

UNIVERSITA' DEGLI STUDI DI MILANO

Facoltà di Medicina e Chirurgia

XXIII CICLO DOTTORATO DI RICERCA IN  
MALATTIE INFETTIVE

**ANDAMENTO DELLA PREVALENZA  
IN ITALIA DEI CEPPI NON-B DI HIV-1  
NEGLI ULTIMI TRENTA ANNI**

Relatore: Prof. Massimo Galli

Tesi di dottorato di

**Ilaria Caramma**

Matricola n. R07898

Anno Accademico 2009-2010

## INDICE

---

<b>INDICE</b>	<b>1</b>
<b>1 INTRODUZIONE</b>	<b>3</b>
CLASSIFICAZIONE DI HIV	4
VARIABILITA' GENETICA DI HIV-1	9
MECCANISMI DELLA VARIAZIONE VIRALE	9
MUTAZIONE	9
RICOMBINAZIONE	10
CLASSIFICAZIONE DELLE FORME GENETICHE DI HIV	13
DISTRIBUZIONE GEOGRAFICA DELLE FORME GENETICHE DI HIV-1	20
IMPLICAZIONI BIOLOGICHE DELLA VARIABILITÀ DI HIV-1	23
IDENTIFICAZIONE E CIRCOLAZIONE DEI SOTTOTIPI NON-B DI HIV-1	24
MODIFICAZIONI DELL'EPIDEMIA DA HIV-1 IN ITALIA	25
<b>2 SCOPO DELLA RICERCA</b>	<b>30</b>
<b>3 MATERIALI E METODI</b>	<b>34</b>
CASISTICA	35
METODI	36
ANALISI FILOGENETICA	36
DISCRIMINAZIONE TRA SOTTOTIPO B E NON-B MEDIANTE STRUMENTI DISPONIBILI IN RETE	37
ANALISI FILOGENETICA DEI SOTTOTIPI PURI E DELLE CRF	37
ANALISI DEI CEPPI RICOMBINANTI NON TIPIZZABILI COME CRF MEDIANTE IL PROGRAMMA SIMPLOT	39
CONTRIBUTO ALLA CARATTERIZZAZIONE E TIPIZZAZIONE DEI CEPPI RICOMBINANTI TRAMITE LA RICERCA BLAST	40
ANALISI STATISTICA	41
<b>4 RISULTATI</b>	<b>42</b>
CARATTERISTICHE DELLA POPOLAZIONE	43
ANDAMENTO TEMPORALE DEI SOTTOTIPI NON-B NEL PERIODO 1980-2008	46
DISTRIBUZIONE DEI CEPPI DI HIV-1 NEL TEMPO RISPETTO ALLE VARIABILI DEMOGRAFICHE	47
FATTORI ASSOCIATI ALL'INFEZIONE CON SOTTOTIPO NON-B	49
ASSEGNAZIONE DEL SOTTOTIPO E DISTRIBUZIONE DEI CEPPI NELLE DIVERSE ETNIE	51
ANALISI DI RICOMBINAZIONE DELLE VARIANTI DI HIV-1	54
<b>5 DISCUSSIONE</b>	<b>55</b>
<b>6 BIBLIOGRAFIA</b>	<b>60</b>

## 1 INTRODUZIONE

---

## CLASSIFICAZIONE DI HIV

---

Il virus dell'immunodeficienza umana (HIV), l'agente eziologico della sindrome da immunodeficienza acquisita (AIDS), è un virus a singola elica a polarità positiva ed è classificato come appartenente alla famiglia delle *Retroviridae*.

I Retrovirus, conosciuti fino agli anni ottanta, soprattutto per le patologie a carattere neoplastico che provocano in alcune specie animali, sono suddivisi in tre sottofamiglie: *Oncovirinae*, *Spumavirinae* e *Lentivirinae*.

Gli Oncovirus raggruppano i virus isolati originariamente come agenti trasformanti; causano sarcomi, leucemie, tumori mammari e talvolta immunosoppressione in diverse specie animali. Appartengono a questa famiglia anche due virus umani: HTLV-I e HTLV-II (*Human T-cell Leucemia Virus*), associati a linfomi delle cellule T e a disordini neurologici.

Gli Spumavirus devono la loro denominazione alla capacità di indurre lesioni vacuolari nelle cellule infettate e conferiscono un aspetto schiumoso alle cellule coltivate *in vitro*.

I Lentivirus sono associati a malattie a lungo decorso ma senza rapporto diretto con le neoplasie. Prima dell'isolamento e caratterizzazione di HIV erano noti solo per patologie provocate in alcune specie di felini ed ungulati. HIV condivide molte caratteristiche biologiche e molecolari con HTLV, come le modalità di trasmissione, il tropismo verso le cellule T e la capacità di indurre formazione di sincizi. Tuttavia viene classificato in modo più appropriato tra i Lentivirus sulla base dell'organizzazione genica, delle analogie di sequenza e delle caratteristiche morfologiche e patologiche. Il *core* di HIV è simile a quello dei Lentivirus animali<sup>1</sup>.

Omologhe sono anche le sequenze aminoacidiche codificate dai geni *pol*<sup>2</sup>. Inoltre i Lentivirus animali e HIV condividono la complessità del genoma, una *open reading frame* del gene *env* particolarmente estesa e la sovrapposizione dei geni *gag* e *pol*. Infine un'altra caratteristica, che contraddistingue l'intera sottofamiglia, è la tendenza a

indurre malattie cronico-degenerative nei loro ospiti, caratterizzate da un lungo periodo di latenza clinica e da un coinvolgimento sia del sistema immunitario sia di quello nervoso.

Sono noti due sierotipi di HIV<sup>3</sup> (Figura 1), denominati HIV-1 e HIV-2 che si differenziano per organizzazione genomica e per virulenza. HIV-2 ha, rispetto ad HIV-1, una più bassa virulenza che gli conferisce un più esteso periodo di latenza clinica e bassa infettività. Probabilmente proprio a causa di queste differenze HIV-1 è ormai diffuso in tutto il mondo<sup>4</sup> ed è considerato la causa primaria della pandemia di AIDS, mentre HIV-2 rimane confinato all'area dell'Africa Occidentale<sup>5</sup> e a paesi con immigrazione di soggetti provenienti da questa zona.

