ399773	Greece	2004
GQ399782	Israel	2003
GQ399848	Spain	2004
GQ399860	Spain	2004
GQ399871	Spain	2003
GQ399880	Spain	2004
GQ399887	Spain	2004
GQ399892	Poland	2003
GQ399893	Spain	2004
GQ399899	Ireland	2003
GQ399900	Greece	2004
GQ399909	Poland	2003
GQ399926	Poland	2003
GQ399936	Spain	2003
GQ399939	Spain	2003
GQ399963	Spain	2003
GQ399968	Greece	2004
GQ400005	Spain	2003
GQ400024	Spain	2003
GQ400029	Slovenia	2004
GQ400041	Spain	2003
GQ400058	Portugal	2003
GQ400088	Spain	2004

GenBank	Country	Year
EU091179	Canada	2007
EU091180	Canada	2007
EU091181	Canada	2007
EU345857	Russia	2007
EU345913	Russia	2007
EU346342	USA	2007
EU375798	Canada	2008
EU375799	Canada	2008
EU375801	Canada	2008
FJ030791	United Kingdom	2000
FJ030794	United Kingdom	1998
FJ376432	Georgia	2006
FJ426617	Georgia	2006
FJ426619	Georgia	2006
FJ426620	Georgia	2007
FJ426621	Georgia	2007
FJ426624	Georgia	2007
FJ426625	Georgia	2007
FJ469749	USA	2002
FJ659544	Venezuela	2005
FJ983533	USA	2005
FN424301	Croatia	2009
GQ207662	USA	2003
GQ209719	USA	2002
GQ210679	USA	2005
GQ251249	Italy	2009
GQ251256	Italy	2009
GQ398832	Spain	2004
GQ398847	Spain	2004
GQ398871	Spain	2004
GQ398875	Spain	2004
GQ398887	Greece	2004
GQ398913	Spain	2003
GQ398929	Spain	2003
GQ398930	Spain	2003
GQ398951	Spain	2004
GQ398963	Spain	2004
GQ398974	Spain	2004
GQ398980	Ireland	2004
GQ398985	Spain	2003
GQ398986	Spain	2003
GQ398987	Greece	2004
GQ398996	Spain	2004
GQ398997	Poland	2003
GQ399000	Greece	2005
GQ399013	Spain	2003
GQ399017	Spain	2004

GenBank	Country	Year
GQ400098	Spain	2004
GQ400110	Spain	2004
GQ400116	Spain	2003
GQ400122	Spain	2004
GQ400123	Poland	2003
GQ400140	Spain	2004
GQ400152	Poland	2003
GQ400153	Spain	2003
GQ400173	Greece	2004
GQ400176	Spain	2003
GQ400178	Spain	2003
GQ400182	Greece	2004
GQ400183	Spain	2003
GQ400197	Spain	2004
GQ400198	Poland	2003
GQ400199	Greece	2005
GQ400228	Poland	2003
GQ400246	Poland	2003
GQ400247	Spain	2003
GQ400251	Spain	2003
GQ400256	Ireland	2004
GQ400271	Spain	2004
GQ400274	Spain	2004
GQ400276	Ireland	2004
GQ400289	Spain	2004
GQ400291	Spain	2003
GQ400293	Spain	2004
GQ400301	Spain	2003
GQ400302	Spain	2003
GQ400306	Greece	2004
GQ400311	Greece	2004
GQ400314	Greece	2004
GQ400317	Ireland	2004
GQ400358	Spain	2003
GQ400416	Israel	2003
GQ400417	Spain	2003
GQ400435	Israel	2003
GQ400440	Greece	2004
GQ400443	Poland	2003
GQ400447	Spain	2003
GQ400474	Slovenia	2004
GQ400481	Spain	2004
GQ400509	Spain	2004
GQ400512	Poland	2003
GQ400516	Spain	2003
GQ400518	Spain	2003
GQ400524	Slovenia	2005

GenBank	Country	Year
GQ399021	Portugal	2003
GQ399024	Ireland	2003
GQ399025	Spain	2004
GQ399034	Spain	2004
GQ399045	Spain	2004
GQ399049	Poland	2003
GQ399060	Spain	2004
GQ399094	Poland	2003
GQ399096	Spain	2004
GQ399113	Slovenia	2004
GQ399119	Spain	2004
GQ399132	Ireland	2003
GQ399136	Greece	2004
GQ399139	Spain	2003
GQ399162	Greece	2004
GQ399173	Spain	2003
GQ399177	Spain	2003
GQ399193	Greece	2004
GQ399195	Poland	2003
GQ399198	Poland	2003
GQ399201	Spain	2003
GQ399212	Greece	2005
GQ399215	Greece	2004
GQ399216	Greece	2004
GQ399218	Spain	2004
GQ399220	Ireland	2003
GQ399244	Slovenia	2004
GQ399245	Slovenia	2004
GQ399247	Poland	2003
GQ399278	Spain	2004
GQ399288	Greece	2005
GQ399293	Serbia	2004
GQ399299	Israel	2003
GQ399300	Spain	2003
GQ399313	Spain	2004
GQ399339	Ireland	2003
GQ399342	Spain	2004
GQ399350	Poland	2003
GQ399360	Greece	2005
GQ399376	Spain	2004
GQ399383	Spain	2004
GQ399395	Slovenia	2004
GQ399407	Greece	2004
GQ399412	Spain	2004
GQ399420	Spain	2004
GQ399422	Ireland	2003
GQ399424	Spain	2004

GenBank	Country	Year
GQ400532	Slovenia	2004
GQ400568	Serbia	2004
GQ400572	Germany	2004
GQ400634	Serbia	2004
GQ400712	Germany	2004
GQ400727	Serbia	2004
GQ400745	Germany	2004
GQ400860	Serbia	2004
GQ400943	Serbia	2004
GQ400971	Serbia	2004
GQ401005	Serbia	2004
GU264223	Spain	2006
GU326139	Spain	2008
GU344279	Switzerland	1997
GU344280	Switzerland	1998
GU344281	Switzerland	1998
GU344283	Switzerland	1997
GU344284	Switzerland	1998
GU345179	China	2006
GU382759	Mexico	2003
GU382761	Mexico	2003
GU382766	Mexico	2003
GU382767	Mexico	2003
GU382769	Mexico	2003
GU382772	Mexico	2003
GU382776	Mexico	2003
GU382778	Mexico	2003
GU382781	Mexico	2003
GU382783	Mexico	2003
GU382785	Mexico	2003
GU382786	Mexico	2003
GU382787	Mexico	2003
GU382799	Mexico	2003
GU382811	Mexico	2003
GU382812	Mexico	2003
GU382813	Mexico	2003
GU382814	Mexico	2003
GU382819	Mexico	2003
GU382823	Mexico	2003
GU382824	Mexico	2003
GU382827	Mexico	2003
GU382830	Mexico	2002
GU382831	Mexico	2003
GU382832	Mexico	2003
GU382833	Mexico	2002
GU382835	Mexico	2003
GU382836	Mexico	2003

GenBank	Country	Year
GQ399447	Spain	2004
GQ399450	Spain	2004
GQ399453	Spain	2004
GQ399454	Norway	2004
GQ399466	Spain	2004
GQ399469	Spain	2003
GQ399484	Spain	2004
GQ399495	Spain	2004
GQ399516	Ireland	2003
GQ399531	Portugal	2003
GQ399539	Spain	2004
GQ399541	Spain	2004
GQ399548	Spain	2004
GQ399564	Greece	2004
GQ399568	Spain	2004

GenBank	Country	Year
GU382847	Mexico	2002
HM460461	Spain	2001
JF895261	Australia	2011
JN100967	United Kingdom	2007
JN101340	United Kingdom	2006
JN214965	USA	2008
JN215231	Honduras	2005
JN215307	Honduras	2005
JN638211	South Africa	2007
JN982147	Romania	2010
JQ292205	France	2010
JQ292207	France	2010
JQ292208	France	2010
JQ292209	France	2010
JQ292677	France	2010

Supplementary Material Table 3.3: Proportion of sequences per country or region according to the number of the closest sequences used in the BLASTn search.

BLASTn Search								
10 sequences			**30 sequences**			**50 sequences**		
Country/Region	**n**	**%**	**Country/Region**	**n**	**%**	**Country/ Region**	**n**	**%**
Spain	99	32.6	Spain	148	26.8	Spain	180	23.3
North America	40	13.2	North America	92	16.6	North America	138	17.9
Southern Europe	52	17.1	Southern Europe	85	15.4	Southern Europe	113	14.6
Rest of Europe	60	19.7	Rest of Europe	84	15.2	Rest of Europe	141	18.2
Central and South America	36	11.8	Central and South America	82	14.8	Central and South America	118	15.3
Others	17	5.6	Others	62	11.2	Others	83	10.7
Total	**304**	**100**	**Total**	**553**	**100**	**Total**	**773**	**100**

Supplementary material Table 3.4: Sampling locations of the dataset retrieved from Los Alamos database (n=7931) named the LANL-Borders dataset. We took the available *pol* subtype B sequences up to 25 for each year and each country, except for the neighboring countries such as Brazil, Ecuador, Panama, Peru, and Venezuela.

Country	N	Country	N	Country	N
Argentina	230	Gabon	2	Norway	46
Austria	55	United Kingdom	275	Panama	134
Australia	107	Georgia	3	Peru	248
Belgium	52	Ghana	1	Philippines	10
Burkina Faso	1	Greenland	72	Poland	52
Bolivia	1	Greece	37	Portugal	70
Brazil	2193	Guyana	6	Paraguay	3
Bahamas	11	Hong Kong	3	Romania	69
Belarus	7	Honduras	166	Russian Federation	86
Belize	9	Haiti	16	Sudan	1
Canada	186	Ireland	6	Sweden	131
Central African Republic	1	India	1	Singapore	26
Switzerland	284	Iran (Islamic Republic of)	13	Slovenia	65
Chile	25	Italy	246	Slovakia	8
China	187	Jamaica	30	Senegal	24
Costa Rica	1	Japan	108	Suriname	5
Cuba	109	Korea, Republic of (South)	305	El Salvador	17
Cyprus	69	Luxembourg	20	Thailand	36
Czech Republic	63	Latvia	18	Trinidad And Tobago	30
Germany	182	Mali	1	Taiwan	272
Denmark	107	Myanmar	1	Ukraine	5
Dominican Republic	75	Mongolia	10	United States	426
Ecuador	46	Mexico	48	Uruguay	6
Spain	248	Malaysia	11	Venezuela	347
Finland	31	Nigeria	2	Yemen	2
France	77	Netherlands	30	South Africa	24

Supplementary material Table 3.5: Results of the Bayesian analyses on the BLASTn-Col dataset. The table shows the migration links within Colombia and with other regions or countries. Abbreviations: [a] The migration rate was identical in the three additional weighted datasets, BCI: Bayesian credible intervals

Link	Bayes Factor	Markov Jumps	
		Mean	95% BCI
a. Bogotá, Colombia and Cali, Colombia	>1000 [a]	44,5	33 - 54
b. Bogotá, Colombia and Santander region, Colombia	>1000	19,1	12 - 24
c. Bogotá, Colombia and Medellín, Colombia	>1000 [a]	37,6	25 - 49
d. Bogotá, Colombia and Coffee region, Colombia	>1000	17,6	9 - 27
e. Spain and North America	>1000 [a]	31,9	26 - 38
f. Spain and West-East-Northern Europe	>1000 [a]	40,6	35 - 46
g. Spain and Southern Europe	>1000 [a]	45,4	39 - 51
h. Spain and Central-South America	>1000 [a]	27,2	21 - 32
i. Spain and Bogotá, Colombia	>1000 [a]	89,4	74 - 105
j. West-East-Northern Europe and Others	>1000 [a]	1,4	1 - 3
k. Spain and Others	605,1 [a]	6,9	4 - 10
l. Spain and Medellín, Colombia	447,6 [a]	20,1	8 - 29
m. Spain and Cali, Colombia	395,5 [a]	21,3	12 - 28
n. Spain and Coffee region, Colombia	220,3 [a]	13,9	6 - 20
o. Bogotá, Colombia and Caribbean region, Colombia	175,8	8,9	3 - 14
p. Spain and Caribbean region, Colombia	80,3 [a]	10,0	4 - 15
q. Bogotá, Colombia and Amazonas region, Colombia	45,4	2,6	0 - 5
r. Bogotá, Colombia and Spain	40,9	5,1	0 - 10
s. Others and West-East-Northern Europe	15,1	1,1	0 - 2
t. Spain and Santander region, Colombia	10,5	4,7	0 - 9
u. North America and West-East-Northern Europe	5,9	1,0	0 - 3
v. Medellín, Colombia and Cali, Colombia	4,5	1,8	0 - 4
w. Cali, Colombia and Santander region, Colombia	3,6	1,6	0 - 4
x. Cali, Colombia and Bogotá, Colombia	3,5	0,7	0 - 3
y. Others and Spain	3,2	0,2	0 - 1
z. Medellín, Colombia and Bogotá, Colombia	3,1	1,5	0 - 5
aa. Cali, Colombia and Amazonas region, Colombia	3,1	0,3	0 - 1

CHAPTER 4: HIV-1 TRANSMITTED DRUG RESISTANCE IN LATIN AMERICA AND CARIBBEAN: WHAT DO WE KNOW?

Supplementary material is available at

http://www.aidsreviews.com/files/2012_14_4_256-267_SUPL.pdf

- Supplementary material table 4.1: Characteristics of published studies on transmitted drug resistance in Latin America and Caribbean
- Supplementary material table 4.2: Levels of transmitted drug resistance in Latin America and Caribbean (articles)
- Supplementary material table 4.3: Levels of transmitted drug resistance in Latin America and Caribbean (abstracts)

CHAPTER 5: TRENDS AND PREDICTORS OF TDR AND CLUSTERS WITH TDR IN A LOCAL BELGIAN HIV-1 EPIDEMIC

Supplementary material is available at

http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0101738#s5

- **Supplementary material Table 5.1:** Characteristics of patients from the Leuven ND cohort and from patients involved in transmission clusters.
- **Supplementary Table 2:** Impact of transmitted drug resistance (TDR) on clinical care.

AIDS Rev. 2012;14:256-67

HIV-1 Transmitted Drug Resistance in Latin America and the Caribbean: What Do We Know?

Andrea-Clemencia Pineda-Peña[1,2], Diana-Carolina Bello[1], Otto Sussmann[1], Anne-Mieke Vandamme[2,3], Jurgen Vercauteren[2], Kristel van Laethem[2,4] and Arley Gómez-López[1]

[1]Clinical and Molecular Infectious Disease Group, Faculty of Sciences and Mathematics, Universidad del Rosario, Bogotá, Colombia; [2]Rega Institute for Medical Research, KU Leuven, Leuven, Belgium; [3]Centro de Malária e outras Doenças Tropicais, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisboa, Portugal; [4]AIDS Reference Laboratory, Universitaire Ziekenhuizen Leuven, Leuven, Belgium

Abstract

Latin America and the Caribbean countries have increased the scaling-up of antiretroviral treatment in the last years. The increase of transmitted drug resistance has been feared due to the worrisome indicators associated with the emergence of drug resistance and monitored by the World Health Organization (WHO). Consequently, our aim was to review all relevant studies on transmitted drug resistance in Latin America and the Caribbean countries, to analyze its levels, to identify the frequency of transmitted drug resistance mutations, and to put these results in the context of the local Latin American and Caribbean countries settings.

A systematic search of Spanish, Portuguese, and English literature was performed in databases and international conferences for the period June 1999 to May 2011. In addition, sequences were downloaded from the Los Alamos and Stanford databases and the transmitted drug resistance was reanalyzed according to the WHO Surveillance Drug Resistance Mutation list 2009.

In total, 50 articles, 27 abstracts, and 1,922 patients were included. The resistance varied geographically, but most of the countries have reached the WHO threshold of 5% of resistance. According to the sequences available in public databases, the overall prevalence in Latin America and the Caribbean countries for the period 1996-2009 was 7.7% and by region it was 4.3% for the Caribbean, 3.9% for Mexico, 9.4% for Brazil, 10.5% for the Andean region and 4.9% for the Southern Cone. For the last four investigated years (2006-2009), the information was restricted to Brazilian and Venezuelan studies and revealed an overall transmitted drug resistance of 10%. Throughout the study period, limited information was available for the Caribbean and Central American countries. These findings support the need for developing comprehensive surveys of transmitted drug resistance in these regions. **(AIDS Rev. 2012;14:256-67)**

Corresponding author: Arley Gómez-López, arley.gomez@urosario.edu.co

Key words

HIV-1. Drug resistance. Latin America. Caribbean. Genotype. Naive.

Introduction

The horizontal transmission of an HIV-1 strain that already displayed the reverse transcriptase (RT) mutation T215F/Y was first described in a homosexual man in 1993[1]. As it was feared that transmitted drug resistance (TDR) could negatively impact the success of first-line antiretroviral therapy (ART), TDR was extensively studied in resource-rich countries. Retrospective studies provided evidence that infection with drug-resistant HIV-1 could also occur in other risk groups, and indicated increasing TDR levels and more extensive mutational patterns. As these studies were often small and based upon convenience sampling, prospective and representative population-based studies were set up in Europe, the USA and Canada to monitor the

Correspondence to:
Arley Gómez-López
Clinical and Molecular Infectious Disease Group
Faculty of Sciences and Mathematics
Calle 63D No. 24-31
Bogotá, Colombia
E-mail: arley.gomez@urosario.edu.co

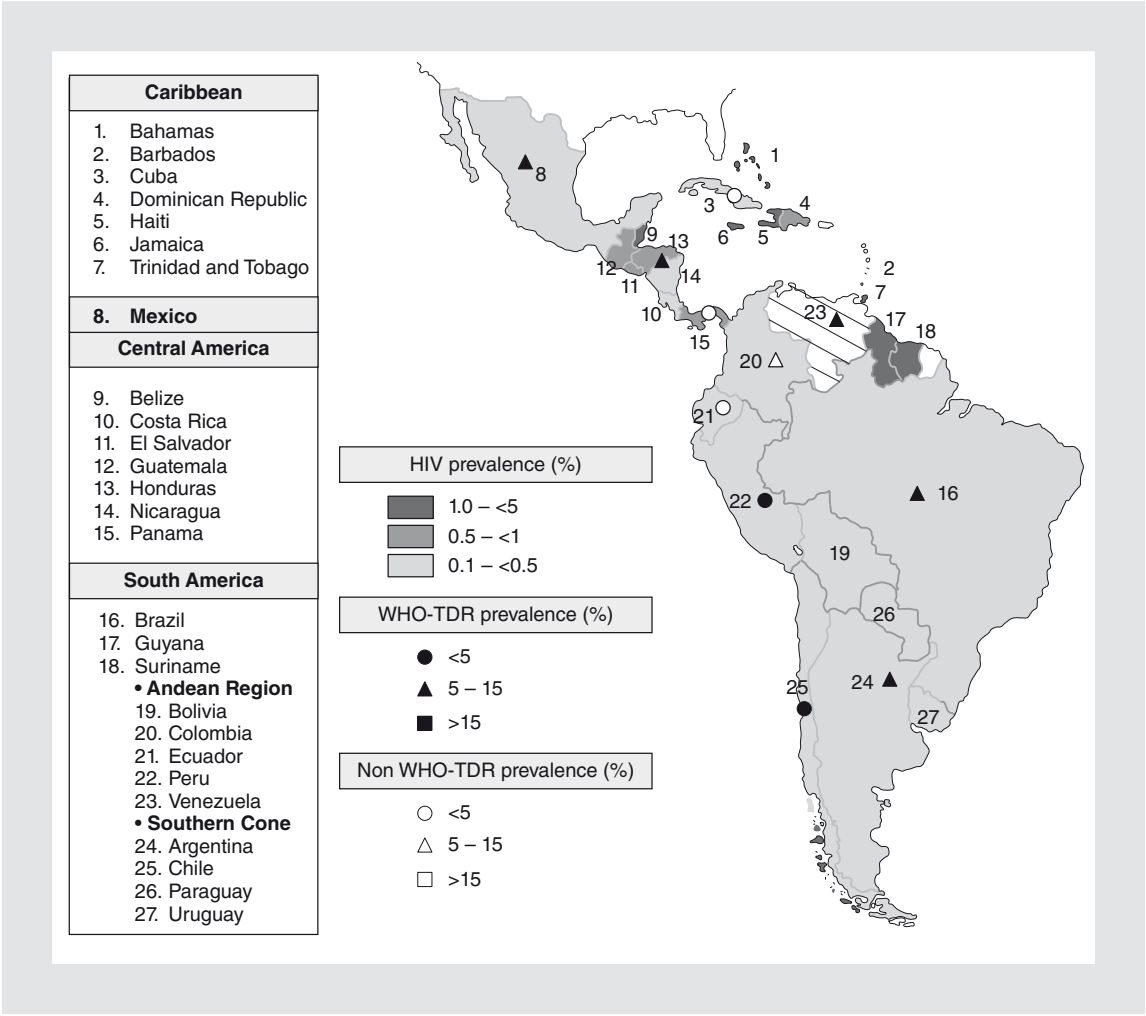


Figure 1. *HIV and transmitted drug resistance (TDR) prevalence in Latin America and the Caribbean. This region is divided geographically in Caribbean (seven countries), Mexico, Central America (seven countries) and South America (12 countries). The latter region is subdivided based upon geographical, social and economic characteristics: the Andean region and Southern Cone. The HIV prevalence according to the UNAIDS report is shown in a grayscale[9]. Although the overall HIV prevalence in Venezuela (hatched) was 0.75% in 2004, the country was not visualized in a grayscale as the prevalence in 2007 varied between 1.62 % in metropolitan areas and 0.08% in more rural areas[90]. The TDR prevalence values from studies that followed the eligibility criteria of WHO surveys are represented in black figures, otherwise they are represented in white. Adapted from UNAIDS report, 2010.*

extent and dynamics of TDR[2-5]. The high rates of TDR (overall prevalence 8-9%) resulted in the recommendation of standard genotypic drug resistance testing to guide ART selection in therapy naive patients[6,7]. Recently, a large European multi-cohort study confirmed the importance of baseline drug resistance testing as patients with TDR and resistance to at least one prescribed drug were three times as likely to virologically fail within one year[8].