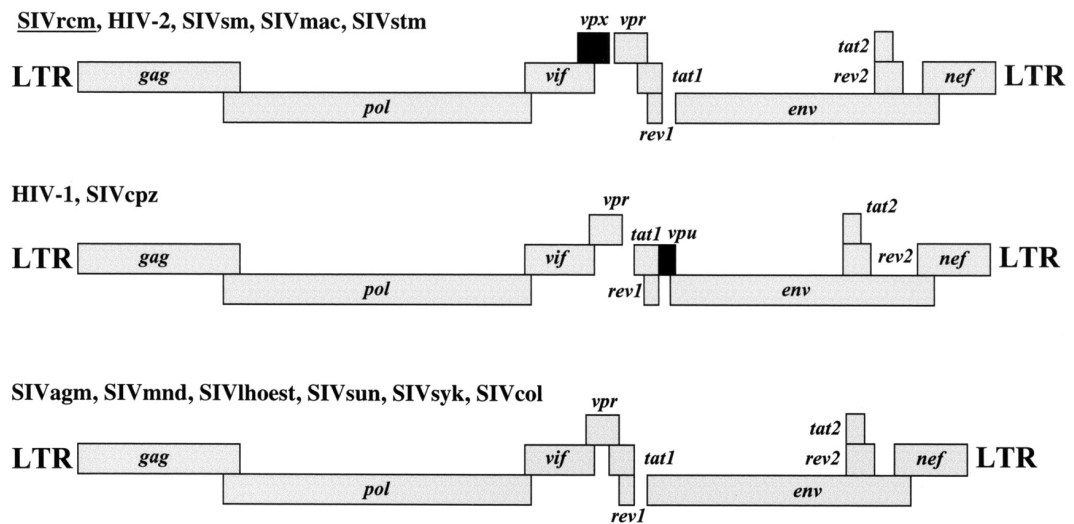


Figura 1. Confronto fra l'organizzazione genomica di HIV-1, HIV-2 e SIV.

La storia evolutiva di HIV-1 e HIV-2 è stata ricostruita nel dettaglio tramite le informazioni ottenute da alberi filogenetici dei lentivirus dei primati<sup>6</sup> (Figura 2).

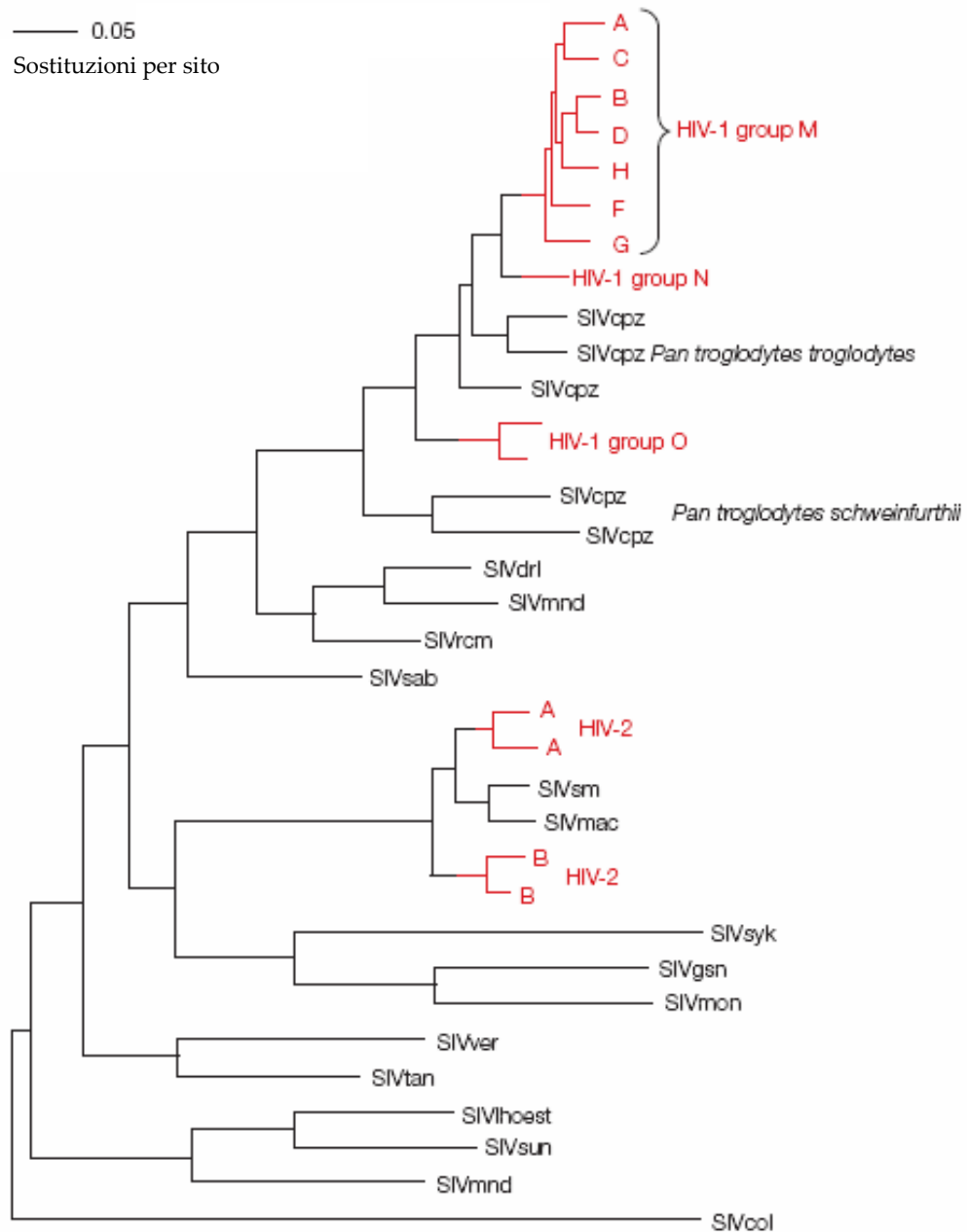


Figura 2. Albero filogenetico dei Lentivirus dei primati che mostra la correlazione tra SIV e HIV.

L'analisi filogenetica ha mostrato come i due sierotipi di HIV siano correlati a differenti forme di SIV<sup>7</sup> (*Simian Immunodeficiency Virus*) e abbiano avuto quindi origini distinte. In particolare, HIV-1 è risultato più vicino a SIVcpz, che infetta alcune sotto specie di scimpanzé, *Pan troglodytes troglodytes* e *Pan troglodytes schweinfurthii*, che popolano rispettivamente aree equatoriali dell'Africa occidentale e centrale. Al contrario HIV-2 sembrerebbe maggiormente correlato a SIVsm che infetta le scimmie sooty mangabey (*Cercocebus atys*), il cui areale di infezione, nell'Africa occidentale, coincide con il probabile luogo di origine di HIV-2. L'analisi filogenetica suggerisce inoltre che siano avvenuti molteplici eventi di trasmissione scimmia-uomo, dato il mix di *lineage* HIV- SIV generato. È difficile stabilire con precisione quanti eventi di trasmissione siano intercorsi, probabilmente almeno 4 per HIV-2 e 3 per HIV-1<sup>8</sup>, uno per ciascun gruppo, in accordo con la loro origine indipendente. Tuttavia, gli eventi di ricombinazione inter-specie, comuni tra i lentivirus dei primati, complicano fortemente questo tipo di analisi. Ad esempio il gruppo N sembrerebbe essersi originato dalla ricombinazione tra un ceppo di SIVcpz e un virus strettamente correlato con l'antecore del gruppo M; tuttavia tale evento sarebbe accaduto prima che i gruppi M ed N si diffondessero nell'uomo.

Dati recenti indicano inoltre che lo stesso SIVcpz di *Pan troglodytes troglodytes* sia un virus ricombinante derivato a sua volta da SIVgsn (*SIV Greater Spotted Monkeys*) e SIVrcm<sup>9</sup> (*SIV Red-capped Mangabeys*), due patogeni che colpiscono scimmie predate dagli scimpanzé.

Il gruppo O non si è invece originato da un ceppo di SIV presente negli scimpanzé, ma da un ceppo presente nei gorilla dell'Africa occidentale (SIVgor)<sup>10</sup>.

Questi risultati sono di estremo interesse poiché indicano che anche SIVcpz, come HIV, è derivato da un evento di infezione interspecie, fornendo così una conferma indiretta all'ipotesi sull'origine dei Lentivirus umani. Secondo l'ipotesi più accreditata, il virus è stato trasmesso all'uomo attraverso l'esposizione cutanea o mucosale a sangue infetto, causata dalle attività venatorie o di macellazione di animali infetti.

L'analisi filogenetica basata sulle sequenze del gene *env*<sup>11</sup> ha permesso di suddividere



## VARIABILITA' GENETICA DI HIV-1

---

Una delle caratteristiche peculiari di HIV-1 è l'estrema variabilità genetica, che si manifesta non solo tra isolati di individui diversi, ma anche tra quelli di uno stesso soggetto nel corso dell'infezione.

Le basi molecolari della variazione virale sono dovute sia all'elevata frequenza di errore nell'incorporazione dei nucleotidi da parte della trascrittasi inversa ( $3.4 \times 10^{-5}$  sostituzioni per ciclo replicativo), sia alla marcata velocità di replicazione del virus. Considerando che il genoma di HIV-1 è lungo circa diecimila paia di basi e che la carica virale totale in soggetti con infezioni da HIV-1 è compresa tra i dieci e i cento miliardi di virioni, si può calcolare che ogni singola mutazione nucleotidica possa verificarsi tra le diecimila e le centomila volte al giorno, in ogni singolo paziente<sup>14</sup>.

Questi fenomeni comportano che all'interno della popolazione virale possa insorgere rapidamente un'ampia divergenza genetica, generando in ciascun soggetto infetto delle quasispecie, varianti virali strettamente correlate ma distinte.

## MECCANISMI DELLA VARIAZIONE VIRALE

---

La variazione genetica di HIV-1 ha alla sua base due meccanismi molecolari: la mutazione e la ricombinazione.

### MUTAZIONE

---

La bassa accuratezza della retrotrascrittasi di HIV-1 ha come diretta conseguenza l'insorgenza di due tipi di mutazioni: sostituzioni puntiformi e inserzioni o delezioni.

La sostituzione è il cambio di un nucleotide con uno diverso durante la sintesi del

DNA virale. Nel caso di una mutazione non sinonima può accadere che la proteina mutata comporti una resistenza farmacologica, una diversa fitness che, a loro volta, possono fornire un vantaggio evolutivo al virus.

L'inserzione e la delezione sono aggiunte o eliminazioni di uno o più nucleotidi. Possono alterare il frame di lettura dell'mRNA, con la conseguente formazione di proteine difettive.

Il gene env sembra possedere un maggior grado di variabilità rispetto alle altre regioni del genoma di HIV-1<sup>15,16,17,18</sup>.

Analizzando sequenze di env provenienti da un singolo paziente, è possibile riscontrare la presenza di quasispecie che differiscono tra loro per il 2-5% della sequenza del gene<sup>19</sup>. Considerando invece isolati provenienti da soggetti diversi, ma appartenenti alla stessa area, la variabilità sale al 6-19% e raggiunge valori del 20-30% se si considerano anche differenti aree geografiche<sup>20,21</sup>.

L'elevato tasso di variabilità di env permette di valutare con precisione le distanze genetiche tra diversi isolati e per questo motivo è stato a lungo utilizzato per l'analisi filogenetica; recentemente però, è stato dimostrato che il gene pol, pur avendo un tasso di variabilità inferiore, contiene sufficiente informazione per ricostruire gli eventi di trasmissione mediante filogenesi<sup>22</sup>; inoltre, il suo frequente utilizzo in ambito clinico per ottenere i saggi genotipici di resistenza, ha reso disponibile una vasta quantità di sequenze.

## RICOMBINAZIONE

---

Come descritto in Figura 4 la retrotrascrittasi di HIV-1 ha la capacità di trasferire la sintesi della catena nascente di DNA da un template all'altro (strand switching); poiché ogni virione contiene due molecole di RNA, attraverso tale meccanismo la RT è in grado di generare ricombinazione tra i template parentali, generando una progenie virale che possiede nuove caratteristiche genetiche.