As access to ART has been implemented and scaled-up in Latin-America and the Caribbean (LAC) since more than a decade, concerns about the emergence and spread of TDR strains in LAC exist. In general, the prevalence of TDR depends mainly on the time since ART implementation, the scale of ART coverage, and the overall virologic success of the prescribed regimens. As the majority of HIV-1 patients in LAC started immediately on HAART, the rise in TDR was anticipated to be less extensive compared to resource-rich countries where sequential mono- and bitherapies were initially prescribed. Nevertheless, other factors unique to the particular setting of LAC could negatively influence TDR rates. For instance, the widespread use of standard first-line regimens based on

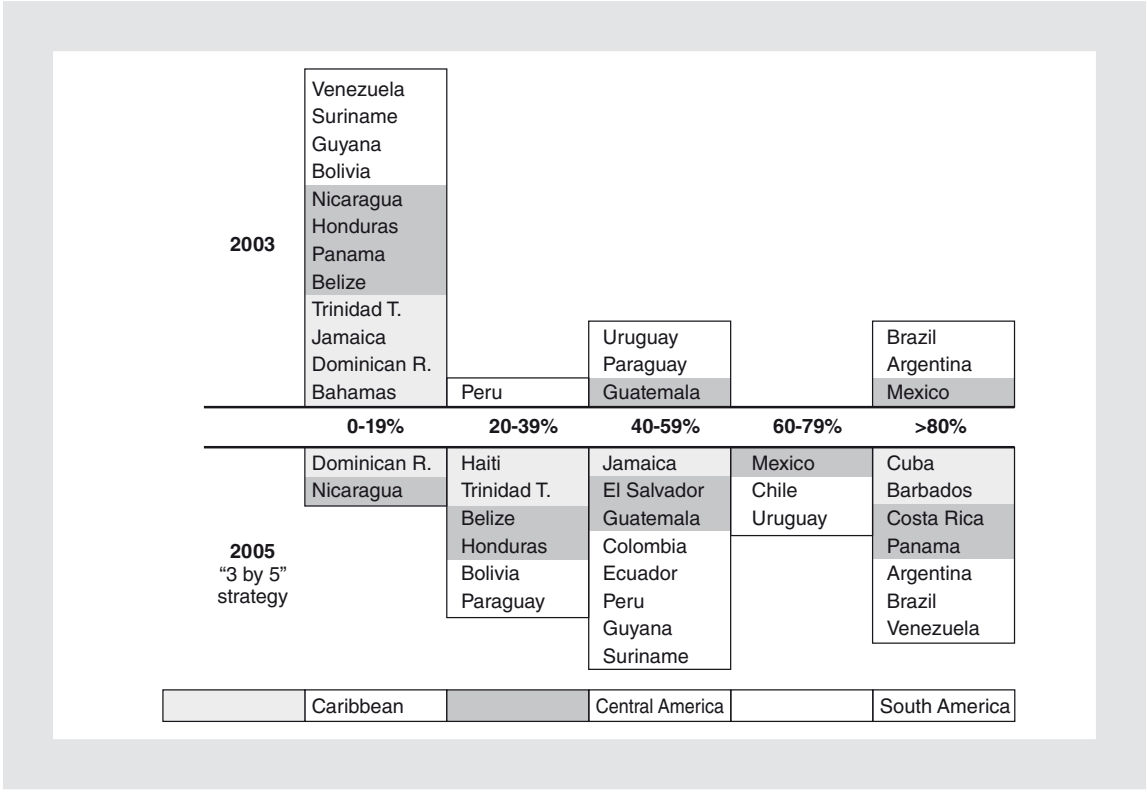


Figure 2. *Scaling-up of antiretroviral therapy (ART). The scaling-up of ART in 2003 and after the "3 by 5" strategy launched by WHO and UNAIDS. The "3 by 5" stands for treatment for three million HIV/AIDS patients in low-middle income countries by 2005 (see details in the UNAIDS report "3 by 5" and beyond)[51]. The regions are shown in a grayscale. Progress in access to treatment is shown as a shift to the right for a country, comparing 2003 (upper panel) with 2005 (lower panel). Countries are scaled as percentage (%) of those diagnosed and eligible for treatment that actually received treatment. The scale is shown in the middle panel. Data were not available for Barbados, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador and Haiti in 2003. Adapted from UNAIDS[51]. Trinidad T: Trinidad and Tobago; Dominican R: Dominican Republic.*

low genetic barrier non-nucleoside RT inhibitors (NNRTI) in the absence of drug resistance testing, together with limited access to second-line therapies, could promote emergence of drug resistance and its subsequent spread to new HIV-1 infections. The LAC is a heterogeneous region that differs extensively in geographical, social and economic characteristics that influence the HIV epidemiology and clinical care facilities. The HIV-1 burden varies considerably between and within the 27 countries belonging to LAC (Fig. 1). In 2009, 1.7 million people were living with HIV-1 in LAC, of which one-third lived in Brazil, and HIV-1 prevalence values ranged between 0.1% in Cuba and 3.1% in Bahamas[9]. Brazil and Argentina took a leading role in the area by providing ART since 1991-1992 and by setting up national programs to provide free access to HAART, clinical care and laboratory monitoring (Fig. 2)[10,11]. Currently, Cuba, Chile and Nicaragua achieve global access, and Argentina, Brazil, Dominican Republic, Mexico and Uruguay are near to universal-coverage levels. However, in other LAC countries ART scale-up has been rather slow and only reached ART coverage between 20% in Bolivia and 66% in Paraguay with still limited access to laboratory testing and second-line regimens[12].

The scope of this review was to collect all relevant studies on TDR in LAC, to analyze TDR levels, to identify the frequency of TDR mutations, and to put these results in the context of the local LAC settings.

Transmitted drug resistance surveillance and early warning indicators

Recently, the fear of TDR in developing countries reappeared due to the worrying reports of early warning indicators in Africa, the Caribbean and Central America[13-16]. Early warning indicators were developed by the WHO as indicators for the prevention of drug

resistance. They include ART prescribing practices, lost-to-follow-up, retention of first-line therapies, viral suppression and percentage of discontinued drug supplies[13,16]. These indicators alert earlier than prospective surveys about the potential of increasing TDR rates.

Few studies have evaluated early warning indicators in LAC. Most of the studies were restricted to Brazil or to the seven AIDS programs from the CCASAnet Cohort[17,18]. Lost-to-follow-up was 6% overall, but large differences were observed between the different AIDS healthcare institutions (between 0.6% in Honduras to 17% in Argentina)[17]. In El Salvador, Guatemala, Honduras and Nicaragua, only seven of 13 sites accomplished the goal of ≤ 20% lost-to-follow-up, and 38 of 55 sites in the Caribbean[13,14]. Low numbers of retention in first-line therapy were also observed in LAC, ranging from 16% in Honduras to 36% in Peru. Adverse events were the main cause in 14% of cases, predominantly linked to zidovudine (hematological toxicity in 7%) and nevirapine (skin rash in 3%)[17,18].

To standardize TDR surveillance studies, the WHO has developed minimum recommendations for developing countries to enable comparison. Briefly, the criteria are: a minimum of 47 samples from small geographic areas within countries where ART has been available for more than three years to at least 20% of the eligible HIV-1 population; HIV-1 infection should have been confirmed by laboratory criteria; participants should not have received any antiviral drug; and if female they should have been younger than 25 years at their first pregnancy. The inclusion of individuals with less than three years of HIV-1 infection or/and laboratory evidence of seroconversion or recent infection are required. If the information is accessible, demographic data like transmission risk group, clinical data like non-AIDS stage, or laboratory data like CD4 > 500 cells/ml, and no previous positive HIV-1 test are desirable[19,20]. In addition, the analysis of resistance should be done with the WHO Surveillance Drug Resistance Mutation (SDRM) list, which is updated periodically and excludes highly polymorphic positions[21]. Nevertheless, some SDRM do occur naturally in therapy naive individuals, which might inflate the estimates of TDR. The working group decided to retain these mutations within the list as stricter inclusion criteria would result in the omission of important drug resistance mutations and these particular mutations might have been the result of unreported therapy exposure or truly TDR within the therapy naive dataset[21].

Epidemiology of transmitted drug resistance in Latin America and the Caribbean

Due to the low achievement of early warning indicators in several LAC countries, available publications on TDR were reviewed using a systematic search strategy (Fig. 3; see Supplementary methods). Fifty articles and 27 abstracts were retained (Tables 1, 2, 3). Approximately half of the studies originated from Brazil (53.2%, 41/77). Ten studies were from Mexico (13%, 10/77), three from the Caribbean (3.9%, 3/77), four from Central America (5.2%, 4/77), 12 from the Southern Cone (15.6%, 12/77) and seven (9.1%, 7/77) from the Andean region.

Heterosexual contacts and men who have sex with men (MSM) were the most frequent transmission risk factors. Six studies were performed in specific patient populations such as blood donors[22-24], female sex workers[25], intravenous drug users[26] and a combination of the latter two[27].

Eleven studies used the Serological Testing Algorithm for Recent HIV Seroconversion (STARHS)[28] strategy to differentiate between recent or chronic infection[24,29-36]. In five other studies, the distinction between recent and chronic infection was based upon previous negative serology test results 9-18 months before diagnosis date[37-41].

Although plasma was widely used as source material, peripheral blood mononuclear cells (PBMC) and dried blood samples (DBS) were also used in nine and two studies, respectively. In most of the studies, in-house methodologies were used to generate sequences, followed by Trugene® (seven) and ViroSeq® (six) (Table 1). The genotypic resistance interpretation algorithm HIVdb was widely used for TDR analysis. Twenty-one studies used the International AIDS Society list and 19 studies the SDRM list. The version of the respective algorithms and lists were often not mentioned in the original publication.

The TDR levels as reported within the original publications are shown in tables 2 and 3. In summary, the prevalence of TDR ranged from 0 to 7.4% in the Caribbean for 1996-2003, from 2.8 to 18% in Mexico for 2001-2010, from 0 to 11.6% in Central America for 2002-2007, from 0 to 41% in Brazil for 1994-2010, from 3 to 11% in the Andean Region for 1998-2007 and from 2.5 to 18.8% in the Southern Cone for 1997-2009.

Due to the heterogeneity of mutation lists used within the original studies, TDR was reanalyzed using the SDRM list 2009 and resulted in TDR prevalence values between 0 and 8.3% in the Caribbean, 2.8 and 6.8% in Mexico, 0 and 7.5% in Central America, 1.3 and

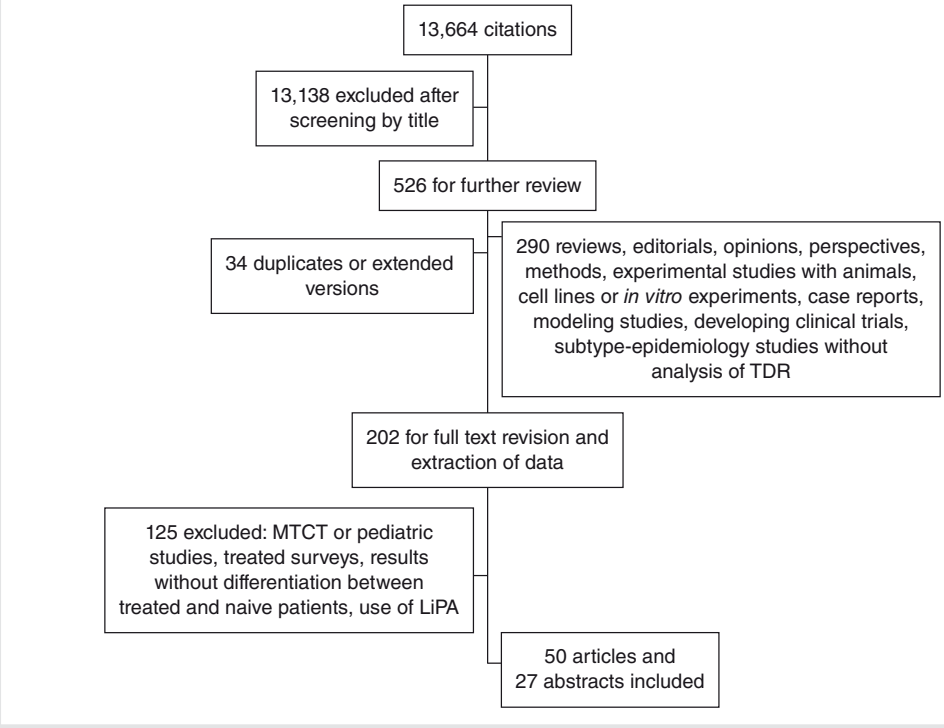


Figure 3. *Selection of studies on transmitted drug resistance (TDR). MTCT: Mother-to-child transmission; LiPA: Line Probe Assay.*

20.3% in Brazil, 0 and 13% in the Andean region and from 1.6 to 12% in the Southern Cone. For comparison, values reported in 14 studies that followed the WHO surveillance criteria for TDR monitoring are also displayed in tables 1 and 2[22,24,29,31-33,35,36,38,39,41-46].

For 10 studies, the presence of TDR in each sample could be directly linked to the respective transmission risk factor[25,31-33,37,41,44,47-49] and combining this data did not reveal any difference in TDR rate between heterosexuals (37/522, 7.1%) and MSM (39/757, 5.2%). One study that solely sampled blood donors from São Paulo (Brazil) showed a TDR of 5.6% in 1998-2002 (19/341)[22], whereas another study sampling in a broader geographical area revealed only a TDR of 1.3% in 2000-2004 (1/74)[24]. The majority of patients included in studies where differentiation between recent and chronic infection was made were chronically infected and therefore differences in TDR between both groups could not be investigated. However, two studies did report a difference[22,44]. For three studies that included at least 50 recent infections, TDR values were 4% (3/74) in Chile[38], 7.7% (4/52) in Argentina[41], and 8.1% (17/210) in Brazil[39].

To obtain a better understanding of the reported TDR rates, public databases were searched and 1,922 relevant HIV-1 sequences were retained for analysis (Fig. 4 A). Based upon automated subtyping[50], HIV-1 subtype B was the most frequent subtype (65.7%, 1,262/1,922), followed by subtype C (14.4%, 276/1,922), BF recombinants (8.7%, 168/1,922) and subtype F1 (4%, 77/1,922); the latter two mainly originated from Brazil and Argentina (BF: 99.4%, 167/168; F1: 98.7%, 76/77). The remaining genetic variants were circulating recombinant forms (CRF) such as CRF31_BC and CRF29_BF in Brazil, CRF_18cpx and CRF19_cpx in Cuba, and BD, BC and CF1 recombinants in Brazil and the Southern Cone.

In LAC, overall TDR prevalence was 7.7% (148/1,922; 95% CI: 6.54-8.98) for the period from 1996 to 2009. The TDR by region was 4.3% in Caribbean for 1996-2005 (8/186; 95% CI: 1.87-8.29), 3.9% in Mexico for 2002-2005 (2/51; 95% CI: 1.87-8.29), 9.4% in Brazil for 1998-2009 (99/1,056; 95% CI: 7.68-11.3), 10.5% in the Andean Region for 1999-2007 (15/142; 95% CI: 6.03-16.82), and 4.9% in the Southern Cone for 1997-2005 (24/487; 95% CI: 3.2-7.24). The NRTI resistance was

A

Country	Year							Total
	96-97	98-99	00-01	02-03	04-05	06-07	08-09	
BR	–	10/66	23/244	6/68	14/213	33/337	13/128	99/1,056
AR	1/28	–	1/47	13/274	4/59	–	–	19/408
VE	–	2/31	–	0/12	–	8/62	–	10/105
CH	–	–	5/79	–	–	–	–	5/79
TT	–	–	1/61	0/2	–	–	–	1/63
MX	–	–	–	2/40	0/11	–	–	2/51
CO	–	–	5/37	–	–	–	–	5/37
CU	–	3/25	–	–	–	–	–	3/25
BB	2/24	–	–	–	–	–	–	2/24
DO	–	–	–	1/12	0/5	–	–	1/17
HT	–	–	–	–	0/16	–	–	0/16
Other	–	–	1/40	–	0/1	–	–	1/41
Total	3/52	15/122	36/508	22/408	18/305	41/399	13/128	148/1,922

B

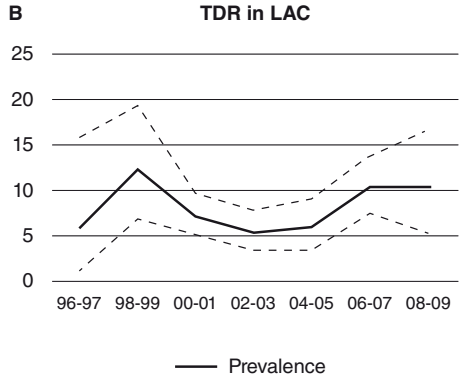


Figure 4. *Prevalence of transmitted drug resistance (TDR) and mutations in Latin America and the Caribbean (LAC). See Supplementary methods.* ***A:*** *TDR prevalence in LAC. The number of sequences with evidence of TDR (numerator) and the total number of sequences included (denominator) divided according to country and sampling period.* ***B:*** *The prevalence of TDR (%) according to sampling period (black line) and the 95% confidence intervals (gray dashes). BR: Brazil; AR: Argentina; VE: Venezuela; CH: Chile; TT: Trinidad and Tobago; MX: Mexico; CO: Colombia; CU: Cuba; BB: Barbados; DO: Dominican Republic; HT: Haiti; Other: include other Caribbean islands.*

4.4% (84/1,922; 95% CI: 3.54-5.38), NNRTI resistance was 2.3% (44/1,922; 95% CI: 1.71-3.06), and protease inhibitor (PI) resistance was 2.08% (40/1,922; 95% CI: 1.53-2.82). Of the TDR cases, 5.4% (8/148) and 6.8% (10/148) had dual class resistance towards NRTI + NNRTI and NRTI + PI, respectively. Triple-class resistance was only detected in 0.7% (1/148) of the TDR cases. The overall TDR time trend was not significant, although two TDR peaks above 9% were observed for the periods 1998-1999 and 2006-2007 (Fig. 4 B). The separate time trend analysis for the countries that contributed the majority of samples (Argentina and Brazil) were also not significant, neither were the trends for the individual drug classes (NRTI, NNRTI and PI).

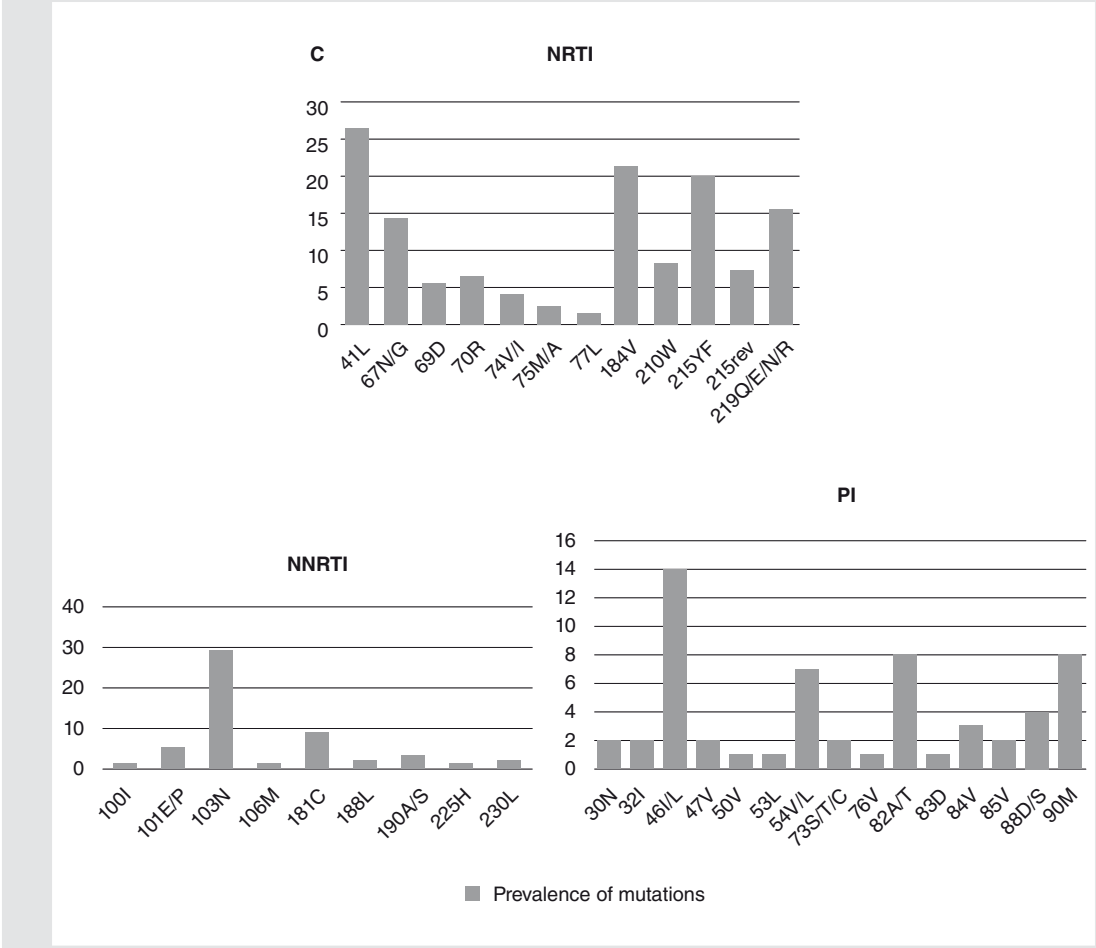


Figure 4. *Prevalence of transmitted drug resistance (TDR) and mutations in Latin America and the Caribbean (LAC). See Supplementary methods.* ***C:*** *The number of sequences with SDRM mutations against nucleoside reverse transcriptase inhibitors (NRTI), non-nucleoside reverse transcriptase inhibitors (NNRTI) and protease inhibitors (PI).*

The 103N mutation was the most frequently observed mutation (29/148, 19.6%), followed by 215Y/F/rev (27/148, 18.2%) and 41L (26/148, 17.6%) (Fig. 4 C). Time trends analysis for the individual mutations 41L and 103N did not reveal any significant changes. The frequency of the most prevalent TDR mutations, such as 41L, 184V, 103N, 46I/L, 82A/T and 90M, was compared with frequencies in the following studies in the USA, Canada and Europe[2,3,5], but no significant differences were observed.

For 447 sequences, information was available on whether it concerned a recent or a chronic infection. However, the TDR rate did not differ significantly between recent (4/52, 7.7%) and chronic (41/395, 10.4%) infections. The prevalence of TDR did differ between sequences obtained from PBMC or plasma (1,712 sequences included). The TDR was significantly higher in PBMC than plasma (33/285, 11.6%, vs. 98/1,427, 6.8%). However, the exclusion of PBMC samples did not significantly impact the overall TDR (7%; 95% CI: 5.83-8.37). Other factors were not evaluated because of lack of information.

Caribbean and Central America

According to the UNAIDS-WHO report, only a minimum ART coverage was implemented in the Caribbean region in 2001[51]. However, coverage increased with extra funding to, for example, 20% in Haiti and

Dominican Republic in 2005 (Fig. 2)[51]. As a result, TDR ranged from 0% in several islands in 2000 to 3.6% in Cuba in 2003[52-55]. In Cuba there was already limited access to mono- and bitherapy before 2000, which could explain the level of TDR observed in 2003 and its frequency of the NRTI drug class[53]. The overall TDR was estimated to be 4.3% and this was based upon reanalysis using the WHO TDR surveillance list of the collected sequences that were available in public databases and that were sampled in the time period 2000-2005[21].

From Central America, only fragmented data was available for two countries of this region. Indeed, only four publications and no sequences were retained after our search in public databases. A Panamanian study showed no TDR mutations in 2005, but NRTI and NNRTI TDR increased in a later report[56,57]. However, in the later study the exact sampling years were unspecified. Reports that studied the TDR rate in Honduras showed a prevalence of 9.2% in 2002-2003 and a prevalence of 7% in 2004-2007[44,58]. In Honduras and Panama, TDR was mainly restricted to NRTI and NNRTI drug classes, explained by the first-line therapy options in these countries[44,57].

Mexico

Reanalysis of Mexican sequences revealed TDR prevalence of 3.9% in the period 2002-2005. This result is substantially lower than the 16% reported by Escoto, et al.[59], which might be partly explained by the algorithm used to interpret drug resistance as it also included polymorphisms. However, the TDR prevalence of 16% also differed from values obtained in other studies[27,60], suggesting that the heterogeneity in ART availability between the different regions in Mexico might also impact TDR rates, although further research is warranted[60]. A recent comprehensive WHO-survey determined that Mexico has reached 6.8% of TDR for the period 2005-2010[45]. This overall rate of TDR remained stable at the national level. However, decreasing trends for NNRTI and increasing trends for PI TDR were observed. The frequency of mutations was 4.2% for NRTI (mainly type TAM I pathway), 1.9% for NNRTI (mostly 103N), and 1.8% for PI (90M)[45].

Andean Region

In the Andean Region, the TDR prevalence was 10.5% according to sequences sampled in the years 1999, 2001, 2003, and 2007. However, these sequences were mainly collected in Venezuela[46,61,62]. A WHO survey performed in Venezuela reported at least 7% TDR according to the French ANRS algorithm, and this value rose to 12.9% when the SDRM list was used for interpretation[46]. In this study, drug resistance to all three drug classes was observed. The high TDR rates may be a consequence of the early introduction of ART as part of an AIDS National Program that was initiated in 2000 and mainly focused in the capital where the surveys were performed. Additionally, one report mentioned that patients receiving ART and displaying poor adherence or engagement with the healthcare system were an important source of HIV-1 transmission and TDR in Venezuela[63].

In that era, free access to ART was also guaranteed in other countries belonging to that region (Fig. 2). However, scaling-up moved slowly and, for example, in Colombia and Peru scale-up reached only 34 and 57%, respectively, in 2010 according to UNAIDS[12]. One Colombian survey showed a TDR of 5.8% according to the HIVdb algorithm. The sequences from this study were not available for reanalysis with SDRM 2009, but other Colombian sequences that could be accessed did not reveal significant differences in TDR[64]. Two studies in Peru and Ecuador revealed low TDR rates of 3.3 and 4.3%, respectively, in 2002-2003[32,65]. However, it is worrying that all drug classes were affected in these countries. It is unknown whether polymorphisms might have affected these values as data was unavailable for reassessment with SDRM 2009.

Brazil

The overall TDR rate in Brazil was investigated by reanalyzing 1,056 sequences that covered the period 1998-2009 and this resulted in a value of 9.4%.

Before 1998, only a few small-sized studies were performed[26,66]. Barreto, et al.[22] reported a TDR prevalence of 5.6% in 1998-2002 in a group of 341 blood donors, using a WHO survey approach. Noteworthy, this study was able to include 16% recent infections and they displayed a TDR rate of 12%, indicating high transmission rates of drug-resistant strains in São Paulo City at that particular time point[22].

Several studies performed between 2000 and 2004, one WHO survey[24], and one large nationwide study that sampled in 2001[67], showed that the TDR was below 5%[35,49,68-71]. This apparent decrease of TDR could be related to the predominant inclusion of chronically infected individuals. However, other reports showed that TDR was still above 5%[72-74]. These discordances

are presumably due to the heterogeneity of the included studies and the geographical variation of TDR.

Since 2005, the TDR rates remained above 5%, with 7% in Curitiba and 8% in Porto Alegre and Santa Catarina[30,36,67,72]. Nationwide studies reported 5.2% TDR in 2007[75], 8.1% in 2008[39] and 12.3% in 2008-2010[43]. The apparent nationwide increase of TDR in Brazil could not be confirmed in our analysis. This might be explained by the large variations in TDR between and within Brazilian regions that were included in our dataset.

Since 2007, TDR per drug class remained stable, with overall values ranging between 5-15%, NNRTI 3-6%, NRTI 2-6%, and PI 1-3%[29,39,47,48,75]. The last WHO survey showed resistance against NRTI in 7.6%, followed by 4% against NNRTI and PI[43].

Southern Cone

As 84% of the Southern Cone sequences originated from Argentina, general conclusions cannot be made for this region. Nevertheless, their reanalysis using SDRM 2009 resulted in a TDR prevalence of 4.9% for the years 1996-2005. In Argentina, the number of studies and their respective sample size remained limited before 2000-2003[25,76]. However, since 2003 several nationwide surveillance studies have been developed that reported TDR rates that ranged from 3.9% in 2003-2005 to 8.4% in 2006-2008[31,33,41,77]. The last reported WHO survey showed a substantial level of resistance towards NNRTI (5.6%), NRTI (4.2%), and PI (3%)[33]. Another study that was performed in recently infected individuals between 2004 and 2009 also supported the 8% of TDR with predominance of NNRTI resistance[78]. Almost all Argentinean studies documented TDR against the three common drug classes. In Chile, TDR was less extensive, with an overall TDR of 6.2% in 2000-2005 that slightly decreased to 4% in 2006-2008[38,79]. In this country, TDR mutations were mainly restricted to NRTI and NNRTI.

Transmitted drug resistance in Latin America and the Caribbean: What do we know?

This review showed a TDR of 7.7% in 1,922 therapy naive HIV-1 patients. However, the majority of sequences were obtained from Brazil, Argentina, Mexico, and Chile, countries that correspond to half of the people living with HIV-1 in LAC. This percentage is higher than for other developing countries[80-82], but similar to the 6.4% reported by the WATCH study that included seven studies since 2000-2006[22,41,61,66,68,76,80,83], and similar to TDR rates in Europe, Canada, and the USA[2-5]. Although no significant trend in TDR was observed in this analysis, two peaks were observed in 1998-1999 and 2006-2007. These peaks overlapped, with a high inclusion of sequences from Brazil and Venezuela, countries characterized with the highest TDR rates in LAC.

Although the analysis has been standardized, implementing the SDRM 2009 list for interpretation of drug resistance, we acknowledge the limitations of this study. The included studies differed, for example, in the intrinsic characteristics of the analyzed population (age, risk factor of transmission, geographical and socioeconomic factors), design of the study (method of sampling, type of study, criteria or exclusion criteria), clinical characteristics (recent or chronic infection, immunological status, viral load), drug resistance testing (sample type, sequencing method, sequence quality, algorithm used for the analysis of mutations), ART-related factors in the treated population (mono- or bitherapy history, HAART scale-up, delay of start therapy, availability of new compounds, pharmacogenetics, toxicity, adherence problems, drug failure) and viral factors (subtype, fitness)[84-86]. For example, in one Brazilian study 20% TDR was reported in PBMC samples[34]. However, PBMC testing is more apt to detect drug resistance mutations than plasma, and the intrinsic regional characteristics of the population might impede the generalization of these results[34,87].

Although the MSM group is one of the main drivers of the LAC epidemic and they are 33.3 times more likely to be HIV-1 positive[88], we did not find differences in the TDR rates between heterosexuals and MSM. More studies are necessary to understand the trends of TDR, especially in MSM because the risk of imprisonment in some countries of the Caribbean and the stigma in LAC could hinder the accurate evaluation of this population, especially when TDR prevalence has been reported to range between 3.3% in Peru for 2002 to 17% in Brazil for 2010[32,88,89].

In regard to distribution of mutations, since NNRTI have been used widely in LAC as first-line therapy consequently, the amino acid change in position 103 was the most frequent in the sequences analyzed and in the studies included (Fig. 4 C). The amino acid change 215 and 41 were frequently found in this review. However, 184V was more frequently detected in this study than in European or North American reports, even though the difference was not statistically significant[2-5]. In general, similar patterns have been found in other developing countries but, in the resource-rich settings, NNRTI resistance seems to be stabilizing, probably

due to the changes in therapy prescriptions[4,81]. Overall, this review did not observe trends in particular drug resistance patterns in the analyzed sequences obtained from LAC.

Conclusions

The overall TDR prevalence in LAC was 7.7% for the period 1996-2009 and reached 10% in the last four investigated years when data was solely restricted to Brazil and Venezuela. Additionally, large geographical differences have been found. Although lack of information for the Caribbean and Central America prevents us from drawing general conclusions, the reported TDR in Cuba and other Caribbean islands is still under the 5% WHO threshold. However, Honduras, Mexico and South American countries have passed the 5% threshold, with the exception of Chile. The limited information available for most of the countries should encourage the development of surveys for drug resistance in newly infected people.

Acknowledgements

We thank Dr. Adriana Rojas-Villarraga and Gertjan Beheydt for methodological advice.

Disclaimers: The authors have no conflicts of interest to disclose.

Sources of support: The research work of AC Pineda-Peña was supported by the Doctoral Research Training Program "Francisco Jose de Caldas" by the *Departamento Administrativo de Ciencia, Tecnología e Innovación*, COLCIENCIAS, Republic of Colombia and by ERACOL, a scholarship for academic exchange in medicine and the health sciences between Europe and Latin America. This work was supported in part by the European Community's Seventh Framework Programme (FP7/2007-2013) under the project "Collaborative HIV and Anti-HIV Drug Resistance Network (CHAIN)" grant agreement N° 223131; by the Fonds voor Wetenschappelijk Onderzoek – Flanders (FWO) grant G.0611.09; and by the AIDS Reference Laboratory of Leuven that receives support from the Belgian Ministry of Social Affairs through a fund within the Health Insurance System.

References

1. Erice A, Mayers D, Strike D, et al. Brief report: primary infection with zidovudine-resistant human immunodeficiency virus type 1. N Engl J Med. 1993;328:1163-5.
2. Jayaraman G, Archibald C, Kim J, et al. A population-based approach to determine the prevalence of transmitted drug-resistant HIV among recent versus established HIV infections: results from the Canadian HIV strain and drug resistance surveillance program. J Acquir Immune Defic Syndr. 2006;42:86-90.
3. Weinstock H, Zaidi I, Heneine W, et al. The epidemiology of antiretroviral drug resistance among drug-naive HIV-1-infected persons in 10 US cities. J Infect Dis. 2004;189:2174-80.
4. Vercauteren J, Wensing A, van de Vijver D, et al. Transmission of drug-resistant HIV-1 is stabilizing in Europe. J Infect Dis. 2009;200:1503-8.
5. SPREAD. Transmission of drug-resistant HIV-1 in Europe remains limited to single classes. AIDS. 2008;22:625-35.
6. Vandamme A, Camacho R, Ceccherini-Silberstein F, et al. European recommendations for the clinical use of HIV drug resistance testing: 2011 update. AIDS Rev. 2011;13:77-108.
7. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Department of Health and Human Services. Available at http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf.; 2011;1-167.
8. Wittkop L, Gunthard H, de Wolf F, et al. Effect of transmitted drug resistance on virological and immunological response to initial combination antiretroviral therapy for HIV (EuroCoord-CHAIN joint project): a European multicohort study. Lancet Infect Dis. 2011;11:363-71.
9. UNAIDS-WHO. GLOBAL REPORT: UNAIDS Report on the global AIDS epidemic. Geneva; 2010.
10. Chequer P, Cuchi P, Mazin R, Garcia Calleja J. Access to antiretroviral treatment in Latin American countries and the Caribbean. AIDS. 2002;16(Suppl 3):S50-7.
11. Teixeira P, Vitoria M, Barcarolo J. Antiretroviral treatment in resource-poor settings: the Brazilian experience. AIDS. 2004;18(Suppl 3):S5-7.
12. UNAIDS-WHO. GLOBAL HIV/AIDS RESPONSE. Epidemic update and health sector progress towards Universal Access. Geneva, Switzerland; 2011.
13. Ravasi G, Pérez-Rosales M, Mejía Cruz V, et al. WHO HIV drug resistance early warning indicators as treatment program monitoring and evaluation tool in Central American countries. 6th IAS Conference on HIV Pathogenesis, Treatment and Prevention, Rome, Italy; 2011. [Abstract TUAD0105].
14. Jack N, Ravasi G, Schrooten W, Del Riego A, Hirnschall G. Implementing HIV drug resistance early warning indicators in the Caribbean. In: XVIII International AIDS Conference, Vienna, Austria; 2010. [Abstract THPE0354].
15. Jordan M, Kelley K, Saadani-Hassani A, Obeng-Aduasare Y, Bertagnolio S. Monitoring ART Clinic and Program Performance Using World Health Organization HIV Drug Resistance Early Warning Indicators. 18th CROI; Boston, USA; 2011.
16. Bennett D, Jordan M, Bertagnolio S, et al. HIV drug resistance early warning indicators in cohorts of individuals starting antiretroviral therapy between 2004 and 2009: World Health Organization global report from 50 countries. Clin Infect Dis. 2012;54(Suppl 4):S280-9.
17. Tuboi S, Schechter M, McGowan C, et al. Mortality during the first year of potent antiretroviral therapy in HIV-1-infected patients in 7 sites throughout Latin America and the Caribbean. J Acquir Immune Defic Syndr. 2009;51:615-23.
18. Crabtree-Ramírez B, Caro-Vega Y, Shepherd B, et al. Cross-sectional analysis of late HAART initiation in Latin America and the Caribbean: Late testers and late presenters. PLoS One. 2011;6:e20272.
19. Bennett D, Bertagnolio S, Sutherland D, Gilks C. The World Health Organization's global strategy for prevention and assessment of HIV drug resistance. Antivir Ther. 2008;13(Suppl 2):1-13.
20. Bennett D, Myatt M, Bertagnolio S, Sutherland D, Gilks C. Recommendations for surveillance of transmitted HIV drug resistance in countries scaling up antiretroviral treatment. Antivir Ther. 2008;13(Suppl 2):25-36.
21. Bennett D, Camacho R, Otelea D, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS One. 2009;4:e4724.
22. Barreto C, Nishyia A, Araujo L, Ferreira J, Busch M, Sabino E. Trends in antiretroviral drug resistance and clade distributions among HIV-1--infected blood donors in Sao Paulo, Brazil. J Acquir Immune Defic Syndr. 2006;41:338-41.
23. Brindeiro R, Vanderborght B, Caride E, et al. Sequence diversity of the reverse transcriptase of human immunodeficiency virus type 1 from untreated Brazilian individuals. Antimicrob Agents Chemother. 1999;43:1674-80.
24. Sa-Ferreira J, Brindeiro P, Chequer-Fernandez S, Tanuri A, Morgado M. Human immunodeficiency virus-1 subtypes and antiretroviral drug resistance profiles among drug-naive Brazilian blood donors. Transfusion. 2007;47:97-102.
25. Pando M, Eyzaguirre L, Carrion G, et al. High genetic variability of HIV-1 in female sex workers from Argentina. Retrovirology. 2007;4:58.
26. Maia Teixeira S, Bastos F, Hacker M, Guimaraes M, Morgado M. Trends in drug resistance mutations in antiretroviral-naive intravenous drug users of Rio de Janeiro. J Med Virol. 2006;78:764-9.
27. Eyzaguirre L, Brouwer K, Nadai Y, et al. First molecular surveillance report of HIV type 1 in injecting drug users and female sex workers along the U.S.-Mexico border. AIDS Res Hum Retroviruses. 2007;23:331-4.
28. Janssen R, Satten G, Stramer S, et al. New testing strategy to detect early HIV-1 infection for use in incidence estimates and for clinical and prevention purposes. JAMA. 1998;280:42-8.