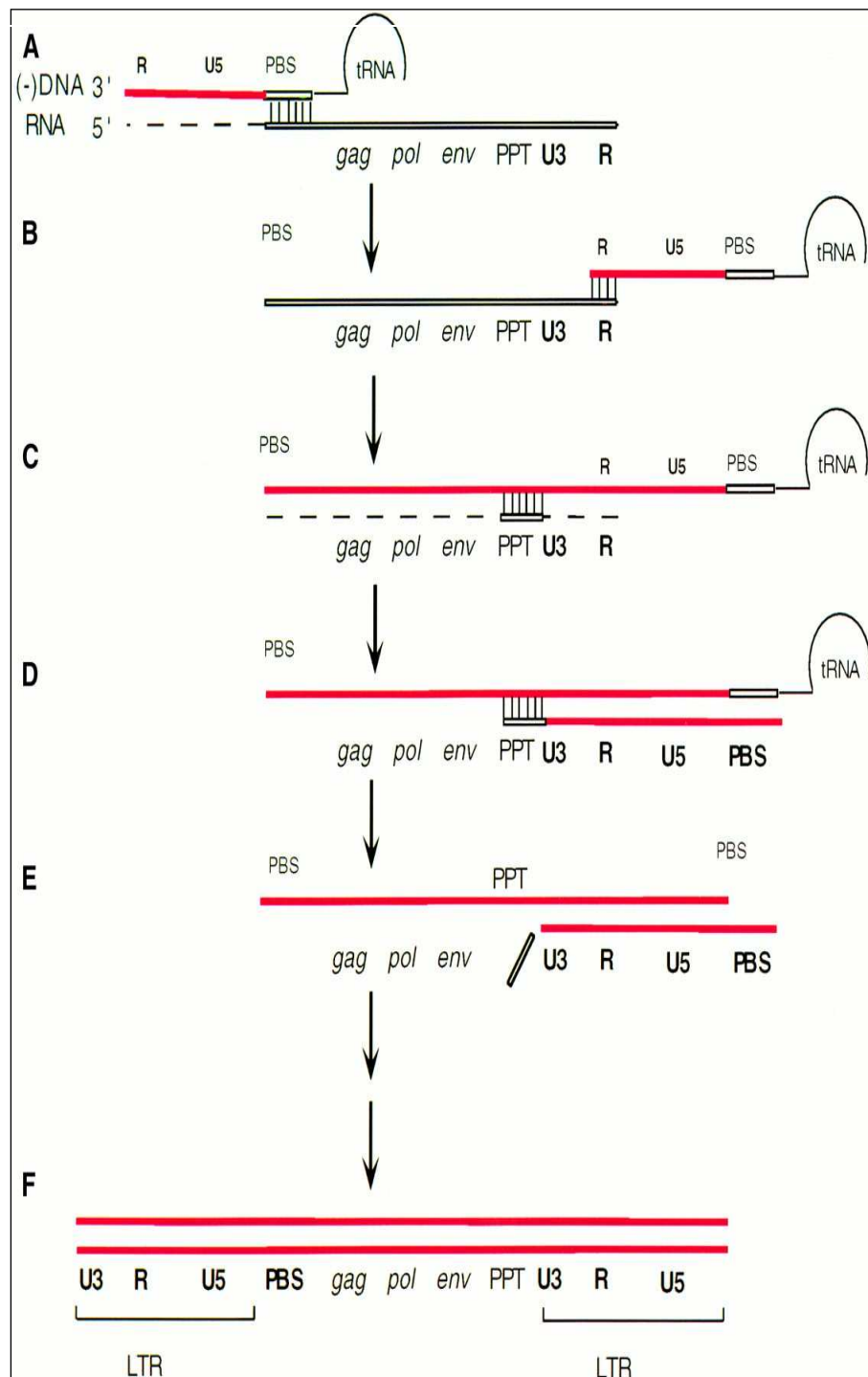


Figura 4. Meccanismo di trascrizione inversa del genoma di HIV-1.

(A) inizia la sintesi del filamento negativo di DNA, utilizzando come *primer* un tRNA cellulare che si associa al PBS. L'RNA template viene degradato dall'RNasi H dell'RT.

(B) il primo salto permette l'associazione della catena neosintetizzata di DNA con l'estremità 3' del genoma virale. Il trasferimento è mediato dalle sequenze R delle LTR.

(C) riprende la sintesi di DNA, accompagnata dalla digestione di tutto l'RNA template tranne la regione PPT.

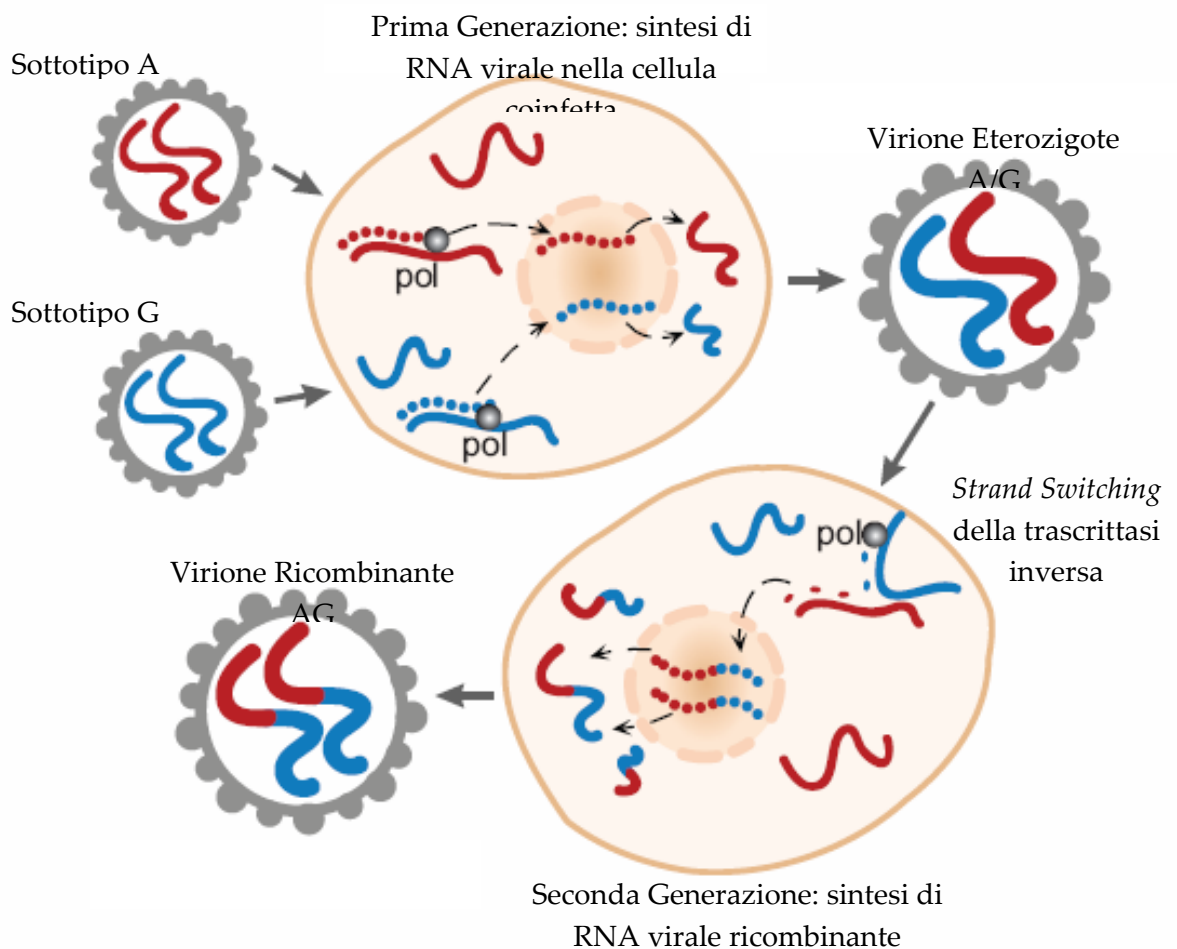
(D) PPT viene usato come *primer* per sintetizzare la seconda catena di DNA.