29. Arruda E, Simoes L, Sucupira C, et al. Short communication: intermediate prevalence of HIV type 1 primary antiretroviral resistance in Ceara State, Northeast Brazil. AIDS Res Hum Retroviruses. 2011;27:153-6.
30. Brigido L, Nunes C, Oliveira C, et al. HIV type 1 subtype C and CB Pol recombinants prevail at the cities with the highest AIDS prevalence rate in Brazil. AIDS Res Hum Retroviruses. 2007;23:1579-86.
31. Dilernia D, Lourtau L, Gomez A, et al. Drug-resistance surveillance among newly HIV-1 diagnosed individuals in Buenos Aires, Argentina. AIDS. 2007;21:1355-60.
32. Lama J, Sanchez J, Suarez L, et al. Linking HIV and antiretroviral drug resistance surveillance in Peru: a model for a third-generation HIV sentinel surveillance. J Acquir Immune Defic Syndr. 2006;42:501-5.
33. Pando M, Gomez-Carrillo M, Vignoles M, et al. Incidence of HIV type 1 infection, antiretroviral drug resistance, and molecular characterization in newly diagnosed individuals in Argentina: A Global Fund Project. AIDS Res Hum Retroviruses. 2011;27:17-23.
34. Sucupira M, Caseiro M, Alves K, et al. High levels of primary antiretroviral resistance genotypic mutations and B/F recombinants in Santos, Brazil. AIDS Patient Care STDS. 2007;21:116-28.
35. Varella R, Ferreira S, de Castro M, Zalis M, Tavares M. Human immunodeficiency virus type 1 protease and reverse transcriptase mutation patterns among treatment-naive patients in different stages of infection in Rio de Janeiro, Brazil. J Med Virol. 2007;79:1033-9.
36. Ferreira J, Thomaz M, Rodrigues R, et al. Molecular characterisation of newly identified HIV-1 infections in Curitiba, Brazil: preponderance of clade C among males with recent infections. Mem Inst Oswaldo Cruz. 2008;103:800-8.
37. Acevedo W, Gallardo A, Galaz J, Afani A, Cortes E. [Detection of primary antiretroviral resistance in Chilean patients recently infected with human immunodeficiency virus (HIV)]. Rev Med Chil. 2007;135:1406-13.
38. Afani A, Beltran C, Maria Gallardo A, Roessler P, Acevedo W, Vasquez P. [Prevalence of primary antiretroviral resistance among HIV infected patients in Chile]. Rev Med Chil. 2010;138:669-76.
39. Inocencio L, Pereira A, Sucupira M, et al. Brazilian Network for HIV Drug Resistance Surveillance: a survey of individuals recently diagnosed with HIV. J Int AIDS Soc. 2009;12:20.
40. Kijak G, Pampuro S, Gomez Carrillo M, et al. Resistance Profiles to Antiretroviral Drugs in HIV-1-Naive Patients. In: 7th CROI; San Francisco, USA; 2000.
41. Petroni A, Deluchi G, Pryluka D, et al. Update on primary HIV-1 resistance in Argentina: emergence of mutations conferring high-level resistance to nonnucleoside reverse transcriptase inhibitors in drug-naive patients. J Acquir Immune Defic Syndr. 2006;42:506-10.
42. Arreseigor T, Valle M, Rimoldi I. Primary drug resistance in newly diagnosed HIV-1 infected individuals in Buenos Aires province. In: XVIII IAC: Vienna, Austria; 2010. [Abstract CDC0464].
43. De Moraes Soares C, Vergara T, Sucupira M, et al. Prevalence of transmitted HIV-1 antiretroviral resistance among patients initiating ART in Brazil: A surveillance using dried blood spots. In: 18th CROI. Boston, USA; 2011.
44. Murillo W, Paz-Bailey G, Morales S, et al. Transmitted drug resistance and type of infection in newly diagnosed HIV-1 individuals in Honduras. J Clin Virol. 2010;49:239-44.
45. Avila-Rios S, Garcia-Morales C, Garrido-Rodríguez D, et al. National Prevalence and Trends of HIV Transmitted Drug Resistance in Mexico. In: 18th CROI. Boston, USA; 2011.
46. Castillo J, Comegna M, Quijada W, et al. Surveillance of HIV type 1 drug resistance among naive patients from Venezuela. AIDS Res Hum Retroviruses. 2009;25:1329-33.
47. Cardoso L, Queiroz B, Stefani M. HIV-1 pol phylogenetic diversity and antiretroviral resistance mutations in treatment naive patients from Central West Brazil. J Clin Virol. 2009;46:134-9.
48. Carvalho B, Cardoso LP, Damasceno S, Stefani M. Moderate prevalence of transmitted drug resistance and interiorization of HIV Type 1 Subtype C in the Inland North State of Tocantins, Brazil. AIDS Res Hum Retroviruses. 2011;271081-7.
49. Medeiros M, Arruda E, Guerrant R, et al. Genotype testing and antiretroviral resistance profiles from HIV-1 patients experiencing therapeutic failure in northeast Brazil. Braz J Infect Dis. 2007;11:390-4.
50. 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.
51. UNAIDS-WHO. Progress on Global Access to HIV Antiretroviral Therapy: A Report on "3 by 5" and Beyond. Geneva, Switzerland: WHO; 2006.
52. Gittens M, Roth W, Roach T, et al. The molecular epidemiology and drug resistance determination of HIV type 1 subtype B infection in Barbados. AIDS Res Hum Retroviruses. 2003;19:313-19.
53. Perez L, Alvarez L, Carmona R, et al. Genotypic resistance to antiretroviral drugs in patients infected with several HIV type 1 genetic forms in Cuba. AIDS Res Hum Retroviruses. 2007;23:407-14.
54. Ruibal-Brunet I, Cuevas M, Diaz-Torres H, et al. Genotypic resistance mutations to antiretroviral drugs in HIV-1 B and non-B subtypes from Cuba. Rev Panam Salud Publica. 2001;10:174-80.
55. Vaughan H, Cane P, Pillay D, Tedder R. Characterization of HIV type 1 clades in the Caribbean using pol gene sequences. AIDS Res Hum Retroviruses. 2003;19:929-32.
56. Ahumada-Ruiz S, Flores-Figueroa D, Toala-Gonzalez I, Thomson M. Analysis of HIV-1 pol sequences from Panama: identification of phylogenetic clusters within subtype B and detection of antiretroviral drug resistance mutations. Infect Genet Evol. 2009;9:933-40.
57. Castillo J, Mendoza Y, Martinez A, et al. Prevalence of transmitted antiretroviral drug resistance in recently HIV infected individuals and newborns from Panama. XVIII IAC, Vienna, Austria; 2010. [Abstract THPE0358].
58. Lloyd B, O'Connell R, Michael N, et al. Prevalence of resistance mutations in HIV-1-Infected Hondurans at the beginning of the National Antiretroviral Therapy Program. AIDS Res Hum Retroviruses. 2008;24:529-35.
59. Escoto-Delgadillo M, Vazquez-Valls E, Ramirez-Rodriguez M, et al. Drug-resistance mutations in antiretroviral-naive patients with established HIV-1 infection in Mexico. HIV Med. 2005;6:403-9.
60. Valle-Bahena O, Ramos-Jimenez J, Ortiz-Lopez R, et al. Frequency of protease and reverse transcriptase drug resistance mutations in naive HIV-infected patients. Arch Med Res. 2006;37:1022-7.
61. Bouchard M, Masquelier B, Moreno M, et al. HIV type 1 drug resistance among naive patients from Venezuela. AIDS Res Hum Retroviruses. 2007;23:482-5.
62. Delgado E, Leon-Ponte M, Villahermosa M, et al. Analysis of HIV type 1 protease and reverse transcriptase sequences from Venezuela for drug resistance-associated mutations and subtype classification: a UNAIDS study. AIDS Res Hum Retroviruses. 2001;17:753-8.
63. Hernandez M, Matos D, Rodrigues R, Cedeño C, Rivera A. Resumen del estudio sobre perfil del paciente con VIH y adherencia al tratamiento antirretroviral en Venezuela. Venezuela: Ministerio del Poder Popular para la Salud; 2008.
64. DiazGranados C, Mantilla M, Lenis W. Antiretroviral drug resistance in HIV-infected patients in Colombia. Int J Infect Dis. 2010;14:e298-303.
65. Guanira J, Lama J, Montoya O, et al. HIV drug resistance in the Andean Region: a look after the universal access to HAART in the Andean region. In: 18th International HIV Drug Resistance Workshop Basic Principles & Clinical Implications; Fort Myers, USA. Antivir Ther. 2009;14(Suppl 1):A174. [Abstract 151 2009].
66. Dumans A, Soares M, Pieniazek D, et al. Prevalence of protease and reverse transcriptase drug resistance mutations over time in drug-naive human immunodeficiency virus type 1-positive individuals in Rio de Janeiro, Brazil. Antimicrob Agents Chemother. 2002;46:3075-9.
67. Brindeiro R, Diaz R, Sabino E, et al. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. AIDS. 2003;17:1063-9.
68. Soares M, Brindeiro R, Tanuri A. Primary HIV-1 drug resistance in Brazil. AIDS. 2004;18(Suppl 3):S9-13.
69. Soares E, Santos R, Pellegrini J, Sprinz E, Tanuri A, Soares M. Epidemiologic and molecular characterization of human immunodeficiency virus type 1 in southern Brazil. J Acquir Immune Defic Syndr. 2003;34:520-6.
70. Eyer-Silva W, Morgado M. A genotyping study of human immunodeficiency virus type-1 drug resistance in a small Brazilian municipality. Mem Inst Oswaldo Cruz. 2005;100:869-73.
71. Lacerda H, Medeiros L, Cavalcanti A, Ximenes R, Albuquerque M. Comparison of the epidemiology, profile of mutations, and clinical response to antiretrovirals among subtypes B and F of the human immunodeficiency virus type 1. Mem Inst Oswaldo Cruz. 2007;102:693-9.
72. Gonsalez C, Alcalde R, Nishiya A, et al. Drug resistance among chronic HIV-1-infected patients naive for use of anti-retroviral therapy in Sao Paulo city. Virus Res. 2007;129:87-90.
73. Pires I, Soares M, Speranza F, et al. Prevalence of human immunodeficiency virus drug resistance mutations and subtypes in drug-naive, infected individuals in the army health service of Rio de Janeiro, Brazil. J Clin Microbiol. 2004;42:426-30.
74. Rodrigues R, Scherer L, Oliveira C, et al. Low prevalence of primary antiretroviral resistance mutations and predominance of HIV-1 clade C at polymerase gene in newly diagnosed individuals from south Brazil. Virus Res. 2006;116:201-7.
75. Sprinz E, Netto E, Patelli M, et al. Primary antiretroviral drug resistance among HIV type 1-infected individuals in Brazil. AIDS Res Hum Retroviruses. 2009;25:861-7.
76. Kijak G, Pampuro S, Avila M, et al. Resistance profiles to antiretroviral drugs in HIV-1 drug-naive patients in Argentina. Antivir Ther. 2001;6:71-7.
77. Rubio A, Pando M, Dilernia D, Trinchero J, Rovner N, Salomon H. Primary resistance in newly diagnosed HIV-1 infected drug-naïve individuals from Buenos Aires, Argentina: a two-period comparison. In: 18th International HIV Drug Resistance Workshop Basic Principles & Clinical Implications; Fort Myers, Florida. Antivir Ther. 2009;14(Suppl 1):A175. [Abstract 152].
78. Petroni A, Deluchi G, Lopardo G, et al. Update on transmission of HIV resistance to drug-naive patients in Argentina. In: International HIV & Hepatitis Virus Drug Resistance Workshop and Curative Strategies; Dubrovnik, Croatia. Antivir Ther. 2010;15(Suppl 2): A179. [Abstract 141].
79. Rios M, Delgado E, Perez-Alvarez L, et al. Antiretroviral drug resistance and phylogenetic diversity of HIV-1 in Chile. J Med Virol. 2007;79:647-56.
80. Frentz D, Van de Vijver D, Wensing A, Bowles E, Boucher C. Worldwide comparison of mutational patterns in studies on transmission of drug resistant HIV-1 across the World. Update on WATCH. In: 6th European HIV Drug Resistance Workshop. Budapest, Hungary; 2008.

81. Hamers R, Derdelinckx I, van Vugt M, Stevens W, Rinke de Wit T, Schuurman R. The status of HIV-1 resistance to antiretroviral drugs in sub-Saharan Africa. Antivir Ther. 2008;13:625-39.
82. Shekelle P, Maglione M, Geotz M, et al. Antiretroviral (ARV) drug resistance in the developing world. Evid Rep Technol Assess (Full Rep). 2007;156:1-74.
83. Afani A, Ayala M, Meyer A, Cabrera R, Acevedo W. [Primary resistance to antiretroviral therapy in patients with HIV/AIDS in Chile]. Rev Med Chil. 2005;133:295-301.
84. Masquelier B, Bhaskaran K, Pillay D, et al. Prevalence of transmitted HIV-1 drug resistance and the role of resistance algorithms: data from seroconverters in the CASCADE collaboration from 1987 to 2003. J Acquir Immune Defic Syndr. 2005;40:505-11.
85. Booth C, Geretti A. Prevalence and determinants of transmitted antiretroviral drug resistance in HIV-1 infection. J Antimicrob Chemother. 2007;59:1047-56.
86. Little S, Smith D. HIV treatment decisions and transmitted drug resistance. Clin Infect Dis. 2005;41:233-5.
87. Bon I, Gibellini D, Borderi M, et al. Genotypic resistance in plasma and peripheral blood lymphocytes in a group of naive HIV-1 patients. J Clin Virol. 2007;38:313-20.
88. Beyrer C, Baral S, Walker D, Wirtz A, Johns B, Sifakis F. The expanding epidemics of HIV type 1 among men who have sex with men in low- and middle-income countries: diversity and consistency. Epidemiol Rev. 2010;32:137-51.
89. Bermudez E, Kerr L, Pinho A, et al. Antiretroviral drug resistance in men who have sex with men in Brazil. In: XVIII IAC. Vienna, Austria; 2010. [Abstract FRLBC103].
90. ONUSIDA-Republica Bolivariana de Venezuela. Informe Anual 2009. Fortaleciendo la Respuesta Nacional y Conjunta al VIH y sida. Caracas, Venezuela; 2009.
91. Soares E, Martinez A, Souza T, et al. HIV-1 subtype C dissemination in southern Brazil. AIDS. 2005;19(Suppl 4):S81-6.
92. de Medeiros L, Lacerda H, Cavalcanti A, de Albuquerque M. Primary resistance of human immunodeficiency virus type 1 in a reference center in Recife, Pernambuco, Brazil. Mem Inst Oswaldo Cruz. 2006;101:845-9.
93. Eyer-Silva W, Couto-Fernandez J, Silva-de-Jesus C, Morgado M. Prevalence of HIV type 1 drug resistance mutations in treatment-naive and experienced patients from resource-limited settings with universal access to antiretroviral therapy: a survey in two small Brazilian cities. Mem Inst Oswaldo Cruz. 2008;103:143-9.
94. de Sa-Filho D, Ambar R, Duarte N, et al. HIV type 1 diversity from newly diagnosed patients in Santos metropolitan area/Brazil. AIDS Res Hum Retroviruses. 2009;25:925-9.
95. Rangel H, Garzaro D, Torres J, et al. Prevalence of antiretroviral drug resistance among treatment-naive and treated HIV-infected patients in Venezuela. Mem Inst Oswaldo Cruz. 2009;104:522-5.
96. Fuentes-Romero L, Rodriguez-Diaz R, Viveros-Rogel M, et al. Genotypic HIV Drug resistance among newly diagnosed, never treated persons in Mexico. In: 13th International HIV Drug Resistance Workshop; Canary Islands, Spain. Antivir Ther. 2004;9:S111.
97. Garrido-Rodríguez D, Avila-Rios S, Vázquez-Pérez J, et al. Molecular epidemiology of HIV in Mexico: a clade B-dominated epidemic with low prevalence of primary drug resistance. In: XVII IAC. Mexico City, Mexico; 2008. [Abstract TUPE0019].
98. Escoto-Delgadillo M, López-Lomelí M, Gonzáles-Morales L, Demi-Castellanos E, Vásquez-Valls E. Frequency of primary mutations to antiretrovirals in Mexican population In: XIV IAC. Barcelona, Spain; 2002. [Abstract B10411].
99. Andrade-Villanueva J, Amador-Lara F, Ortiz-Macias X, Mora-Jimenez E. Primary HIV Drug Resitance in Referral HIV Unit in Mexico. In: 42nd Annual Meeting of IDSA. USA; 2004. [Abstract 910].
100. Rodriguez-Diaz R, Fuentes-Romero L, Viveros-Rogel M, et al. Does Mexico need to use genotypic resistance testing before starting treatment in naive patients with recent HIV diagnosis? In: 4th IAS Conference on HIV Pathogenesis, Treatment and Prevention: Bangkok, Thailand; 2007. [Abstract CDB235].
101. Silva F, Andrade J, Espinosa-Aguilar L, et al. Primary genotypic resistance in chronically infected patients screened for clinical trials in Mexico. In: XVII IAC: Mexico City, Mexico; 2008. [Abstract TUPE0032].
102. Patterson P, Krolewiecki A, Vázquez L, et al. Frequency distribution of mutations in HIV-1 subtypes B and BF in treatment-naive patients from Argentina. In: XVII IAC: Mexico City, Mexico; 2008. [Abstract TUPE0021].
103. Parenti P, David D, Casiro A, et al. Prevalence of primary drug resistance in HIV-1 infected patients in Argentina. In: XVII IAC: Mexico City, Mexico; 2008. [Abstract MOPDA206].
104. Argel M, Gonzalez P, Esposto A, et al. HIV-1 Genomic characterization from naive patients in La Plata City, province of Buenos Aires, Argentina. In: XVII IAC: Mexico City, Mexico; 2008. [Abstract CDA0033].
105. Dias-Tavares M, Brandão-Varella R, Baía-Fereira S, Braga de Castro M, Correa-Malta F, Schechter M. Large scale use of antiretroviral drugs is not associated with a high prevalence of genotypic resistance mutations in treatment-naive HIV-1 infected individuals in Rio de Janeiro, Brazil. In: 2nd IAS Conference on HIV Pathogenesis and Treatment: Paris, France; 2003. [Abstract 778].
106. Velasco de Castro C, Grinsztejn V, Veloso V, et al. HIV-1 subtypes and antiretroviral drug resistance mutations among recent (RS) and long-term seroconverters (LTS) from VCTs of Rio de Janeiro state, Brazil. In: XVI IAC: Toronto, Canada; 2006. [Abstract MOPE0385].
107. Waléria-Aleixo A, Tanuri A, Cleto S, Greco D. Mutations related to primary resistance to antiretrovirals among newly HIV-1 infected individuals in Belo Horizonte-Brazil. In: 2nd IAS Conference on HIV Pathogenesis and Treatment: Paris, France; 2003. [Abstract 242].
108. Rodrigues R, Oliveira C, Custodio R, et al. Diversity in Sáo Paulo: low ARV resistance in HIV incident infections. In: XV IAC: Bangkok, Thailand; 2004. [Abstract MoPeC3426].
109. Couto-Fernandez J, Brindeiro R, Inocencio L, et al. Time trends in primary resistance to HIV drugs in the chronically infected ARV-naïve population in Rio de Janeiro, Brazil. In: International HIV & Hepatitis Virus Drug Resistance Workshop and Curative Strategies; Dubrovnik, Croatia: Antivir Ther. 2010;15(Suppl 1):A152.
110. Gagliani L, Alkmim Maia W, Caseiro M, Sá-Filho D, Teixeira C, Diaz R. Primary drug antiretroviral resistance and HIV-1 subtype among antiretroviral-naïve patients immediately before starting HAART in Santos, Brazil. XVII IAC: Mexico City, Mexico; 2008. [Abstract TUPE0025].
111. Sucupira M, Sanabani S, Côrtes R, et al. Primary antiretroviral resistance among individuals with recent HIV infection in the city of São Paulo, Brazil: use of full-length genome sequencing to evaluate six antiretroviral classes. In: 18th International HIV Drug Resistance Workshop Basic Principles & Clinical Implications; Fort Myers, USA. Antivir Ther. 2009;14(Suppl 1):A172. [Abstract 149].
112. Rodrigues R, Knoll R, Nunes C, et al. Low primary HIV drug resistance mutations in South Brazil. In: XVII IAC: Mexico City, Mexico; 2008. [Abstract TUPE0027].
113. Sousa T, Dumans A, Kurck I, Moreira de Sá C, Sion F, Soares M. Increasing in primary antiretroviral resistance among untreated HIV-infected individuals in Rio de Janeiro city, Brazil. In: 4th IAS Conference on HIV Pathogenesis, Treatment and Prevention: Bangkok, Thailand; 2007. [Abstract CDB234].
114. Ambar R, Duarte N, Thome M, et al. HIV-1 pol gene diversity and antiretroviral drug resistance mutations in drug naïve patients from Santos metropolitan area, Brazil. In: XVII IAC: Mexico City, Mexico; 2008. [Abstract MOPDA203].
115. Diaz R, De Souza R, Sucupira M, et al. Lower prevalence of primary antiretroviral resistance among clade C infected individuals as compared to clade B in South Brazil's emerging clade C HIV-1 epidemic In: XVII IAC: Mexico City, Mexico; 2008. [Abstract TUAA02].

OPEN ACCESS Freely available online

PLOS ONE

Trends and Predictors of Transmitted Drug Resistance (TDR) and Clusters with TDR in a Local Belgian HIV-1 Epidemic

Andrea-Clemencia Pineda-Peña[1,2], Yoeri Schrooten[1,3], Lore Vinken[1], Fossie Ferreira[1], Guangdi Li[1], Nídia Sequeira Trovão[1], Ricardo Khouri[1], Inge Derdelinckx[4,5], Paul De Munter[4,5], Claudia Kücherer[6], Leondios G. Kostrikis[7], Claus Nielsen[8], Kirsi Littsola[9], Annemarie Wensing[10], Maja Stanojevic[11], Roger Paredes[12], Claudia Balotta[13], Jan Albert[14,15], Charles Boucher[16], Arley Gomez-Lopez[2], Eric Van Wijngaerden[4,5], Marc Van Ranst[1,3], Jurgen Vercauteren[1], Anne-Mieke Vandamme[1,17], Kristel Van Laethem[1,3]*

1 Clinical and Epidemiological Virology, Rega Institute for Medical Research, Department of Microbiology and Immunology, KU Leuven, Leuven, Belgium, **2** Clinical and Molecular Infectious Diseases Group, Faculty of Sciences and Mathematics, Universidad del Rosario, Bogotá, Colombia, **3** AIDS Reference Laboratory, University Hospitals Leuven, Leuven, Belgium, **4** Clinical Infectious and Inflammatory Disorders, Department of Microbiology and Immunology, KU Leuven, Leuven, Belgium, **5** Internal Medicine, University Hospitals Leuven, Leuven, Belgium, **6** Robert Koch-Institut, Berlin, Germany, **7** University of Cyprus, Nicosia, Cyprus, **8** Statens Serum Institut, Copenhagen, Denmark, **9** National Institute of health and welfare, Helsinki, Finland, **10** Department of Virology, University Medical Center Utrecht, The Netherlands, **11** University of Belgrade, Faculty of Medicine, Belgrade, Serbia, **12** IrsiCaixa Foundation, Badalona, Spain, **13** Luigi Sacco University Hospital, Milan, Italy, **14** Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden, **15** Department of Clinical Microbiology, Karolinska University Hospital, Stockholm, Sweden, **16** Erasmus MC, University Medical Center, Rotterdam, The Netherlands, **17** Centro de Malária e outras Doenças Tropicais and Unidade de Microbiologia, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisboa, Portugal

Abstract

We aimed to study epidemic trends and predictors for transmitted drug resistance (TDR) in our region, its clinical impact and its association with transmission clusters. We included 778 patients from the AIDS Reference Center in Leuven (Belgium) diagnosed from 1998 to 2012. Resistance testing was performed using population-based sequencing and TDR was estimated using the WHO-2009 surveillance list. Phylogenetic analysis was performed using maximum likelihood and Bayesian techniques. The cohort was predominantly Belgian (58.4%), men who have sex with men (MSM) (42.8%), and chronically infected (86.5%). The overall TDR prevalence was 9.6% (95% confidence interval (CI): 7.7–11.9), 6.5% (CI: 5.0–8.5) for nucleoside reverse transcriptase inhibitors (NRTI), 2.2% (CI: 1.4–3.5) for non-NRTI (NNRTI), and 2.2% (CI: 1.4–3.5) for protease inhibitors. A significant parabolic trend of NNRTI-TDR was found ($p = 0.019$). Factors significantly associated with TDR in univariate analysis were male gender, Belgian origin, MSM, recent infection, transmission clusters and subtype B, while multivariate and Bayesian network analysis singled out subtype B as the most predictive factor of TDR. Subtype B was related with transmission clusters with TDR that included 42.6% of the TDR patients. Thanks to resistance testing, 83% of the patients with TDR who started therapy had undetectable viral load whereas half of the patients would likely have received a suboptimal therapy without this test. In conclusion, TDR remained stable and a NNRTI up-and-down trend was observed. While the presence of clusters with TDR is worrying, we could not identify an independent, non-sequence based predictor for TDR or transmission clusters with TDR that could help with guidelines or public health measures.

Citation: Pineda-Peña A-C, Schrooten Y, Vinken L, Ferreira F, Li G, et al. (2014) Trends and Predictors of Transmitted Drug Resistance (TDR) and Clusters with TDR in a Local Belgian HIV-1 Epidemic. PLoS ONE 9(7): e101738. doi:10.1371/journal.pone.0101738

Editor: Cecilio López-Galíndez, Centro Nacional de Microbiología - Instituto de Salud Carlos III, Spain

Received February 15, 2014; **Accepted** June 10, 2014; **Published** July 8, 2014

Funding: This work was supported in part by the AIDS Reference Laboratory of Leuven that receives support from the Belgian Ministry of Social Affairs through a fund within the Health Insurance System; by the Fonds voor Wetenschappelijk Onderzoek – Flanders (FWO) grant G069214N; and by the European Community's Seventh Framework Programme (FP7/2007-2013) under the project "Collaborative HIV and Anti-HIV Drug Resistance Network (CHAIN)" grant agreement N° 223131; by the University of Leuven (Program Financing no. PF/10/018); by the Doctoral Research Training Program "Francisco Jose de Caldas" by the Departamento Administrativo de Ciencia, Tecnología e Innovación, COLCIENCIAS, Republic of Colombia and by ERACOL, a scholarship for academic exchange in medicine and the health sciences between Europe and Latin America to A.C.P.P. The computational resources and services used in this work were provided by the Hercules Foundation and the Flemish Government – department EWI-FWO Krediet aan Navorsers (Theys, KAN2012 1.5.249.12). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: kristel.vanlaethem@uzleuven.be

Introduction

In recent years, the number of newly diagnosed HIV-1 patients increased in Belgium [1] with a rate of 10.7 per 100,000 population in 2011, one of the highest rates in Europe [2]. Studies carried out in Europe and America highlighted the important role of transmission networks in the spread of transmitted drug resistance (TDR) [3–7]. TDR is a clinical and public health issue because it can compromise the response to antiretroviral therapy (ART) at the individual and population level

[8]. Three nationwide studies were performed previously in Belgium and reported a TDR prevalence of 29% (67/231; 95% CI: 23.5–35.2) between 1995 and 1998 [9], 7.2% (6/83; 95% CI: 3.4–14.9) in 2000 [10] and 9.5% (27/285, 95% CI: 6.6–13.4) between 2003 and 2006 [11]. However, due to differences in methodology and the lack of a recent study, no up-to-date information is yet available on TDR trends in Belgium. Nevertheless, recent reports revealed the rapid onward transmission of an HIV-1 strain with K103N mutation [12] and the involvement of transmission clusters (TCs) in approximately half of patients with TDR [4] in a local HIV epidemic in Belgium.

Because other studies consistently showed regional differences between the drivers of the HIV-1 epidemic [13,14], this study aimed to characterize the temporal trend in TDR, the factors associated with TDR including TCs and the clinical impact of TDR for a period of 15 years in a regional epidemic, serviced by the Leuven University Hospitals. The data included socio-demographic, clinical and virological variables.

Materials and Methods

Ethics Statement

The research was conducted according to the Declaration of Helsinki. Only patients for whom written informed consent was obtained were included in this study, except patients enrolled in care after 2009. In 2009, UZ Leuven implemented a generic "opt out" system. Patients, who logged an objection to use their medical data for research purposes, were not included in this study. The protocol and this consent procedure were approved by the Ethical Committee UZ Leuven (reference ML-8627, approval B322201316521 S52637).

Study Population

We analysed data from the cohort of the AIDS Reference Centre (ARC) in Leuven, the capital of the province of Flemish Brabant (Belgium). The ARC in Leuven has been collecting information since 1997 on treated HIV-1 patients and since 1999, also for naive HIV-1 patients, including epidemiological, clinical and virological data, related with the routine patient healthcare services. The prospective clinical use of baseline genotypic drug resistance testing was implemented in 1999 and stored plasma samples from before 1999 were available to retrospectively perform drug resistance testing upon clinician's request. Therefore HIV-1 sequences for drug naive patients were either prospectively or retrospectively obtained from a sample taken at diagnosis, except for 135 patients for whom a later pre-therapy sample was used. The inclusion criteria for the analysis of TDR in the present study were newly HIV-1 diagnosed between January 1998 and December 2012, availability of a nucleotide sequence before antiviral therapy initiation and age older than 18 years, and this cohort was called the Leuven newly-diagnosed (ND) cohort for the purpose of this study. The only exclusion criterion used was documented vertical transmission. Recent infections were defined using clinical and laboratory information such as p24 ELISA, HIV-specific antibody ELISA, and Inno-Lia profile. Patients with the following criteria were classified as recently infected: Fiebig stages I-V [15] or no more than 6 months difference between the last seronegative and first seropositive HIV-1 test [11], CD4 count >200 cells/μl and absence of AIDS-defining conditions [16].

Drug Resistance Testing

Drug resistance testing was performed using population-based Sanger sequencing of the *pol* gene fragment encoding protease (PR) (amino acids 1 to 99) and 5′-prime end of reverse transcriptase (RT) (amino acids 1 to 320). Sequences were obtained using the ViroSeq HIV-1 Genotyping System version 2 (Celera Diagnostics, Alameda, CA) or with an in-house method upon failure of the commercial test [17]. Sequences with associated information are available through Euresist (http://www.euresist.org).

TDR mutations were defined according to the 2009 list of surveillance drug resistance mutations from the World Health Organization [18]. Therefore, the nucleotide sequences were submitted to the Calibrated Population Resistance tool version 6.0 (http://cpr.stanford.edu/cpr.cgi). The clinical impact of genotypic drug resistance on first line therapy was evaluated using Rega algorithm [19] version 9.1.0 (available at http://rega.kuleuven.be/cev/avd/software/rega-algorithm).

HIV-1 Subtyping

HIV-1 subtypes and circulating recombinant forms (CRF) were determined using two HIV-1 subtyping tools, namely Rega version 3 (http://www.bioafrica.net/typing-v3/hiv) and COMET version 0.3 (http://comet.retrovirology.lu/) [20–22]. Sequences with discordant results were analyzed using manual phylogenetic analysis as was explained previously [22]. Briefly, maximum likelihood (ML) phylogenetic trees under the GTR+Γ nucleotide substitution model were built with RAxML [23] and recombination was verified using SimPlot [24].

Transmission Cluster Analysis

To investigate the factors associated with TDR and onward transmission of TDR, cluster analyses were performed on the Leuven ND cohort and four additional datasets as controls: (i) all other *pol* sequences from the ARC in Leuven, including treated HIV-1 patients, HIV-1 patients younger than 18 years old and HIV-1 patients with vertical mode of transmission, (ii) HIV-1 *pol* sequences obtained with the search term "Belgium" as sampling country in the Los Alamos HIV sequence database (retrieved from http://www.hiv.lanl.gov date in April 2013), (iii) HIV-1 *pol* sequences from the collaborative study SPREAD that enrolled patients with newly diagnosed HIV-1 infection from 22 European countries including Belgium between 2002 and 2008 (see details of the study in [25,26]), and (iv) the 30 most similar sequences to the Leuven ND cohort (retrieved by Basic Local Alignment Search Tool (BLAST) from http://blast.ncbi.nlm.nih.gov/Blast.cgi). The quality control of the sequences was performed using the tool available at http://www.hiv.lanl.gov/content/sequence/QC/ and the criteria previously described [27]. Separate datasets were constructed according to subtype. As an out-group, two or three reference sequences of subtype D or B (retrieved from http://www.hiv.lanl.gov) were included for B and non-B subtypes, respectively. Sequences were aligned with Muscle as accessory application in the program Mega version 5 [28,29]. Duplicates were removed and the positions encoding surveillance drug resistance mutations were excluded [18], which resulted in an average final length of 950 nucleotides. In the resulting dataset, we had 755 sequences for subtype A, 4225 for subtype B, 1036 for subtype C, 202 for subtype F, 665 for subtype G, 440 for CRF01_AE and 677 for CFR02_AG to perform phylogenetic analyses on. Subtypes with frequencies less than 1% were not included in TCs analyses.

A ML tree was inferred with the nucleotide substitution GTR+Γ model and 1000 bootstrap replicates in RAxML [23]. TCs, including pairs (two individuals) and larger clusters ($\geq$3 individuals), were identified by using Cluster Picker (retrieved from http://hiv.bio.ed.ac.uk/software.html) [30] with a genetic distance less or equal than 0.06 substitutions per site and bootstrap support

≥98%. A sensitivity analysis was performed to evaluate the effect of other genetic distances (0.015, 0.030 and 0.045) [31–34].

The robustness of the identified TCs was evaluated using Bayesian phylogenetic analysis. TCs and the closest control sequences together with two reference sequences as an out-group were selected and trees were constructed with BEAST v1.7.5 [35] using a lognormal relaxed molecular clock with the SRD06 model of nucleotide substitutions [36] and a Bayesian skyline coalescent prior. The analyses were run in triplicate for 100 million states and trees were sampled every 10000^{th} states. Maximum clade credibility trees (MCC) were summarized using TreeAnnotator after 10% of the burn-in was discarded and visualized with FigTree v.1.4 (available at http://tree.bio.ed.ac.uk). The TCs with a Bayesian posterior probability of 1 were considered robust enough and included in the analysis.

Finally, we defined TCs with TDR as any pair or cluster with more than 3 patients that included at least one patient with TDR from the Leuven ND cohort. The TCs with TDR with more than 3 individuals with similar TDR mutation profile can be suggestive of onward transmission of TDR, they are specifically indicated as TCs with TDR-OT.

Statistical Analysis

Prevalence of TDR and TDR mutations were calculated with a 95% Wilson score confidence interval (95% CI) on the basis of a binomial distribution, and their trend was calculated by logistic regression analysis. Socio-demographic, virological and clinical variables that were significantly associated with TDR or with TCs with TDR were evaluated in the Leuven cohort. Analyses were performed on patients involved in TCs from the Leuven ND cohort, and from Leuven ND cohort with the other four control datasets. Categorical data were compared using the Chi-square test, the Fisher's exact test or regression techniques as appropriate. The t-test or Mann–Whitney U test was used to compare continuous data. The statistical significance was set at $p<0.05$ two-sided. All data were analysed using the statistical R software version 2.13.1.

Bayesian Network Learning

Those factors that were found to be significantly associated with TDR or TCs with TDR in univariate analysis were included in a Bayesian network analysis. This is a probabilistic model that describes statistical conditional dependencies between multiple variables and was performed using the B-course software adapted by Deforche et al [37]. In this analysis, the arcs were scored based on the stability of the conditional dependency assessed with 100 non-parametric bootstrap replicates. The arcs with bootstrap over 75% were considered and depicted in the consensus network.

Results

General Characteristics Of The Study Population

778 of the 795 patients who were newly diagnosed with an HIV-1 infection and who received a baseline genotypic drug resistance test between January 1998 and December 2012 at University Hospitals Leuven were included in the analysis, they are referred to as the Leuven ND cohort. Two patients were excluded because their risk group was vertical transmission. For 15 patients, the baseline nucleotide sequence did not fulfill the preset quality criteria: 14 sequences did not have the gene fragments encoding PR or RT, and one sequence was excluded due to the presence of more than four stop codons and indels. The included HIV-1 patients were between 18 and 78 years old and were predominantly male (73.7%), of Belgian origin (58.4%), chronically infected (86.5%) with CDC stage 1 or 2 (67.0%) (Table 1). Patients originating from Belgium were more frequently diagnosed with a recent infection and displayed higher viral loads and CD4 counts ($p<0.001$). Of all included HIV-1 patients originated from Belgium, 66.7% reported men who have sex with men (MSM) or bisexual contacts as risk factor, whereas 22.5% reported heterosexual contacts. In contrast, HIV-1 patients originating from Sub-Saharan countries reported infection through heterosexual contacts predominantly (79.3%). HIV-1 patients from Sub-Saharan countries were more likely to be co-infected with hepatitis B than patients from Belgium (60.9% vs. 34.8%; OR: 4.49, 95% CI 1.75–12.15, $p<0.001$). There were 36 HIV-1 therapy-naive patients who did not receive a baseline drug resistance test in this period. This group included more patients of non-Belgian origin (75.0%) and with CD4 count above 500 cells/mL (42.4%, 14/33).

The demographic characteristics of the Leuven ND cohort were compared to the general HIV-1 population in Belgium, as reported by the Belgian Scientific Institute of Public Health (information until 2011) (available at www.wiv-isp.be) [1]. The Leuven ND cohort contained more men (73.5% vs. 61.0%, $p<0.0002$) and Belgians (58.4% vs. 40.6%, $p<0.0002$) and more MSM (55.9% vs. 42.5%, $p<0.0002$). National data only covered gender and country of origin from 1998 to 2011, and transmission risk from 2005 to 2011.

Subtypes

52.2% of the HIV-1 patients were infected with subtype B, followed by CRF02_AG (11.2%), subtype C (10.3%), subtype A (7.7%), CRF01_AE (6.6%), subtype F (2.8%), subtype G (2.1%) and unique recombinant forms (4.8%). Subtypes D, H, J, CRF09_cpx, CRF12_BF, CRF13_cpx, CRF14_BG, CRF18_cpx, CRF22_01A1, CRF37_cpx, and CRF45_cpx were each found in less than 1%. Of the patients with a subtype B infection, 81.0% were of Belgian origin and 71.9% were MSM. Whereas in patients with non-B infections, 33.6% and 51.6% had a Belgian or sub-Saharan origin, respectively, and 69.9% were infected through heterosexual contacts, followed by bisexual/MSM risk factor (13.2%).

Levels And Trends Of Transmitted Drug Resistance

The overall TDR prevalence was 9.6% (75/778; 95% CI 7.7–11.9). The prevalence of TDR against nucleoside RT inhibitors (NRTI) was 6.5% (51/778; 95% CI 5.0–8.5), against non-NRTI (NNRTI) was 2.2% (17/778; 95% CI 1.4–3.5), and against protease inhibitors (PI) 2.2% (17/778; 95% CI 1.4–3.5). In recently infected individuals, the prevalence of overall TDR was 16.2% (17/105; 95% CI 10.4–24.4), significantly higher than in patients with chronic or unknown duration of infection (8.6%, 58/673; 95% CI 6.7–11.0). The prevalence of TDR by drug class also varied in recently infected individuals. The prevalence of TDR against NRTI was 12.4% (13/105; 95%CI 7.4–20.0), against NNRTI 1.9% (2/105; 95%CI 0.5–6.7), and against PI 6.7% (7/105; 95%CI 3.3–13.1). Dual resistance was detected in 10 patients (1.3%): 3 displayed TDR against NRTI and NNRTI, 6 against NRTI and PI and one against NNRTI and PI. The latter patient with NNRTI and PI resistance and 4 out of 6 individuals with NRTI and PI resistance were recently infected patients (5/105; 4.8%). No triple class resistance was observed.

The majority of the 75 TDR patients displayed one single mutation (70.7%), mainly related to NRTI (58.5%) and NNRTI resistance (24.5%). The revertants at RT position 215 were the most prevalent (44%), followed by M41L (18.7%), K103N (17.3%), L210W (10.7%), K219Q (10.7%), D67N (6.7%), K219R (5.3%), G190A (4.0%), M184V (2.7%), L74V (1.3%),

Table 1. Characteristics of the Leuven ND cohort and factors associated with TDR.

Characteristics at time of sampling	Total		TDR		Univariate		Multivariate	
	n	%	n	%	OR (95% CI)	*p*	OR (95% CI)	*p*
Patients	778	100	75	100				
Male	573	73.7	67	89.3	3.25 (1.52–7.99)	<0.001*		
Pregnant women	15	1.9	1	1.3				
Age in years at enrolment, Mean (SD)	37.5 (±10.6)		39.6 (±12.4)					
<25	69	8.9	6	8.0				
25–34	287	36.9	23	30.7				
35–44	241	31.0	23	30.7				
45–54	121	15.6	12	16.0				
>55	60	7.7	11	14.7				
Country or region of origin								
Belgium	454	58.4	55	73.3	2.09 (1.20–3.77)	0.006*		
Western Europe (except Belgium)	23	3.0	4	5.3				
High-prevalent regions[†]					0.23 (0.08–0.52)	<0.001*		
Sub-Saharan Africa	198	25.4	7	9.3	0.27 (0.10–0.61)	<0.001*		
Other	24	3.1	0	0				
Other	75	9.6	9	12.0				
Unknown	4	0.5	0	0				
Risk of transmission								
MSM	333	42.8	47	62.7	2.44 (1.46–4.15)	<0.001*		
Heterosexual (high-prevalent country)	179	23.0	5	6.7				
Heterosexual (non-endemic)	121	15.6	13	17.3				
Bisexual	32	4.1	5	6.7				
IVDU	14	1.8	1	1.3				
Unknown	75	9.6	1	1.3				
Other	24	3.1	3	4.0				
Type of infection								
Chronic	673	86.5	58	77.3				
Recent	105	13.5	17	22.7	2.04 (1.06–3.75)	0.02*		
CDC stage[§]								
1 and 2	521	67.0	55	73.3				
3	246	31.6	19	25.3				
Unknown	11	1.4	1	1.3				
CD4 cell count, median (IQR) [‖]	335	(163–493)	365	(236–505)				

Table 1. Cont.

Characteristics at time of sampling	Total		TDR		Univariate		Multivariate	
	n	%	n	%	OR (95% CI)	*p*	OR (95% CI)	*p*
<200 cells/mm^3	228	29.3	16	21.3				
200–349 cells/mm^3	174	22.4	18	24.0				
350–499 cells/mm^3	181	23.3	21	28.0				
≥500	182	23.4	19	25.3				
Unknown	13	1.7	1	1.3				
HIV-RNA load, median (IQR), log copies/ml [¶]	4.76	(4.12–5.31)	4.76	(4.16–5.25)				
Co-infection								
Hepatitis B	23	3.0	4	5.3				
Negative	412	53.0	36	48.0				
Unknown	343	44.1	35	46.7				
Hepatitis C	406	52.2	1	1.3				
Negative	18	2.3	40	53.3				
Unknown	354	45.5	34	45.3				
Subtype								
B	406	52.2	59	78.7	3.77 (2.09–7.17)	<0.001*	3.04 (1.45–6.35)	0.003
Part of transmission cluster	226	29.0	32	42.6	1.95 (1.15–3.25)	0.010*		

*Fisher's test or logistic regression techniques were used. [†]High prevalent countries were defined as HIV-prevalence over 1% in adult population (UNAIDS 2012), [§] CDC stage was defined according to 2008 definitions [47], [‖]CD4 count at diagnosis (median 345, IQR: 158–496) was not statistically different from CD4 count at time of sampling. [¶]Viral load at diagnosis (median 4.76 IQR: 4.04–5.31) was not statistically different from viral load at time of sampling. Abbreviations: %: percentage, CDC: Center for Disease Control and Prevention, CI: Confidence interval, IVDU: Intravenous drug user, IQR: interquartile range, MSM: men who have sex with men, n: number, OR: odds ratio, SD: standard deviation.

doi:10.1371/journal.pone.0101738.t001

Y115F (1.3%) and Y181C (1.3%). Within PR, I54VT (10.7%) was the most frequent mutation followed by M46IL (9.3%), N88D (6.7%), V82TS (2.7%), L24I (1.3%), I54T (1.3%) and I85V (1.3%). As the inclusion of PR position 46 within the TDR mutation list has been debated due to its polymorphic nature [38], TDR was recalculated excluding this position. This resulted in an overall TDR of 8.9% (69/778; 95% CI 7.0–11.0) and a PI-TDR of 1.4% (11/778; 95% CI 0.8–2.5).