(E) l'RNasi H rimuove la sequenza PPT e il tRNA.

(F) col secondo salto la catena di DNA appena formata (il filamento positivo) si associa alla sequenza PBS del primo filamento.

Il completamento della seconda catena di DNA produce una molecola lineare di DNA a doppio filamento, con le LTR ad entrambe le estremità.

Questo fenomeno ha particolare rilevanza in presenza di coinfezione o superinfezione da HIV-1 (Figura 5), ossia quando la stessa cellula viene infettata da virus di ceppi diversi; in tal caso si avrà l'integrazione di differenti provirus nel nucleo cellulare. Alla prima espressione e maturazione dei provirus si genera una popolazione mista, contenente sia virioni parentali omozigoti, sia virioni eterozigoti (prima generazione). Nelle successive infezioni, sostenute dai virioni eterozigoti di prima generazione, l'attività di *strand switching* della retrotrascrittasi porta alla formazione di un DNA provirale chimerico, contenente tratti di entrambi i genomi parentali. Le particelle virali prodotte da questo nuovo provirus costituiscono una seconda generazione di virioni ricombinanti omozigoti<sup>23</sup>, e vengono oggi definite forme circolanti ricombinanti (*Circulating Recombinant Forms, CRF*).



**Figura 5. Modello di formazione delle forme ricombinanti di HIV-1.**

## CLASSIFICAZIONE DELLE FORME GENETICHE DI HIV

Inizialmente, gli isolati di HIV-1 sono stati raggruppati in distinti sottotipi sulla base dell'analisi filogenetica di sequenze parziali dei geni gag ed env. All'interno dell'albero filogenetico le varianti virali appartenenti al medesimo sottotipo risultavano raggruppate tra loro, costituendo dei gruppi monofiletici con un nodo comune; questo ha suggerito l'esistenza di un medesimo antecessore per tutti i sottotipi noti. In seguito all'identificazione di nuovi ceppi di HIV-1 e di CRF, caratterizzati mediante l'analisi dell'intero genoma virale, si è resa necessaria una revisione della tassonomia di HIV-1 e l'introduzione di linee guida internazionali per una corretta nomenclatura:

Classificazione	Definizione	Esempi
<b>Sottotipo o <i>Clade</i></b>	Ceppi di HIV-1 equidistanti tra loro, con distribuzione filogenetica a stella piuttosto che ad albero.	A, B, C, D, F, G, H, J e K.
<b>Sotto-sottotipo</b>	Distinti <i>lineage</i> all'interno di un sottotipo; la distanza genetica tra di essi è circa la metà rispetto a quella tra i sottotipi.	I sottotipi A ed F sono ulteriormente divisi in A1, A2, A3, A4, A5 e F1, F2 rispettivamente. La maggior parte di essi circolano in Africa centrale e orientale.
<b><i>Intersubtype recombinant forms</i></b>	Ceppi con mosaicismi formato da frammenti provenienti da più sottotipi che si alternano lungo il genoma.	Comuni in aree dove più sottotipi co-circolano. Derivano da eventi di superinfezione o coinfezione.
<b><i>Circulating Recombinant forms</i></b>	Specifiche forme ricombinanti che si diffondono nella popolazione; vengono definite tali quando identificate in almeno tre individui non epidemiologicamente correlati; il nome contiene il numero progressivo di identificazione ed i sottotipi che la compongono; <i>cpx (complex)</i> viene utilizzato per le forme che contengono tre o più sottotipi.	Fino ad ora ne sono state descritte 48 (di cui le ultime tre in attesa di pubblicazione).
<b><i>Unique recombinant forms</i></b>	Forme ricombinanti caratterizzate in un singolo individuo.	In numero imprecisato, descritte sulla base di sequenze genomiche complete o parziali. Il loro potenziale nella pandemia è sconosciuto.

<b>Lineage geograficamente distinti</b>	<i>Lineage</i> specifici associati ad alcune regioni che sono distinguibili filogeneticamente; al contrario dei sotto-sottotipi non sono equidistanti all'interno di un sottotipo.	Thai B, Indian C, West vs. East African D e FSU-A (Former Soviet Union)
---	--	---

Nella classificazione attuale<sup>24</sup>, vengono riconosciuti tre gruppi filogenetici di HIV-1: M (*Major*), O (*Outlier*) e N (*Non-M, Non-O o New*), ai quali recentemente si è aggiunto il gruppo P.

La maggior parte delle infezioni da HIV su scala globale è sostenuta da virus del gruppo M, a cui appartengono tutti i sottotipi inizialmente identificati. I virus del gruppo O sono limitati alle popolazioni che vivono in Paesi dell'Africa centrale o epidemiologicamente correlate ad esse, in particolare il Camerun e i paesi limitrofi; tuttavia anche in queste aree rappresentano una minoranza delle infezioni. Attualmente sono state osservate solo pochi casi di infezione riconducibili al gruppo N, tutti documentati in Camerun.

All'interno del gruppo M si riconoscono nove sottotipi (A, B, C, D, F, G, H, J e K), ritenuti originari dell'Africa centrale. Il sottotipo E era stato inizialmente classificato sulla base di sequenze parziali del gene *env* come filogeneticamente distinto dai sottotipi A, B, C e D allora noti. Una volta valutata l'intera sequenza genomica, è risultato invece essere un ricombinante dei sottotipi A ed E, ed è stato la prima forma ricombinante ufficialmente riconosciuta, la CRF01\_AE.

Tuttavia un ceppo E puro non è mai stato isolato; si può al momento solo ipotizzare che sia esistito in passato e che successivamente, in seguito a selezione o eventi di ricombinazione, si sia estinto lasciando porzioni del ceppo parentale all'interno dei ricombinanti. Dati recenti inoltre mettono in dubbio l'origine ricombinante della CRF01\_AE, ipotizzando una divergenza genetica dal ceppo A dovuta ad un'elevata frequenza di mutazione della regione *env*; secondo tale ipotesi quindi, il ceppo ricombinante dovrebbe essere più correttamente considerato un sotto-sottotipo A<sup>25</sup>.

Anche il sottotipo I, inizialmente classificato sulla base di un numero limitato di

sequenze del gene *env*, è risultato essere un ricombinante di tre sottotipi (A, G e I) sulla base dell'analisi dell'intero genoma ed è quindi stato ridefinito CRF04\_cpx (cpx=*complex*). Tuttavia una successiva e più accurata analisi ha evidenziato che questa CRF presenta una organizzazione genomica di complessità unica, in cui il ceppo I non è più rappresentato e che comprende regioni di almeno quattro distinti sottotipi (A, G, H e K) oltre a regioni non classificabili<sup>26</sup>.

I sottotipi rappresentano linee evolutive di HIV che hanno raggiunto una notevole divergenza genetica e spesso hanno una certa associazione con specifiche aree geografiche. Negli alberi filogenetici sono rappresentati da gruppi equidistanti tra loro, con una distanza evolutiva pari al 25-35% calcolata sulle sequenze aminoacidiche codificate dal gene *env*.

All'interno dei sottotipi A e F si distinguono due distinti sottogruppi, denominati sotto-sottotipi, A1, A2, A3 e F1, F2 ciascuno dei quali risulta più strettamente correlato all'altro sotto-sottotipo che ad un diverso sottotipo.

Anche i sottotipi B e D sono maggiormente correlati tra loro di quanto non lo siano con gli altri, essendosi probabilmente evoluti dallo stesso ceppo progenitore. Infatti all'analisi filogenetica risultano essere poco divergenti e dovrebbero essere più correttamente definiti sotto-sottotipi B1 e B2. Tuttavia, per non generare confusione, la designazione di sottotipo D è stata mantenuta nella nomenclatura.

Infine, dalla topologia degli alberi filogenetici è possibile rilevare, all'interno di alcuni sottotipi, raggruppamenti geograficamente localizzati che hanno un progenitore comune. Ad esempio all'interno del sottotipo C si distinguono i ceppi indiani da quelli etiopi; tra i G identificati in Europa, è possibile distinguere gli spagnoli dai portoghesi<sup>27,28,29</sup>, tra gli isolati di sottotipo B si riconosce un raggruppamento dei ceppi thailandesi.

Un nuovo sottotipo, per essere definito tale, deve essere valutato sulla base della sua intera sequenza e deve essere identificato almeno in tre soggetti epidemiologicamente non correlati tra loro. Recentemente è stato individuato e sequenziato l'intero genoma di un nuovo ceppo, provvisoriamente definito come sottotipo U, che essendo stato

isolato unicamente in due soggetti, non può ancora essere considerato come il decimo sottotipo di HIV-1, non essendo soddisfatte le raccomandazioni definite dal Comitato Internazionale per la Nomenclatura di HIV<sup>30</sup>.

Si ritiene che le forme genetiche di HIV-1 il cui genoma presenta porzioni che sono filogeneticamente correlate a distinti sottotipi si siano originate a seguito di eventi di ricombinazione tra distinti sottotipi in soggetti infettati con due o più ceppi virali.

I virus ricombinanti, identificati in almeno tre soggetti epidemiologicamente non correlati tra loro e caratterizzati dopo sequenziamento dell'intero genoma, vengono designati *Circulating Recombinant Forms* o CRF<sup>31,102</sup>. Tuttavia, le linee guida definite dalla Commissione Internazionale nominata per la revisione e l'aggiornamento continuo della nomenclatura di HIV, ha stabilito che, anche se tre genomi *full-length* sono preferibili, qualora non fosse possibile ottenere la sequenza completa del terzo caso, sarebbero sufficienti anche due *full-length* in associazione con sequenze parziali di un terzo soggetto; in quest'ultimo caso le sequenze parziali devono contenere regioni che confermino la struttura a mosaico.

Le CRF (Figura 6) sono denominate con un numero progressivo che corrisponde all'ordine in cui vengono identificate, seguito dalle lettere dei sottotipi parentali (U indica le porzioni non classificabili) o da cpx (*complex*) per i ricombinanti che comprendono più di due sottotipi.

Ad oggi, sono state identificate 48 CRF<sup>32</sup>, che, per la maggior parte, sono geograficamente localizzate. Per la maggior parte di esse è stata pubblicata la caratterizzazione molecolare della struttura ricombinante; resta da chiarire il pattern di alcune di cui sono comunque disponibili in rete le sequenze.