No significant time trends were found in the overall TDR prevalence, nor in transmitted NRTI and PI resistance (see Figure 1A). A parabolic trend was observed for NNRTI-TDR ($p = 0.019$) with a peak in 2008. That corresponded with a peak in occurrence of K103N ($p = 0.026$), the only mutation with a temporal trend. Surprisingly, the parabolic temporal NNRTI-TDR trend was not observed in the recently infected individuals. Instead, a stable temporal trend was observed for overall TDR and individual drug classes in this subset of patients. When the analysis was performed according to region of origin, a significant parabolic trend of NNRTI resistance was only found in patients originating from Belgium ($p = 0.039$).

Factors Associated With Transmitted Drug Resistance

Univariate analysis was performed to identify predictors of TDR (Table 1), which were male gender (odds ratio (OR) 3.25, 95% CI 1.52–7.99, $p<0.001$), Belgian origin (OR 2.09, 95% CI 1.20–3.77, $p = 0.006$), MSM transmission (OR 2.44, 95% CI 1.46–4.15, $p<0.001$), recent infection (OR 2.04, 95% CI 1.06–3.75, $p = 0.02$), being part of TCs (OR 1.95, 95% CI 1.15–3.25, $p = 0.010$) and infected with subtype B virus (OR 3.77, 95% CI 2.09–7.17, $p<0.001$). Only the latter remained a significant factor (OR 3.04, 95% CI 1.45–6.35, $p = 0.003$) in multivariate analysis. Since these types of analyses do not display possible interdependencies of the variables and subtype B was most frequently found in MSM originating in Belgium ($p<0.001$), we evaluated the interdependencies of the variables with a Bayesian network approach. TDR was directly associated only with subtype B, but this subtype was strongly associated with MSM (100% bootstrap support) and with Belgian origin (97% bootstrap support), and to a lesser extent to being part of TCs (78% bootstrap support) (See Figure 1B). To verify whether we could find an important predictor of TDR that could be used in guidelines to target a subpopulation of newly diagnosed for preferential drug resistance testing, we repeated the analysis excluding any information that results from the genotype itself. When subtype B was thus excluded from the analysis, then male gender became directly associated with TDR (64% bootstrap support), and with MSM and Belgian origin (100% bootstrap support), whereas the association between the two latter variables with TCs had lower bootstrap support (31%).

Transmission Clusters

We identified 114 TCs, 16 of which harbored 32 of the 75 TDR patients from our Leuven ND cohort. Five pairs and eight larger clusters of ≥3 individuals were found among subtype B infected patients, one cluster of 17 individuals with CRF02_AG, and one pair for each subtypes C and CRF01_AE (see Table 2). Six of these 16 TCs with TDR included only a single patient with TDR whereas six clusters were TCs with TDR-OT and included 20 individuals from the Leuven ND cohort (26.7%, 20/75). Singletons were frequently found in patients from the Leuven ND cohort involved in TCs (81.3%, 26/32). Likewise, prevalence against NRTI was the most frequent (81.3%, 26/32), followed by PI (25.0%) and NNRTI (12.5%). Thymidine analogue mutations (TAMs) were predominantly detected in TCs with TDR (81.3%, 26/32), mainly represented by the revertant at position 215 (59.4%, 19/32), followed by the mutations K219QR (18.8%), L210W (15.6%) and M41L (6.3%). Mutations for NNRTI and PI were I54V, N88D (each 15.6%), K103N (12.5%) and M46IL (9.4%).

The characteristics of the Leuven ND cohort patients involved in TCs were evaluated. Patients carrying TDR were significantly more associated with TCs compared to patients without TDR (OR: 1.95, see Table 1 and Table S1) and the association remained when considering only larger clusters by excluding pairs (OR: 2.86, 95% CI 1.59–5.01, $p = <0.001$). Similarly, when including only TCs with TDR-OT, TDR remained significantly

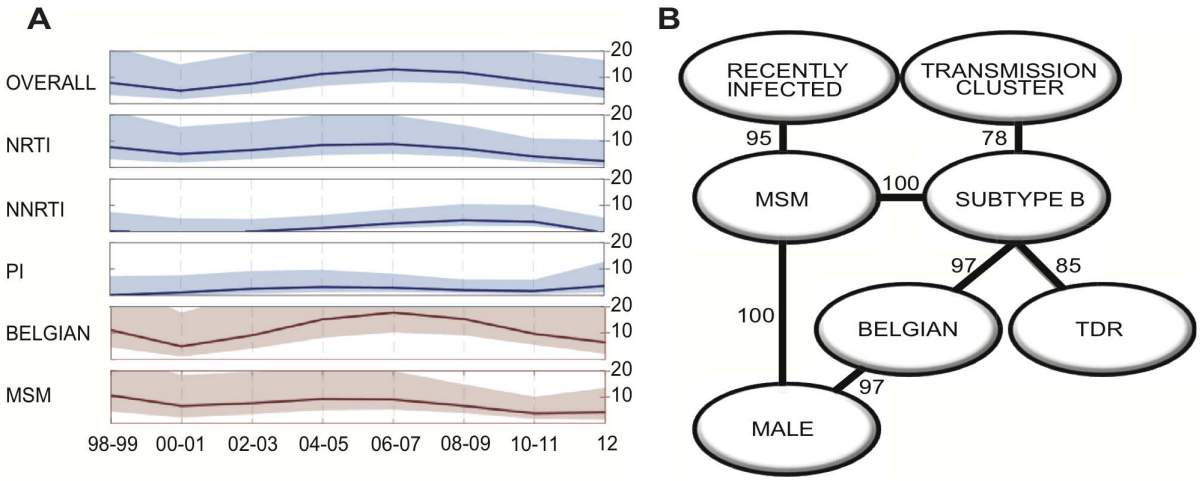


Figure 1. Temporal trends and factors associated with transmitted drug resistance (TDR). (**A**) Trends of prevalence of TDR (percentage) and the 95% confidence intervals (light shading) among newly diagnosed HIV-1 patients at ARC Leuven (Belgium) from 1998 to 2012 are shown for the overall-TDR, NRTI-TDR, NNRTI-TDR, PI-TDR in blue, MSM overall-TDR and Belgian overall-TDR in red. (**B**) The significant variables associated with TDR in the univariate analysis were included in the Bayesian network, the number next to the arcs represents the bootstrap support. Abbreviations: NRTI: nucleoside reverse transcriptase inhibitors, NNRTI: non-nucleoside reverse transcriptase inhibitors, MSM: men who have sex with men, PI: protease inhibitors.
doi:10.1371/journal.pone.0101738.g001

Table 2. Characteristics of transmission clusters containing Leuven patients with TDR.

									TDR mutations		
Cluster	**Patient ID**[†]	**Subtype**	**Risk of transmission**	**Country of sampling**	**Country of origin**	**Country of infection**	**Type of infection**	**Year of diagnosis**	**NRTI**	**NNRTI**	**PI**
1	AY165241	B	Unknown	Sweden	Unknown	Unknown	Naive - Unknown	2000*	-	-	-
	ARCL-1	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*2002*	*M41LM*	-	-
	ESAR-1	B	Heterosexual - Male	Spain	Pakistan	Spain	Naive - Recent	2003	-	-	-
	FJ481828	B	MSM	Spain	Unknown	Unknown	Naive - Unknown	2006*	-	-	-
	ESAR-2	B	MSM	Finland	Russia	Finland	Naive - Chronic	2007	-	-	-
2	EU248399	B	Unknown	Belgium	Unknown	Unknown	Naive - Unknown	2003*	-	K103N	-
	JN101707	B	Unknown	United Kingdom	Unknown	Unknown	Unknown	2004*	-	K103N	-
	ARCL-2	*B*	*Heterosexual - Male*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*	*2007*	-	*K103N*	-
	ARCL-3	*B*	*Bisexual - Male*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*2009*	-	*K103N*	-
	JF683797	B	Unknown	Cyprus	Unknown	Unknown	Naive - Unknown	2009*	-	K103N	-
3	*ARCL-4*	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive -Recent*[§]	*2001*	*L210W, T215S*	-	*I54V, N88D*
	EU248438	B	Unknown	Belgium	Unknown	Unknown	Naive - Unknown	2003*	L210W, T215S	-	I54V, N88D
	ARCL-5	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Recent*	*2005*	*L210W, T215S*	-	*I54V, N88D*
	ARCL-6	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*[§]	*2007*	*L210W, T215S*	-	*I54V, N88D*
	ARCL-7	*B*	*MSM*	*Belgium*	*Italy*	*Belgium*	*Naive - Recent*	*2010*	*L210W,T215S*	-	*I54V,N88D*
	ARCL-8	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Recent*	*2011*	*L210W, T215S*	-	*I54V, N88D*
4	DQ177230	B	Unknown	Belgium	Unknown	Unknown	Naive - Unknown	2002*	T215E	-	-
	ARCL-9	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*	*2005*	*T215E*	-	-
5	*ARCL-10*	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*[§]	*2007*	*T215D*	-	-
	ARCL-11	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*2007*	*T215D*	-	-
	ARCL-12	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*2007*	*T215D*	-	-
	ARCL-13	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*	*2007*	*T215D*	-	-
	ARCL-14	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*2010*	-	-	-
	ARCL-15	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*	*2010*	-	-	-
	ARCL-16	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Recent*	*2011*	*T215D*	-	-
	ARCL-17	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Recent*	*2011*	*T215D*	-	-
	ARCL-18	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*2012*	*T215D*	-	-
6	DQ206665	B	Unknown	Argentina	Unknown	Unknown	Naive - Unknown	2004*	-	K103N, P225H	-
	JN670104	B	Unknown	Argentina	Unknown	Unknown	Treated	2005*	M41L, M184V, T215Y	-	D30N, N88D
	ARCL-19	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*2007*	*T215S*	*K103N*	-
	ARCL-20	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Recent*	*2007*	-	*K103N*	-
7	*ARCL-21*	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Recent*	*2009*	-	-	*M46L*

Table 2. Cont.

									TDR mutations		
Cluster	**Patient ID†**	**Subtype**	**Risk of transmission**	**Country of sampling**	**Country of origin**	**Country of infection**	**Type of infection**	**Year of diagnosis**	**NRTI**	**NNRTI**	**PI**
	ARCL-22	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*2010*	*K219Q*	-	-
	JQ650683	B	MSM	Netherlands	Unknown	Unknown	Unknown	Unknown	-	-	M46L
8	*ARCL-23*	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Recent*	*2011*	*T215E*	-	-
	JQ650714	B	MSM	Netherlands	Unknown	Unknown	Unknown	Unknown	T215E	-	-
9	*ARCL-24*	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*	*2006*	-	-	-
	ARCL-25	*B*	*Transfusion - Male*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*2010*	*K219RK*	-	-
10	*ARCL-26*	*B*	*MSM*	*Belgium*	*Indonesia*	*Asia*	*Naive - Chronic*	*2001*	*T215D*	-	-
	ARCL-27	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*	*2003*	*T215D*	-	-
11	*ARCL-28*	*B*	*MSM*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*1998*	*T215C*	-	-
	ARCL-29	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*	*2002*	*T215S*	-	-
12	*ARCL-30*	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Recent*	*2003*	-	-	-
	ARCL-31	*B*	*Bisexual - Male*	*Belgium*	*Belgium*	*Belgium*	*Naive - Recent*	*2009*	*K219R*	-	-
	ARCL-32	*B*	*Bisexual - Male*	*Belgium*	*Belgium*	*Unknown*	*Naive - Chronic*	*2009*	-	-	-
	ARCL-33	*B*	*Unknown*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*	*2010*	-	-	-
13	EU817049	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	1998*	K219Q	-	-
	EU817059	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2001*	K219Q	-	-
	EU817062	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2001*	K219R	-	-
	ARCL-34	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Chronic*	*2007*	*K219Q*	-	-
	ARCL-35	*B*	*MSM*	*Belgium*	*Belgium*	*Belgium*	*Naive - Recent*	*2006*	*K219Q*	-	-
	ESAR-1	B	MSM/bisexual	Italy	Italy	Unknown	Naive - Chronic	2003	K219Q	-	-
	EU817050	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2003*	K219Q	-	-
	EU817048	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2004*	K219R	-	-
	EU817061	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2004*	K219Q	-	-
	DQ345509	B	Unknown	Argentina	Unknown	Unknown	Naive - Recent	2005*	-	-	-
	EU817058	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817060	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817065	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817056	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817051	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817068	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817047	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817055	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	EU817063	B	Unknown	United Kingdom	Unknown	Unknown	Naive - Unknown	2005*	K219QR	-	-

Table 2. Cont.

									TDR mutations		
Cluster	**Patient ID**[†]	**Subtype**	**Risk of transmission**	**Country of sampling**	**Country of origin**	**Country of infection**	**Type of infection**	**Year of diagnosis**	**NRTI**	**NNRTI**	**PI**
	FJ469703	B	Unknown	United States	Unknown	Unknown	Naive - Unknown	2005*	K219Q	-	-
	ESAR-2	B	MSM/bisexual	Germany	Germany	Unknown	Naive - Recent	2006	K219Q	-	-
	ESAR-3	B	MSM/bisexual	Cyprus	Cyprus	Unknown	Naive - Chronic	2007	-	-	-
	ESAR-4	B	MSM/bisexual	Germany	Germany	Germany	Naive - Recent	2007	-	-	-
	JF683765	B	Unknown	Cyprus	Unknown	Unknown	Naive - Unknown	2007*	-	-	-
	JF683791	B	Unknown	Cyprus	Unknown	Unknown	Naive - Unknown	2009*	-	-	-
	JF683808	B	Unknown	Cyprus	Unknown	Unknown	Naive - Unknown	2009*	-	-	-
14	***ARCL-36***	***C***	***Heterosexual - Female***	***Belgium***	***Belgium***	***Unknown***	***Naive - Chronic***	***2012***	***-***	***-***	***M46LM***
	ARCL-37	***C***	***Heterosexual - Male***	***Belgium***	***Ethiopia***	***Unknown***	***Naive - Chronic***	***2012***	***-***	***-***	***-***
15	***ARCL-38***	***CRF01_AE***	***Heterosexual - Female***	***Belgium***	***Thailand***	***Thailand***	***Naive - Chronic***	***2005***	***-***	***-***	***-***
	ARCL-39	***CRF01_AE***	***Heterosexual - Male***	***Belgium***	***Belgium***	***Belgium***	***Naive - Chronic***	***2005***	***-***	***-***	***M46IM***
16	JX290261	CRF02_AG	Unknown	Russia	Unknown	Unknown	Unknown	2001 or 2002[‡]	-	-	-
	AY829204	CRF02_AG	IVDU - Male	Uzbekistan	Unknown	Unknown	Unknown	2002*	-	-	-
	AY829207	CRF02_AG	IVDU - Male	Uzbekistan	Unknown	Unknown	Unknown	2002*	-	-	-
	AY829214	CRF02_AG	IVDU - Male	Uzbekistan	Unknown	Unknown	Unknown	2002*	-	-	-
	HQ449394	CRF02_AG	Unknown - Female	Russia	Unknown	Unknown	Unknown	2005[‡]	-	-	-
	DQ465230	CRF02_AG	MTCT	USA	Unknown	Unknown	Unknown	2006*	-	-	-
	GQ290726	CRF02_AG	Unknown	South Korea	Unknown	Unknown	Naive - Unknown	2008*	-	-	-
	GQ290743	CRF02_AG	Unknown	South Korea	Unknown	Unknown	Naive - Unknown	2008*	-	-	-
	HQ412530	CRF02_AG	Unknown - Male	Russia	Unknown	Unknown	Unknown	2008*	L74I, M184V, K219E	L100I, K101E, Y181C, G190S	-
	HQ115069	CRF02_AG	Unknown - Male	Ukraine	Unknown	Unknown	Unknown	2009[‡]	-	-	-
	JX500703	CRF02_AG	Unknown	Russia	Unknown	Unknown	Unknown	2010*	-	-	-
	ARCL-40	***CRF02_AG***	***Heterosexual - Female***	***Belgium***	***Kazakhstan***	***Unknown***	***Naive - Chronic***	***2011***	***K219RK***	***-***	***-***
	JX500697	CRF02_AG	Unknown	Russia	Unknown	Unknown	Unknown	2011*	-	-	-
	JX500706	CRF02_AG	Unknown	Russia	Unknown	Unknown	Unknown	2011*	-	-	-
	KC509858	CRF02_AG	Unknown - Female	Russia	Unknown	Unknown	Unknown	2012*	-	-	-
	KC120872	CRF02_AG	Unknown	South Korea	Unknown	Unknown	Unknown	Unknown	-	-	-
	KC120881	CRF02_AG	Unknown	South Korea	Unknown	Unknown	Unknown	Unknown	-	-	-

Abbreviations: ARCL: AIDS Reference Center Leuven, CRF: Circulating recombinant form, ESAR: European Society for Translational Antiviral Research, IVDU: intravenous drug user, NRTI: nucleoside reverse transcriptase inhibitors, NNRTI: non-nucleoside reverse transcriptase inhibitors, MSM: men who have sex with men, MTCT: mother to child transmission, PI: protease inhibitors,[†]Patient ID includes patients of the Leuven cohort (bold and italics), ESAR controls and accession numbers of NCBI database *Control sequences have available year of sampling. [‡]Control sequences with year of diagnosis available. [§]Sequences were also included when the patient was on antiretroviral treatment.

doi:10.1371/journal.pone.0101738.t002

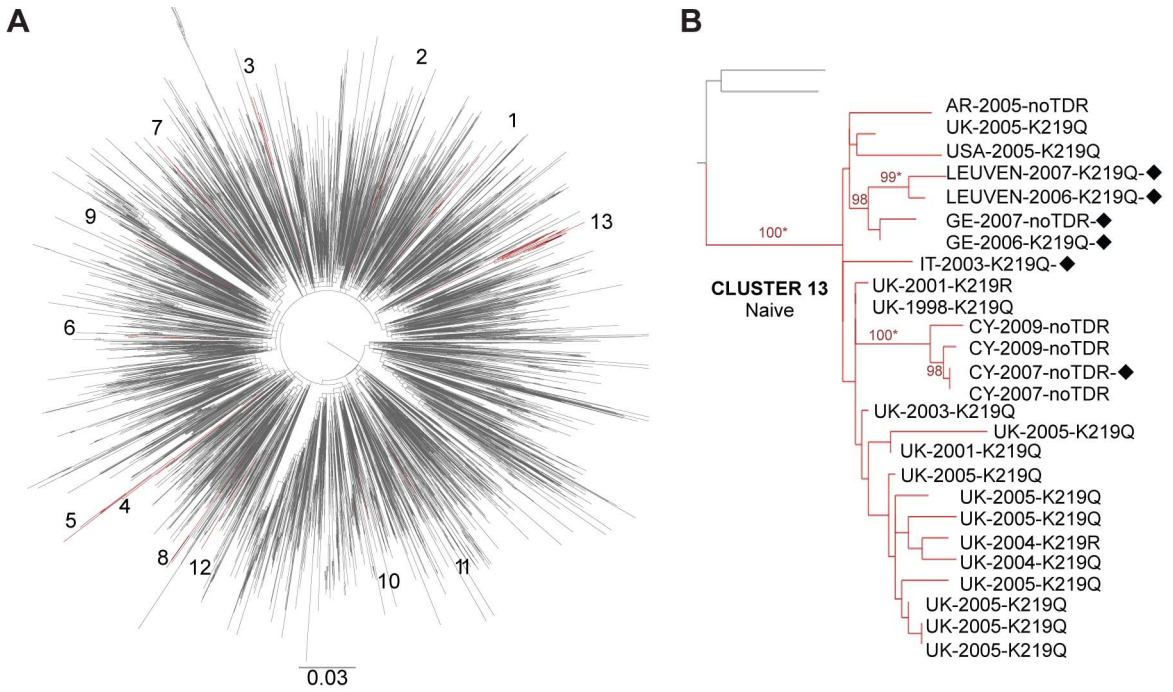


Figure 2. Examples of subtype B transmission clusters (TCs) with TDR: A maximum likelihood (ML) tree per subtype was constructed, and TCs were confirmed by Bayesian Phylogenetic analyses. (**A**) The ML tree for subtype B (Leuven ND cohort and control sequences) with the TCs colored in dark red. (**B**) The largest TC of subtype B: composed of therapy-naive patients, several nationalities and mutations at RT position 219; bootstrap values above 98% are shown. Abbreviations: AR: Argentina, CY: Cyprus, GE: Germany, IT: Italy, UK: United Kingdom, USA: United States of America, black diamond: men who have sex with men, asterisk: posterior distribution equal to 1 in the Bayesian phylogenetic analysis.
doi:10.1371/journal.pone.0101738.g002

associated with TCs (OR: 2.44, 95% CI 1.32–4.36, p = 0.002, see Table S1). 32 out of 75 TDR patients were involved in TCs (42.6%, 95% CI 32.1–53.9), while of the 703 patients without TDR, only 194 were found in TCs (27.6%, 95% CI 24.4–31.0). As expected, Leuven ND cohort patients with TDR and involved in TCs were significantly more of Belgian origin (90.6% versus 65.5%; OR: 4.84, 95% CI 1.41–25.78, p = 0.005), infected with subtype B (90.6% versus 63.9%; OR: 5.42, 95% CI 1.59–28.85, p = 0.001) and characterized with MSM risk factor (78.1% versus 52.6%; OR: 2.71, 95% CI 1.07–7.84, p = 0.03) than their counterparts without TDR (see Table S1). When we focussed on predictors for TCs with TDR by including the data of controls, subtype B remained significantly associated with TCs with TDR (77.4% versus 65.1%, OR: 1.83 95% CI 1.03–3.35, p = 0.036). However, TCs of patients with TDR were larger than TCs without TDR (median 3.5 vs. 2 patients per TC, OR: 1.43 95% CI 1.08–1.91, p = 0.001). Belgium as sampling country or as country of origin and heterosexual contact were then more frequent in the group of TCs that included solely patients without evidence of TDR. Multivariate analysis did not show any significant factor associated with patients in TCs with TDR versus other TCs. Finally, we also performed separate analyses on TCs with TDR-OT (see Table S1). The same variables remained significantly associated with TCs in the univariate analysis, with the exception of recent infection that became significant. The median of the TCs with TDR-OT was larger than TCs without TDR (5.5 versus 2 patients per TCs, OR: 2.01 (1.29–3.11). Likewise, none of the variables were significant in the multivariate analysis.