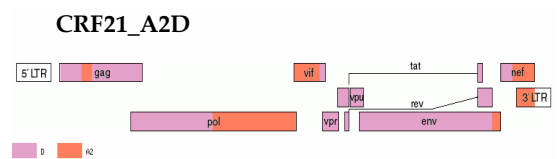
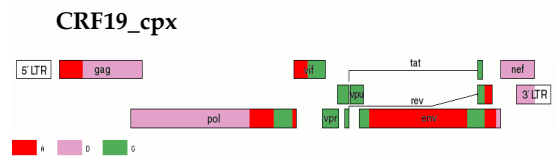
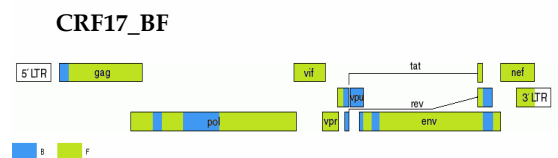
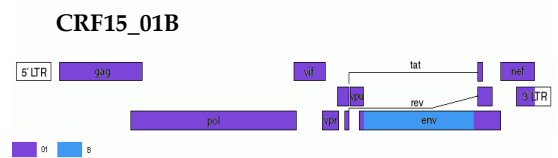
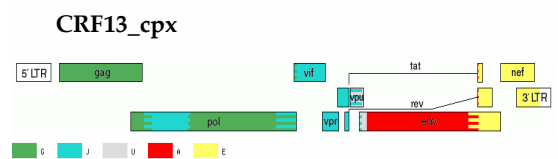
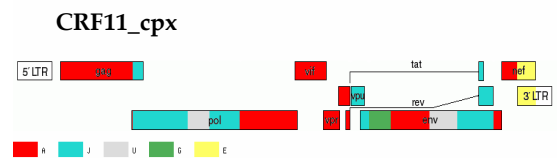
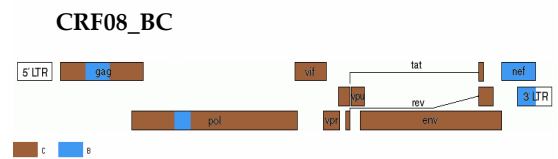
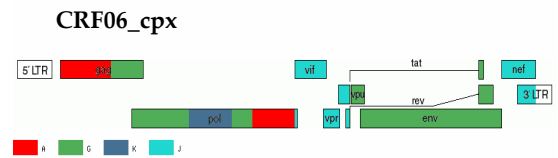
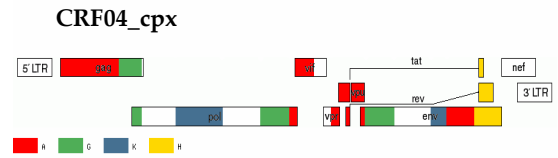
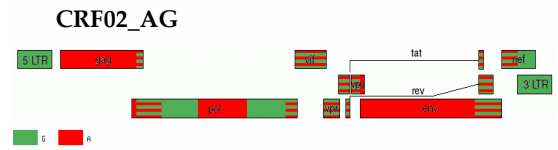
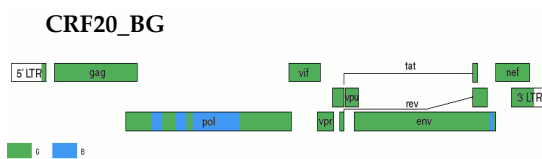
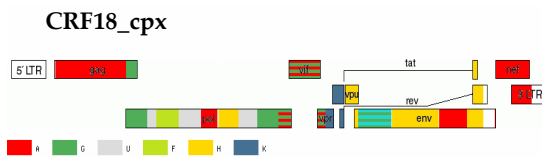
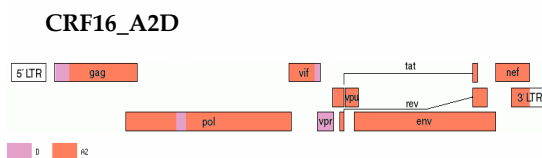
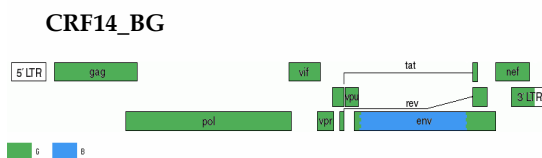
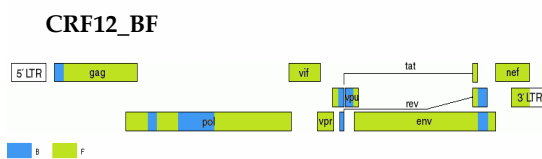
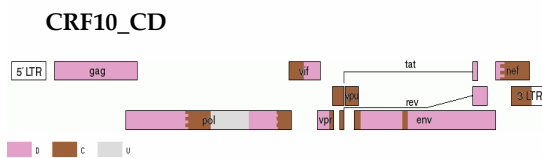
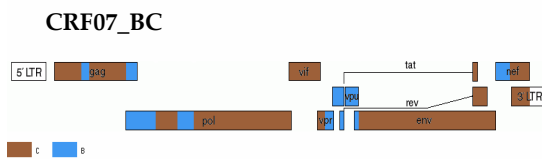
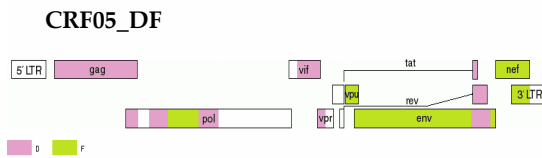
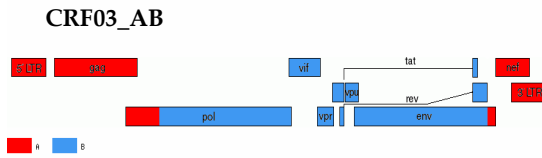
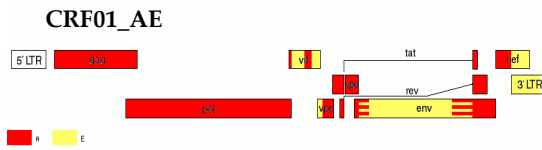
La maggior parte delle CRF hanno origine africana, tuttavia alcune si sono originate fuori dall'Africa, in aree dove i ceppi parentali sono cocircolanti ad alta prevalenza.