The characteristics of the Leuven TDR patients involved in TCs are shown in Table 2. Five pairs included mainly naive MSM originating from Belgium, infected with a subtype B strain carrying one of the 215 revertants, whereas two pairs with subtype C and CRF01_AE strains displaying mutations at PR position 46 included heterosexual Belgians with a foreign partner. Two large subtype B TCs with TDR included 9 or more patients. Cluster number 5 was composed of nine therapy-naive individuals originating from Belgium with MSM as a risk factor and diagnosed between 2007 and 2012. Seven of them displayed a revertant at RT position 215 while two strains had no TDR. Cluster number 13 (Figure 2B) included 26 naive patients mainly from United Kingdom and countries of western and southern Europe, western Asia and America. The main mode of transmission was MSM infected with viruses carrying mutations at position 219, except for six individuals who mainly originated from Cyprus and did not display any mutations. The remaining subtype B TCs were composed of three to six individuals. All had at least one individual originating from Belgium and one from another country. For instance, cluster number 1 included one patient originating from Belgium and two patients from Russia and Pakistan but infected in other countries like Finland and Spain. One peculiar subtype B cluster with extensive NRTI and PI resistance (cluster number 3) contained 6 therapy-naive MSM with TDR all with a Belgian connection, either infected in Belgium or originating from Belgium. On the other hand, the largest cluster of non-B subtypes was composed of 17 individuals infected with CRF02_AG and sampled in different countries from Central-eastern Asia and Eastern Europe and characterized by different risks factors including heterosexual orientation, intravenous drug

user (IVDU) or vertical transmission. The resistance pattern was also heterogeneous in this cluster. The majority of patients did not display TDR mutations, whereas one naive patient sampled in Belgium and originating from Kazakhstan displayed a mixture of K219RK, and one patient who was tested in Russia was probably treated and displayed high-level resistance against NRTI and NNRTI.

Since the definition of clustering is still a matter of debate, additional analyses were performed to assess the impact of the genetic distance on the identification of TCs with values of 0.015, 0.03 and 0.045. As a result, the number of TCs with TDR decreased to 9, 13 and 17, respectively, including 22.7%, 29.3% and 36.0% of the 75 patients with TDR. The number of TCs that included TDR patients was larger with the genetic distance 0.045 because the cluster with 9 individuals identified as number 5 in Table 2 was split in one pair (ARCL-16-TDR and ARCL-18-TDR) and one cluster with the remaining patients. With the stringent criteria of 0.015, TDR was still associated with TCs (OR: 2.56 05% CI 1.32–4.76, $p = 0.003$). Subtype B infection remained as the factor associated with TCs with TDR ($p = 0.002$), whereas sociodemographic factors such as Belgian nationality and MSM contact were not significantly linked.

Potential Impact Of TDR On First-Line Regimen

For each patient with TDR, we analysed the genotypic susceptibility score (GSS) of the three regimens most frequently prescribed during the year the patient was diagnosed, and for each patient who in the meantime had started treatment we also estimated the GSS of the first-line therapy received. Rega algorithm 9.1.0 was used for these analyses, which has a recommended GSS for patients with TDR. This recommendation was introduced in March 2007 and suggests a regimen with GSS of ≥ 3.5 when TDR is detected, thereby suggesting a triple therapy with a fully active boosted PI which receives a score 1.5, but not with an NNRTI which receives a score of 1 when fully active.

Amongst the 636 patients who started ART, the GSS of the actual prescribed first-line regimen was ≥ 3.5 in 261 patients (41.0%), 3 in 357 patients (56.1%), and <3 in 18 patients (2.8%). The latter group included five patients without TDR, of which three patients were treated with bi-therapy in 1998, one with bi-therapy including a PI in 2002 and one with a mono PI regimen in 2010. Although the demographic characteristics were similar between the patients with a GSS ≥ 3 and <3, the latter group more often started ART before 2002 (8/18, $p<0.001$).

In the group of patients with viruses carrying TDR, 60 started therapy before the end of 2012. The GSS was ≥ 3.5 in 34 patients (56.7%), equal to 3 in 13 patients (21.7%), and <3 in 13 patients (21.7%). The latter group displayed resistance only to NRTIs (6/13), only to PI (2/13) or dual resistance (5/13).

Sustained undetectable viral load during the first-line therapy was reached in 83.3% (50/60) of the TDR patients, whereas 3.3% of the patients (2/60) had sustained low level viremia without evidence of virological rebound above 500 copies/ml. Four patients (6.7%) had early changes in ART due to toxicity and 1 patient (1.7%) died shortly after therapy initiation. Only 5.1% of the patients (3/60) displayed virological failure and the emergence of major NNRTI resistance-related mutations. Two of these patients displayed evidence of only NRTI TDR mutations at baseline and received a NRTI+NNRTI regimen with a GSS equal to 3. The third patient started a NRTI+NNRTI therapy with a GSS equal to 1 in 2004, two days after the first contact date and two weeks before the drug resistance results were available. In this patient, the therapy was quickly changed after receiving the baseline drug resistance report indicating NRTI and NNRTI TDR mutations and observing no virological response. Subsequent drug resistance testing on a later sample revealed further accumulation of NNRTI resistance.

To appreciate the value of baseline drug resistance testing, the GSS was calculated for the three most frequently prescribed first-line regimens in patients displaying TDR in our cohort per year (see Table S2). Theoretically, 49.3% of TDR patients (37/75) were likely to receive a potential suboptimal regimen ($GSS<3$) if baseline drug resistance testing had not been available to the treating physician. However, the frequency of these patients is decreasing over time ($p = 0.002$). Up until 2003, the most common first-line regimens included a thymidine analogue with either a NNRTI- or unboosted PI. Thereafter, tenofovir or abacavir were more commonly used as supporting NRTI, but a $GSS<3$ was still mainly observed for NRTI+NNRTI regimens. A low GSS under a boosted PI based regimen would only have accounted for approximately 15% of the TDR cases per year between 2005 and 2011.

Discussion

The TDR surveillance in the 778 included patients who were newly diagnosed with HIV-1 at our clinic in Leuven showed a stable overall prevalence of 9.6% between 1998 and 2012. This result is in line with the 9.5% of the latest national survey that included 285 patients who were newly diagnosed in Belgium between 2003 and 2006 [11]. It was also consistent with the stable overall TDR levels of 9.7% in Spain and of 9.0% in France, results from national surveys with a similar design and time frame as our study [34,39], and with the overall trend of 8.9% in Europe between 2002 and 2007 [26,40]. However, the overall TDR in our local epidemic was higher than the 5.6% between 2003 and 2010 in Sweden [5] and the 6.5% between 2001 and 2009 in Ghent (Belgium) [4]. These regional differences highlight the importance of studying local epidemics and suggest that TDR prevalence may vary within a single country. Indeed, our findings may not be generalizable to the HIV-1 epidemic in Belgium because our cohort has a higher prevalence of MSM and of individuals originating from Belgium. It should be pointed out, however, that the demographic characteristics data of our study population was more complete than the national database for which nationality and mode of transmission were not available in approximately 25% of the cases. Although our study could have overestimated the level of TDR, due to patients who were unwilling to disclose their ART status, measures were taken to decrease the number of misclassifications. Patient records were exhaustively revised by clinicians and virologists, and individuals with evidence of drug resistance and viral load profiles suggestive of treatment were not considered drug naive.

While subtype B, being MSM, male gender, originating from Belgium, recently infected, and involvement in TCs were all significant predictors for TDR, only subtype B remained significantly associated in the multivariate analysis. Bayesian analyses however, showed the dependency of this factor on being of Belgian origin and MSM. Similarly when subtype B was excluded, male gender became directly associated with TDR with low support but significantly dependent on Belgian origin and MSM. These findings are in agreement with previous reports from other Belgian, European and American studies [4,5,11,39,41]. In a recent study, the association of subtype with country of sampling, risk group and gender has been interpreted as evidence for highly compartmentalized epidemics in Europe [42]. Therefore, the early introduction of subtype B in European MSM and their broad access to HIV care and ART for decades might explain the single

direct association of subtype B with TDR in many resource-rich settings. Nevertheless, recent studies revealed an increasing prevalence of TDR among Sub-Saharan African migrants residing in Spain and Sweden, potentially linked to the increasing drug resistance levels in Africa [5,43]. However in our cohort, Sub-Saharan African patients were still associated with less TDR and we did not observe a time trend in those patients (data not shown).

Fluctuations in TDR levels were observed over the entire study period, but the only significant trend was a parabolic trend detected for transmitted NNRTI resistance. The overall NNRTI TDR prevalence was 2.2%, with a maximum of 6.5% in 2008-2009. This parabolic trend, mainly linked to the detection of K103N and to a Belgian origin, was not observed among recently infected patients. Although a parabolic trend with a peak in 2004 was also described in the SPREAD study that included data up until 2005 [26], the same surveillance up to 2007 showed a linear increase over time [40] potentially associated with the frequent use of NNRTI in first-line regimens as the authors suggested. Similarly, the local change of prescribing practices to more potent regimens in later years, use of drug resistance testing and the longer time period analyzed in this study could explain the parabolic trend. In the total cohort, TDR associated with NRTI and PI resistance fluctuated around 6.5% and 2.2%, respectively. Among recently infected patients, NRTI- and PI-TDR levels increased to 12.4% and 6.7% respectively. This increase was not observed for NNRTI resistance, presumably due to the lower impact of NNRTI mutations on viral fitness with consequently a lower likelihood of reversion to wild-type and of a TDR underestimation by population-based Sanger sequencing in chronically infected patients.

Singletons were predominantly detected in our cohort, with TAMs as the most commonly observed. Although the use of zidovudine has decreased over the last few years, we did not find any time trend for TAMs. The peak in NNRTI TDR in 2008 was not related to clustered transmission of TDR or migration from other countries as has been suggested in other settings [12,43]. However, it might have been linked with the enhanced use of NNRTI-containing combinations in the years before. From 2009 onwards, the most commonly prescribed regimen was the potent combination tenofovir+emtricitabine+efavirenz. Up until 2012, no other available NNRTI- or PI/ritonavir-based regimen had proven superior to this regimen with respect to virologic responses.

If resistance testing had not been performed and patients with TDR would have received one of the preferred first-line regimens at that time, approximately half of them would likely have received a regimen in which the virus had lost susceptibility to at least one of the prescribed drugs. Irrespective of the detected TDR by population-based Sanger sequencing, all of them would have had a higher risk of virological failure, as NNRTI-based regimes were commonly prescribed from 2002 onwards [8]. However, 83% of the patients with TDR and who started ART achieved undetectable viral load thanks to the prescription of potent regimens enabled by the availability of drug resistance results. Only 3 patients with baseline resistance to NRTI had virological failure with development of NNRTI resistance after the initiation of a NRTI+NNRTI regimen.

In this cohort, 42.6% of the TDR patients were involved in TCs, which included nine clusters and seven pairs. Because we were interested in observing TCs over a period of 15 years, a genetic distance of 0.06 substitutions per site and a bootstrap support of $\geq$98% were used to define TCs. These TCs were also confirmed using Bayesian phylogenetic techniques, indicating that the obtained results were robust [4,7,34]. Although, the comparison between studies of transmission networks is difficult due to the differences in sampling, phylogenetic techniques and the lack of a standardized TC definition, our results were in line with a study carried out in Ghent (Belgium) in which 18 out of 33 (55%) TDR patients were involved in pairs or larger clusters [4]. When a more stringent criterion of 0.015 genetic distance was used, the percentage of TDR patients involved in TCs decreased to approximately 23%. In general, some conclusions can be drawn from the TCs analyses. First, none of the factors were significantly associated with TCs with TDR in the multivariate model, although a dependency between subtype B and clustering was found in the Bayesian network and TDR was significantly associated with TCs in the univariate analysis. Therefore we were unable to identify a non-sequence based predictor of being in TCs with TDR, even though the odds were higher for patients who were MSM and originating from Belgium. This result is similar to other studies performed in Europe [4,34]. Second, TAMs were more frequently found in TCs and this was also observed in the Ghent cohort and other settings [3,4,7]. Third, TCs with TDR involved mainly therapy-naive individuals, chronically or recently infected, rather than ART-experienced patients, which could suggest that drug naive people, potentially unaware of their HIV seropositive status, are the main source of TDR instead of patients failing ART [3,6,7]. Fourth, 7 out of 16 TCs with TDR involved patients of different nationalities. Although, we were not able to retrace the country of infection in many instances, this may imply that migration plays an important role in the local spread of subtype B as previously described [44], but also of TDR. Fifth, spread of non-B subtypes in the local epidemic was still limited and was related with heterosexuals as has been described in other epidemics [6]. They were also not prone to spread TDR. They often involved Belgians and other nationalities that could imply a limited intermixing of the HIV-1 epidemic between locals and immigrants.

Although we used all the sequences available in public databases and from a collaborative European dataset, for 7 TCs with TDR we did not find evidence that patients other than the ones followed at our clinic were involved in transmission networks. Similarly, 6 TCs with TDR patients from the Leuven ND cohort included only a single patient with TDR, and for these patients, no evidence of onward transmission of TDR is available. When 3 or more patients in a TC had the same TDR mutation profile, we indicated the cluster as TDR-OT, since this is suggestive of onward transmission of TDR, although we cannot exclude that all these TDR patients received their resistance from a treated patient. The association between TDR and TCs remained, also for those TCs with TDR-OT. Since 27% of the patients from the Leuven ND cohort were involved in those TCs, this could imply an important role of local transmission on the spread of TDR. Nevertheless, we cannot exclude the possibility that the networks might include other intermediary individuals who were not sampled or unaware of their seropositive status, known limitations of phylogenetic analyses.

Singletons and the TAM 215 were predominant in TCs, but the clinical impact on the current first-line therapies remains limited. However, two TCs that involved MSM individuals originating from Belgium with viruses carrying the K103N were detected. The latest diagnosis date was 2009 in these TCs, suggesting that in our local epidemic this mutation was not involved in a recent spread in contrast to a reported outbreak in Namur (Belgium) [12]. Continuous monitoring of the spread of this mutation is required to establish the impact on current practices. On the other hand, two large clusters were detected with the TAM 219 and they involved different nationalities from Europe, Asia and America. One of these TCs contained one patient living in Belgium but

originating from Kazakhstan, and control sequences from Uzbekistan that were part of an outbreak of CRF02_AG among IVDU [45]. Our analyses revealed the involvement of other countries and risk groups and the absence of K219R in many of the clustered sequences. The other large cluster included MSM individuals infected with subtype B and control sequences mainly from United Kingdom [46] and other countries in Europe. The majority of strains in this cluster displayed K219Q with only a few strains displaying K219R or no TDR.

In summary, this study showed a stable trend of almost 10% overall TDR between 1998 and 2012 in our Leuven (Belgium) cohort. TDR associated with NNRTI resistance displayed a parabolic trend that overlapped with an up-and-down NNRTI TDR trend in Belgians and with the trend of K103N. Our cohort was mainly composed of chronically infected patients and around 43% of the patients with TDR were involved in transmission networks, suggesting public health policies that target early diagnoses of recently infected patients are needed. Although the main factor related with TDR was subtype B, this variable was dependent on Belgian nationality and MSM mode of transmission. While these variables were also associated with being in TCs with TDR, we were unable to significantly identify a population that could be targeted for future TDR prevention policies. More local, national and international surveillance studies are needed to confirm the significance and durability of our observations, as changes in TDR levels and patterns are not straightforward to predict due to potential changes in prevention, testing and treatment strategies and changes in other potentially important drivers, such as e.g. behaviour and migration.

Supporting Information

Table S1 Characteristics of patients from the Leuven ND cohort and from patients involved in transmission clusters. Transmission clusters with likely onward transmission included clusters number 2, 3, 5, 6, 7, 13 in Table 2. Multivariate analysis was not significant in any of the analyses. Abbreviations: CI confidence intervals, IVDU intravenous drug user, MSM men who have sex with men, n sample, OR odds ratio, % percentage (DOC)

Table S2 Impact of transmitted drug resistance (TDR) on clinical care: The genotypic susceptibility score (GSS) of each sequence with TDR was calculated for the antiretroviral regimens most frequently prescribed in the year of diagnosis (top three). For instance, the GSS was less than 3 for each of the most frequently prescribed regimens in the only sequence with TDR sampled in 1998. According to the Rega algorithm, a GSS of at least 3.5 is advised for the first-line therapy in a patient carrying a virus with TDR. Abbreviations: ART antiretroviral therapy, 3TC lamivudine, ABC abacavir, ATV atazanavir, ATV/r ritonavir-boosted atazanavir, AZT zidovudine, D4T stavudine, DDI didanosine, DRV/r ritonavir-boosted darunavir, EFV efavirenz, FPV/r ritonavir-boosted fosamprenavir, FTC emtricitabine, IDV indinavir, LPV/r ritonavir-boosted lopinavir, NFV nelfinavir, NVP nevirapine, TDF tenofovir disoproxil fumarate.
(DOCX)

Acknowledgments

We are grateful to the patients for their participation. We would like to thank Nuno Rodriguez Faria, Bram Vrancken, Ana Abecasis and Philippe Lemey for methodological advice. We would like to acknowledge the ESAR group for the provision of the European SPREAD dataset. Part of this work was presented at the 14th European AIDS Conference/European AIDS Clinical Society, Brussels, Belgium, 2013 and the 12th European Workshop on HIV & Hepatitis, Barcelona, Spain, 2014.

Author Contributions

Conceived and designed the experiments: ACPP AMV KVL. Performed the experiments: YS LV. Analyzed the data: ACPP GL NST RK JV AMV KVL. Wrote the paper: ACPP YS LV FF GL NST RK ID PDM CK LGK CN KL AW MS RP CB JA CB AGL EVW MVR JV AMV KVL. Contributed data/reagents/materials/analysis tools: FF ID PDM CK LGK CN KL AW MS RP CB JA CB EVW MVR JV AMV KVL.

References

1. ISP-WIV (2012) Epidemiologie van AIDS en HIV infectie in België. Epidemiologie du SIDA et de l'infection à VIH en Belgique. In: Direction opérationnelle Santé publique et Surveillance, Service Maladies infectieuses au sein de la Population généra and Programme de surveillance VIH/SIDA/IST, editors.
2. European Centre for Disease Prevention and Control/WHO Regional Office for Europe (2012) HIV/AIDS surveillance in Europe 2011. Stockholm: European Centre for Disease Prevention and Control.
3. Brenner BG, Roger M, Routy JP, Moisi D, Ntemgwa M, et al. (2007) High rates of forward transmission events after acute/early HIV-1 infection. The Journal of infectious diseases 195: 951–959.
4. Chalmet K, Staelens D, Blot S, Dinakis S, Pelgrom J, et al. (2010) Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections. BMC infectious diseases 10: 262.
5. Karlsson A, Bjorkman P, Bratt G, Ekvall H, Gisslen M, et al. (2012) Low prevalence of transmitted drug resistance in patients newly diagnosed with HIV-1 infection in Sweden 2003-2010. PLoS One 7: e33484.
6. Yerly S, von Wyl V, Ledergerber B, Boni J, Schupbach J, et al. (2007) Transmission of HIV-1 drug resistance in Switzerland: a 10-year molecular epidemiology survey. AIDS (London, England) 21: 2223–2229.
7. Callegaro A, Svicher V, Alteri C, Lo Presti A, Valenti D, et al. (2011) Epidemiological network analysis in HIV-1 B infected patients diagnosed in Italy between 2000 and 2008. Infect Genet Evol 11: 624–632.
8. Wittkop L, Gunthard HF, de Wolf F, Dunn D, Cozzi-Lepri A, et al. (2011) Effect of transmitted drug resistance on virological and immunological response to initial combination antiretroviral therapy for HIV (EuroCoord-CHAIN joint project): a European multicohort study. The Lancet infectious diseases 11: 363–371.
9. Van Vaerenbergh K, Debaisieux L, De Cabooter N, Declercq C, Desmet K, et al. (2001) Prevalence of genotypic resistance among antiretroviral drug-naive HIV-1-infected patients in Belgium. Antiviral therapy 6: 63–70.
10. Derdelinckx I, Van Laethem K, Maes B, Schrooten Y, De Wit S, et al. (2004) Current levels of drug resistance among therapy-naive HIV-infected patients have significant impact on treatment response. Journal of acquired immune deficiency syndromes (1999) 37: 1664–1666.
11. Vercauteren J, Derdelinckx I, Sasse A, Bogaert M, Ceunen H, et al. (2008) Prevalence and epidemiology of HIV type 1 drug resistance among newly diagnosed therapy-naive patients in Belgium from 2003 to 2006. AIDS research and human retroviruses 24: 355–362.
12. Ruelle J, Ingels MG, Jnaoui K, Ausselet N, Vincent A, et al. (2013) Transmission Network of an HIV Type 1 Strain with K103N in Young Belgian Patients from Different Risk Groups. AIDS research and human retroviruses 29: 1306–1309.
13. Castro E, Khonkarly M, Ciuffreda D, Burgisser P, Cavassini M, et al. (2010) HIV-1 drug resistance transmission networks in southwest Switzerland. AIDS research and human retroviruses 26: 1233–1238.
14. Skoura L, Metallidis S, Buckton AJ, Mbisa JL, Pilalas D, et al. (2011) Molecular and epidemiological characterization of HIV-1 infection networks involving transmitted drug resistance mutations in Northern Greece. J Antimicrob Chemother 66: 2831–2837.
15. Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, et al. (2003) Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS (London, England) 17: 1871–1879.
16. UNAIDS-WHO (2011) When and how to use assays for recent infection to estimate HIV incidence at a population level. In: UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance, editor editors. Geneva, Switzerland.
17. Maes B, Schrooten Y, Snoeck J, Derdelinckx I, Van Ranst M, et al. (2004) Performance of ViroSeq HIV-1 Genotyping System in routine practice at a Belgian clinical laboratory. Journal of virological methods 119: 45–49.
18. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, et al. (2009) Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS One 4: e4724.

19. Vercauteren J, Beheydt G, Prosperi M, Libin P, Imbrechts S, et al. (2013) Clinical evaluation of Rega 8: an updated genotypic interpretation system that significantly predicts HIV-therapy response. PLoS One 8: e61436.
20. de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, et al. (2005) An automated genotyping system for analysis of HIV-1 and other microbial sequences. Bioinformatics (Oxford, England) 21: 3797–3800.
21. Struck D, Perez-Bercoff D, Devaux C and Schmit JC (2010) COMET: A Novel approach to HIV-1 subtype prediction. 8th European HIV Drug Resistance Workshop. Sorrento, Italy.
22. Pineda-Pena AC, Faria NR, Imbrechts S, Libin P, Abecasis AB, et al. (2013) Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: Performance evaluation of the new REGA version 3 and seven other tools. Infect Genet Evol.
23. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics (Oxford, England) 22: 2688–2690.
24. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, et al. (1999) Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. Journal of virology 73: 152–160.
25. SPREAD (2008) Transmission of drug-resistant HIV-1 in Europe remains limited to single classes. AIDS (London, England) 22: 625–635.
26. Vercauteren J, Wensing AM, van de Vijver DA, Albert J, Balotta C, et al. (2009) Transmission of drug-resistant HIV-1 is stabilizing in Europe. The Journal of infectious diseases 200: 1503–1508.
27. Rhee SY, Kantor R, Katzenstein DA, Camacho R, Morris L, et al. (2006) HIV-1 pol mutation frequency by subtype and treatment experience: extension of the HIVseq program to seven non-B subtypes. AIDS (London, England) 20: 643–651.
28. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32: 1792–1797.
29. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and evolution 28: 2731–2739.
30. Ragonnet-Cronin M, Hodcroft E, Hue S, Fearnhill E, Delpech V, et al. (2013) Automated analysis of phylogenetic clusters. BMC bioinformatics 14: 317.
31. Hue S, Clewley JP, Cane PA and Pillay D (2004) HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. AIDS (London, England) 18: 719–728.
32. Frentz D, Wensing AM, Albert J, Paraskevis D, Abecasis AB, et al. (2013) Limited cross-border infections in patients newly diagnosed with HIV in Europe. Retrovirology 10: 36.
33. Kaye M, Chibo D and Birch C (2008) Phylogenetic investigation of transmission pathways of drug-resistant HIV-1 utilizing pol sequences derived from resistance genotyping. Journal of acquired immune deficiency syndromes (1999) 49: 9–16.
34. Yebra G, Holguin A, Pillay D and Hue S (2013) Phylogenetic and demographic characterization of HIV-1 transmission in Madrid, Spain. Infect Genet Evol 14: 232–239.
35. Drummond AJ and Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC evolutionary biology 7: 214.
36. Shapiro B, Rambaut A and Drummond AJ (2006) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Molecular biology and evolution 23: 7–9.
37. Deforche K, Silander T, Camacho R, Grossman Z, Soares MA, et al. (2006) Analysis of HIV-1 pol sequences using Bayesian Networks: implications for drug resistance. Bioinformatics (Oxford, England) 22: 2975–2979.
38. Frentz D, van de Vijver DA, Boucher CA and Albert J (2011) Estimates of HIV transmitted drug resistance can be inflated due to natural sequence polymorphisms. Journal of acquired immune deficiency syndromes (1999) 58: e135–137.
39. Descamps D, Assoumou L, Chaix ML, Chaillon A, Pakianather S, et al. (2013) National sentinel surveillance of transmitted drug resistance in antiretroviral-naive chronically HIV-infected patients in France over a decade: 2001-2011. J Antimicrob Chemother in press.
40. Frentz D, Van de Vijver DA, Abecasis AB, Albert J, Hamouda O, et al. (2012) The Epidemiology of HIV-1 Transmitted Drug Resistance: Increase in transmitted resistance to nonnucleoside reverse transcriptase inhibitors among newly diagnosed HIV-1 infections in Europe. Department of Virology. Rotterdam, the Netherlands: Erasmus MC. pp. 77–94.
41. Readhead AC, Gordon DE, Wang Z, Anderson BJ, Brousseau KS, et al. (2012) Transmitted antiretroviral drug resistance in New York State, 2006-2008: results from a new surveillance system. PLoS One 7: e40533.
42. Abecasis AB, Wensing AM, Paraskevis D, Vercauteren J, Theys K, et al. (2013) HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. Retrovirology 10: 7.
43. Yebra G, de Mulder M, Perez-Elias MJ, Perez-Molina JA, Galan JC, et al. (2011) Increase of transmitted drug resistance among HIV-infected sub-Saharan Africans residing in Spain in contrast to the native population. PLoS One 6: e26757.
44. Paraskevis D, Pybus O, Magiorkinis G, Hatzakis A, Wensing AM, et al. (2009) Tracing the HIV-1 subtype B mobility in Europe: a phylogeographic approach. Retrovirology 6: 49.
45. Carr JK, Nadai Y, Eyzaguirre L, Saad MD, Khakimov MM, et al. (2005) Outbreak of a West African recombinant of HIV-1 in Tashkent, Uzbekistan. Journal of acquired immune deficiency syndromes (1999) 39: 570–575.
46. Leigh Brown AJ, Lycett SJ, Weinert L, Hughes GJ, Fearnhill E, et al. (2011) Transmission network parameters estimated from HIV sequences for a nationwide epidemic. The Journal of infectious diseases 204: 1463–1469.
47. Centers for Disease Control and Prevention (2008) Revised Surveillance Case Definitions for HIV Infection Among Adults, Adolescents, and Children Aged < 18 Months and for HIV Infection and AIDS Among Children Aged 18 Months to <13 Years — United States, 2008. In: MMWR, editor editors. Atlanta. pp. 12.

CURRICULUM VITAE

PERSONAL DATA

Name: Andrea-Clemencia Pineda-Peña

Place of birth: Bogotá, Colombia

E-mail: andreapinedap@gmail.com

PROFESSIONAL BACKGROUND

Research experience	**PhD researcher.** Universidad del Rosario, Colombia and KU Leuven, Belgium	08-2009 – 08-2014 10-2010 – 08-2014
	Research Coordinator. Asistencia Científica de Alta Complejidad S.A.S., Colombia. Multidisciplinary program for assistance of HIV/ aids infected patients**.**	02-2009 – 07-2010
	Member of Clinical and Molecular Aspects of Infectious Diseases Group. Universidad del Rosario, Colombia.	08-2009 – 08-2014 01-2008 – 07-2009
	Member of Medical Basic Sciences Research Group. Universidad del Rosario, Colombia.	08-2005 – 08-2007
	Research assistant. Infectious Disease Unit, School of Medicine, Universidad del Rosario, Colombia.	11-2005 – 11-2006
Teaching experience	**Teaching assistance.** KU Leuven, Belgium Courses of Bioinformatics: Evolutionary and Quantitative Genetics and Bioinformatics and Systems Biology: Sequence, Structure and Evolution.	05-2012 –12-2013
	Lecturer. Infectious Diseases Unit, Universidad del Rosario, Colombia. Clinical Microbiology course for undergraduate medical students.	09-2007 – 11-2009
	Supervision of undergraduate students. Pontificia Universidad Javeriana and Universidad del Rosario, Colombia.	09-2006 – 03-2007
	Peer Tutor of School of Medicine. Universidad del Rosario, Colombia. Human Genetics to undergraduate medical students.	01-2004 – 06-2004
Clinical experience	**Infectious Diseases with emphasis on HIV**. UZ Leuven, Leuven, Belgium. Clinical observer in the *Interne geneeskunde-infectieziekten*	09-2012 – 09-2013

HIV/AIDS Physician. Asistencia Científica de Alta Complejidad S.A.S., Colombia.	11-2007 – 07-2010
HIV/AIDS Physician. Adriana Villalba Foundation, Colombia.	10-2008 – 07-2009

EDUCATIONAL BACKGROUND

Biomedical Sciences PhD	Universidad del Rosario, Colombia KU Leuven, Belgium.	08-2009 – 08-2014 10-2010 – 08-2014
Master on HIV care	Universidad Rey Juan Carlos, Spain.	03-2010 – 03-2011
Certificate on HIV care	Universidad del Rosario, Colombia	03-2008 – 07-2008
Master of Basic Sciences emphasis in Genetics	Universidad del Rosario, Colombia.	08-2005 – 09-2007
Medical Doctor	Universidad del Rosario, Colombia. 3rd place among 31 students	06-1999 – 06-2005

HONORS AND AWARDS

- Peer Tutor of School of Medicine. Teaching assistance only bestowed upon the 10% best qualified students.
- Honours Degree. Master of Basic Sciences with emphasis in Genetics.
- Scholarship for Doctoral Research Training Program “Francisco Jose de Caldas” by the Departamento Administrativo de Ciencia, Tecnología e Innovación, COLCIENCIAS, Republic of Colombia.
- Scholarship by ERACOL (Erasmus Mundus Programme), academic exchange in medicine and the health sciences, between Europe and Latin America. European Union.

PUBLICATIONS IN INTERNATIONAL JOURNALS

1. **Pineda-Peña AC,** Schrooten Y, Vinken L, Ferreira F, Li G, Sequeira Trovão N, Khouri R, Derdelinckx I, De Munter P, Kücherer C, Kostrikis LG, Nielsen C, Littsola K, Wensing A, Stanojevic M, Paredes R, Balotta C, Albert J, Boucher C, Gomez-Lopez A, Van Wijngaerden E,

Van Ranst M, Vercauteren J, Vandamme AM, Van Laethem K. Trends and predictors of transmitted drug resistance (TDR) and clusters with TDR in a local Belgian HIV-1 epidemic. Plos One. 2014; 9(7): e101738

2. Camargo M, Soto de Leon SC, Muñoz M, , Sanchez R, Peña-Herrera D, **Pineda-Peña AC**, Sussmann O, Paez C, Perez-Prados A, Patarroyo ME, Patarroyo MA. Human Papillomavirus detection in women with and without Human Immunodeficiency virus infection in Colombia. BMC Cancer. 2014; 14:451

3. Muñoz M, Camargo M, Soto de Leon SC, Sanchez R, **Pineda-Peña AC**, Perez-Prados A, Patarroyo ME, Patarroyo MA. Classical molecular tests using urine samples as a potential screening tool for human papillomavirus detection in human immunodeficiency virus-infected women. *Journal of Clinical Microbiology*. 2013; 51(11):3688-93

4. **Pineda-Peña AC,** Faria NR, Imbrechts S, Libin P, Gomez-Lopez A, Abecasis AB, Camacho RJ, de-Oliveira T, Vandamme AM. 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.* 2013 Oct (19):337-48.

5. Muñoz M, Camargo M, Soto de Leon SC, Sanchez R, Parra D, **Pineda AC**, Sussmann O, Perez-Prados A, Patarroyo ME, Patarroyo MA. Human Papillomavirus Detection from Human Immunodeficiency Virus-Infected Colombian Women's Paired Urine and Cervical Samples. Plos One. 2013; 8 (2): e56509

6. **Pineda-Peña AC,** Bello DC, Sussmann O, Vandamme AM, Van Laethem K, Gomez-Lopez A. HIV Transmitted Drug Resistance in Latin America: What do we know? Aids Reviews. 2012; 14: 256-267.

7. Gaona MA, Rios DI, **Pineda AC**, Peña MC, Ibáñez M, Ramírez G. Variability of *Staphylococcus aureus* Carriers on a Medicine Student's Population. Revista Ciencias de la Salud. 2009 January; 7(1):37-46.

8. **Pineda A**, Rubio C, Restrepo CM. Juberg-Hayward syndrome: case report. Pediatría. 2006 March; 41(1):57-61.

9. **Pineda A**, Bastidas R, Barragán AM, Bustamante MC, Restrepo CM. Short rib-polydactyly Type II Saldino Noonan Syndrome. Revista Colombiana de Radiologia. 2004 June; 15(2):1581-84.

SUBMITTED PUBLICATIONS TO INTERNATIONAL JOURNALS

1. **Pineda-Peña AC**, Faria NR, Diaz FJ, Olaya P, Møller-Frederiksen C, Guangdi L, Gomez-Lopez A, Paraskevis D, Lemey P, Vandamme AM, The Colombian HIV-1 epidemic is dominated by subtype B and linked with Spain. *Under review in AIDS*

2. Li G, Piampongsant S, Faria NR, Voet A, **Pineda-Peña AC**, Khouri R, Lemey P, Vandamme AM, Theys K. An integrated map of HIV genome-wide diversity from a population perspective. *Under review in Retrovirology*

3. Li G, Theys K, Verheyen J, **Pineda-Peña AC**, Khouri R, Piampongsant S, Eusébio M, Ramon J, Vandamme AM. A new ensemble coevolution system for detecting HIV-1 protein coevolution. *Under review in Biology Direct*

4. Faria NR, Sigaloff KC, Van de Vijver DAMC, Tatem AJ, **Pineda-Peña AC**, Wallis CL, Suchard MA, Rinke de Wit TF, Hamers RL, Lemey P, Ndembi N**.** The impact of highway corridors on HIV-1 spread in East Africa. *Submitted*

5. Kouri V, Khouri R, Alemán Y, Abrahantes Y, Vercauteren J, **Pineda-Peña AC**, Theys K, Megens S, Moutschen M, Pfeifer N, Van Weyenbergh, Pérez AB, Pérez J, Pérez L, Van Laethem K, Vandamme AM. *Submitted*

MEETING ABSTRACTS

1. Eusebio M, Winand R, **Pineda-Peña AC**, Li G, Gomez P, Camacho RJ, Vandamme AM, Abecasis AB. Transmission clusters of drug resistance in subtype B in Portugal. 12th European Workshop on HIV & Hepatitis: Treatment strategies & Antiviral Drug Resistance; 2014; Barcelona, Spain

2. **Pineda-Peña AC**, Schrooten Y, Vinken L, Ferreira F, Li G, Sequeira Trovão N, Khouri R, Derdelinckx I, De Munter P, Kücherer C, Kostrikis LG, Nielsen C, Littsola K, Wensing A, Stanojevic M, Paredes R, Balotta C, Albert J, Boucher C, Gomez-Lopez A, Van Wijngaerden E, Van Ranst M, Vercauteren J, Vandamme AM, Van Laethem K. Predictors of HIV-1 transmitted drug resistance and transmission networks in the Leuven cohort. 12th European Workshop on HIV & Hepatitis: Treatment strategies & Antiviral Drug Resistance; 2014; Barcelona, Spain

3. Faria NR, Sigaloff KC, Van de Vijver DAMC, Tatem AJ, **Pineda AC**, Wallis CL, Suchard MA, Rinke de Wit TF, Hamers RL, Lemey P, Ndembi N. Migration of HIV-1 Subtypes in East Africa is Associated with Proximity to Highway Corridor. Conference on Retroviruses and Opportunistic Infections; 2014 Feb 23-26; Boston, USA.

4. Arruda LB, Beheydt G, Benevides P, Aguiar B, **Pineda AC**, Van Laethem K, Duarte AJDaS, Casseb J**.** Bayesian network approaches to evaluate novel V3 mutations associated to HIV-1 tropism. Congreso Argentino de Bioinformática y Biología Computacional (4CAB2C) y 4ta. Conferencia Internacional de la Sociedad Iberoamericana de Bioinformática (SolBio); 2013 Oct 29-31; Rosario, Argentina.

5. **Pineda-Peña AC**, Ferreira F, Schrooten Y, Vinken L, Derdelinckx I, De Munter P, Van Ranst M, Gomez-Lopez A, Van Wijngaerden E, Vandamme AM, Vercauteren J, Van Laethem K. Drug Resistance in Newly Diagnosed HIV-1 Infected Patients in Leuven, Belgium from 1998 to 2011: Results from Routine Genotyping Testing. 14th European AIDS Conference/EACS; 2013 Oct 16-19; Brussels, Belgium.

6. Eusebio M, Winand R, **Pineda-Peña AC**, Li G, Camacho RJ, Vandamme AM, Abecasis AB. TDR Levels and Mutation Patterns in Portugal Differ Significantly in HIV-1 Subtype G Compared to Subtype B. 14th European AIDS Conference/EACS; 2013 Oct 16-19; Brussels, Belgium.

7. **Pineda-Peña AC**, Faria NR, Diaz FJ, Olaya P, Møller-Frederiksen C, Guangdi L, Gomez-Lopez A, Lemey P, Vandamme AM, The Colombian epidemic is dominated by HIV-1 subtype B over time: a Molecular Epidemiology and Phylodynamics study. 20th International HIV, Dynamics & Evolution Conference; 2013 May 8-11; Utrecht, Netherlands.

8. Faria NR, Sigaloff KC, Van de Vijver DAMC, Tatem AJ, **Pineda AC**, Wallis CL, Suchard MA, Rinke de Wit TF, Hamers RL, Lemey P, Ndembi N**.** Phylodynamic evidence for the impact of highway corridors on HIV-1 spread in east Africa. 20th International HIV, Dynamics & Evolution Conference; 2013 May 8-11; Utrecht, Netherlands.

9. **Pineda-Peña AC**, Faria NR, Imbrechts S, Libin P, Gomez-Lopez A, Abecasis AB, Camacho RJ, de-Oliveira T, Vandamme AM, Performance of the subtyping tools in the surveillance of the HIV-1 epidemic: comparison between the new Rega version 3 and six other automated tools to identify pure subtypes and circulating recombinant forms. BREACH Symposium 2012: Belgian Research Aids&HIV consortium; 2012 September 28-29; Brussels, Belgium. Available in www.breach-hiv.be/p_194.htm.

10. **Pineda-Peña AC**, Faria NR, Imbrechts S, Libin P, Gomez-Lopez A, Abecasis AB, Camacho RJ, de-Oliveira T, Vandamme AM, Performance of the subtyping tools in the surveillance of the HIV-1 epidemic. 17th International BioInformatics Workshop on Virus Evolution and Molecular Epidemiology; 2012 August 27-31; Belgrade, Serbia

11. Rojas Y, **Pineda AC**, Sussmann O. Social characteristics of HIV patients in an Institution of Bogotá, Colombia. Proceedings of VII Encuentro Nacional de Investigación en Enfermedades Infecciosas; 2009 May 20-22; Paipa, Colombia: Infectio: 2010

12. **Pineda AC**, Rodríguez G, Silva CT, Ríos DI, Gaona MA, Vizcaíno C, et al. Epidemiology of Nosocomial *Candida* in Bogotá, Colombia by Phenotyping and Genotyping methods. Proceedings of I Congreso Latinoamericano de Genética Humana y el IX Congreso Colombiano de Genética; 2008 October 8-10; Cartagena de Indias, Colombia. Acta Biológica Colombiana, 2008 supplement.

13. Rodríguez G, **Pineda AC**, Silva CT, Ríos DI, Amaya JC, Lagos A, et al. *Burkholderia cepacea* complex: Report of outbreak in Hospital of Bogotá, Colombia. Saludarte 2006 October;5(1):58.

ORAL PRESENTATIONS

1. **Pineda AC**, Rodríguez G, Silva CT, Ríos DI, Gaona MA, Vizcaíno C, et al. Epidemiology of Nosocomial *Candida* in Bogotá, Colombia by Phenotyping and Genotyping methods. Presented at: Seminario de Micología Médica "De la Clínica al Laboratorio". 2008 April; Bogotá, Colombia.

2. **Pineda A**, Rubio C, Restrepo CM. Juberg-Hayward Syndrome. Presented at: IV Congreso Internacional y VII Congreso Colombiano de Genética; 2006 February 23-25; Bucaramanga, Colombia.