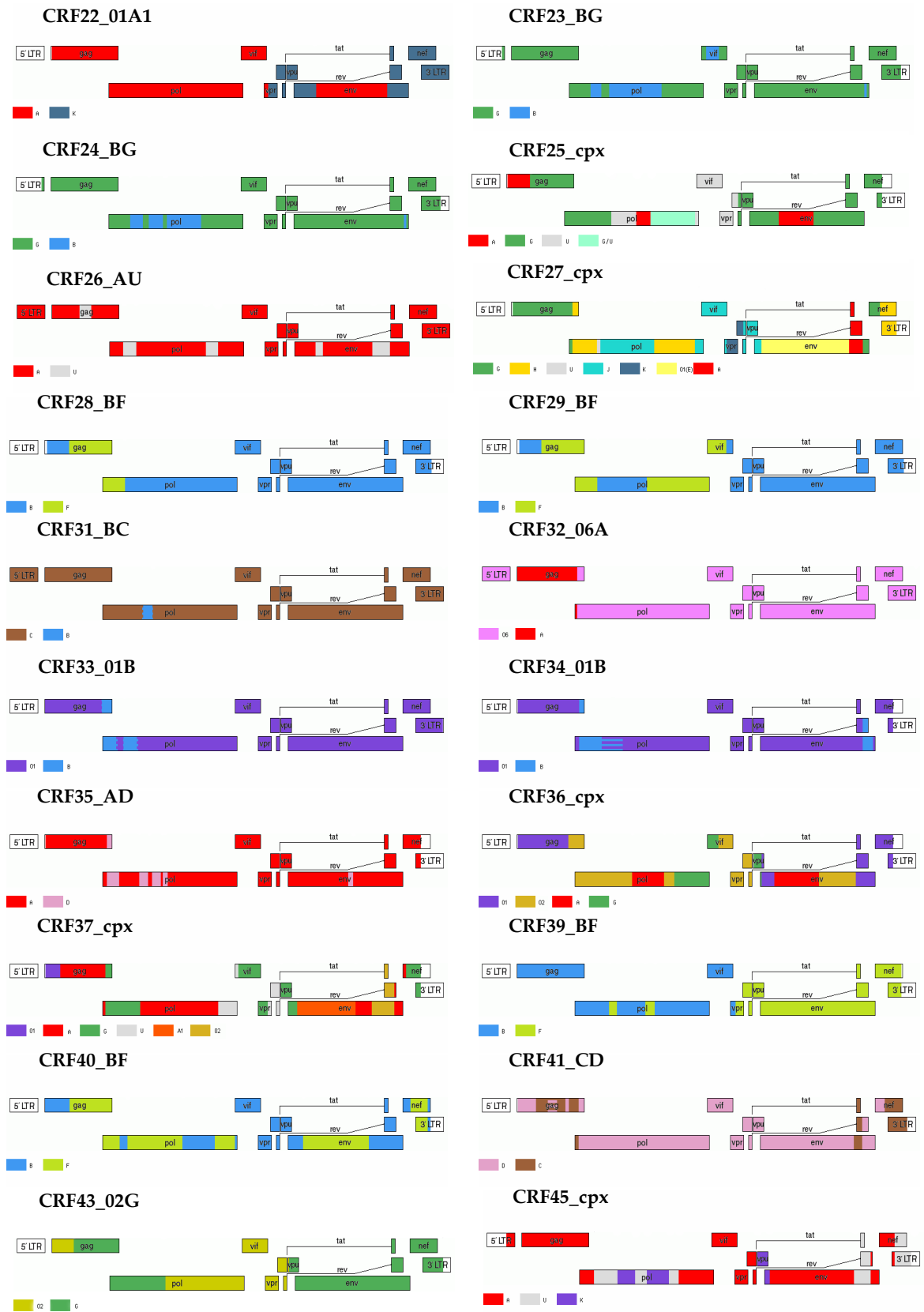
cui la CRF12\_BF<sup>33,34</sup>, oggi diffusa in tutto il mondo, rappresenta solo la prima identificata (CRF17\_BF<sup>35</sup>, CRF28\_BF<sup>36</sup>, CRF29\_BF<sup>37</sup>, CRF39\_BF, CRF40\_BF<sup>38</sup>).

Di recente identificazione sono anche i mosaicismi cubani quali la CRF18\_cpx<sup>39</sup>, la CRF19\_cpx<sup>40</sup>, la CRF23\_BG e la CRF24\_BG<sup>41</sup>.

Il fatto che il numero di CRF identificate recentemente continui ad aumentare suggerisce che l'eterogeneità di questi sottotipi sia significativamente superiore rispetto a quanto ritenuto correntemente. In un recente studio di sorveglianza è stato rilevato che l'impatto delle CRF nella pandemia globale di HIV-1 è in notevole aumento, rappresentando il 18% delle nuove infezioni<sup>42</sup>.

La distribuzione globale delle differenti forme di HIV-1 è un processo dinamico. La probabilità di rilevare nuove forme ricombinanti aumenterà all'aumentare delle varianti di HIV-1 che cocircolano nella stessa area. I pattern dei nuovi mosaicismi diventeranno sempre più complessi, in quanto la ricombinazione coinvolgerà sempre più ceppi virali che sono a loro volta delle forme mosaiche, come avvenuto per la CRF15<sup>43</sup>, la CRF33<sup>44</sup> e la CRF34<sup>45</sup> che coinvolgono la CRF01 e la CRF32<sup>46</sup> che coinvolge la CRF06<sup>47</sup>. Virus mosaici contenenti la CRF02\_AG sono la CRF36<sup>48</sup> e CRF37, entrambe di origine africana.





**Figura 6. Pannello delle Forme Circolanti Ricombinanti oggi classificate. Nonostante le CRF ad oggi identificate siano 45, la struttura ricombinante di alcune di esse non è stata ancora caratterizzata e pubblicata. Fonte: Los Alamos National Laboratory ([www.lanl.gov](http://www.lanl.gov))**

## DISTRIBUZIONE GEOGRAFICA DELLE FORME GENETICHE DI HIV-1

E' stato dimostrato che tutti i sottotipi di HIV-1 sono originati nell'Africa centrale. L'iniziale diversificazione del gruppo M dovrebbe essere avvenuta all'interno o nei pressi della Repubblica Democratica del Congo, dove sono state registrate le più alte differenze all'interno gruppo<sup>49</sup> e dove è stato documentato il più antico caso di infezione oggi noto, risalente al 1959<sup>50</sup>.

Al di fuori dell'Africa centrale, nella maggior parte dei Paesi prevalgono una o due forme genetiche (Figura 7). Ciò suggerisce un effetto fondatore che ha permesso alla variante genetica che ha iniziato a circolare più precocemente e con maggior successo, di stabilirsi come forma predominante nella popolazione, acquisendo un vantaggio sulle varianti arrivate in seguito.

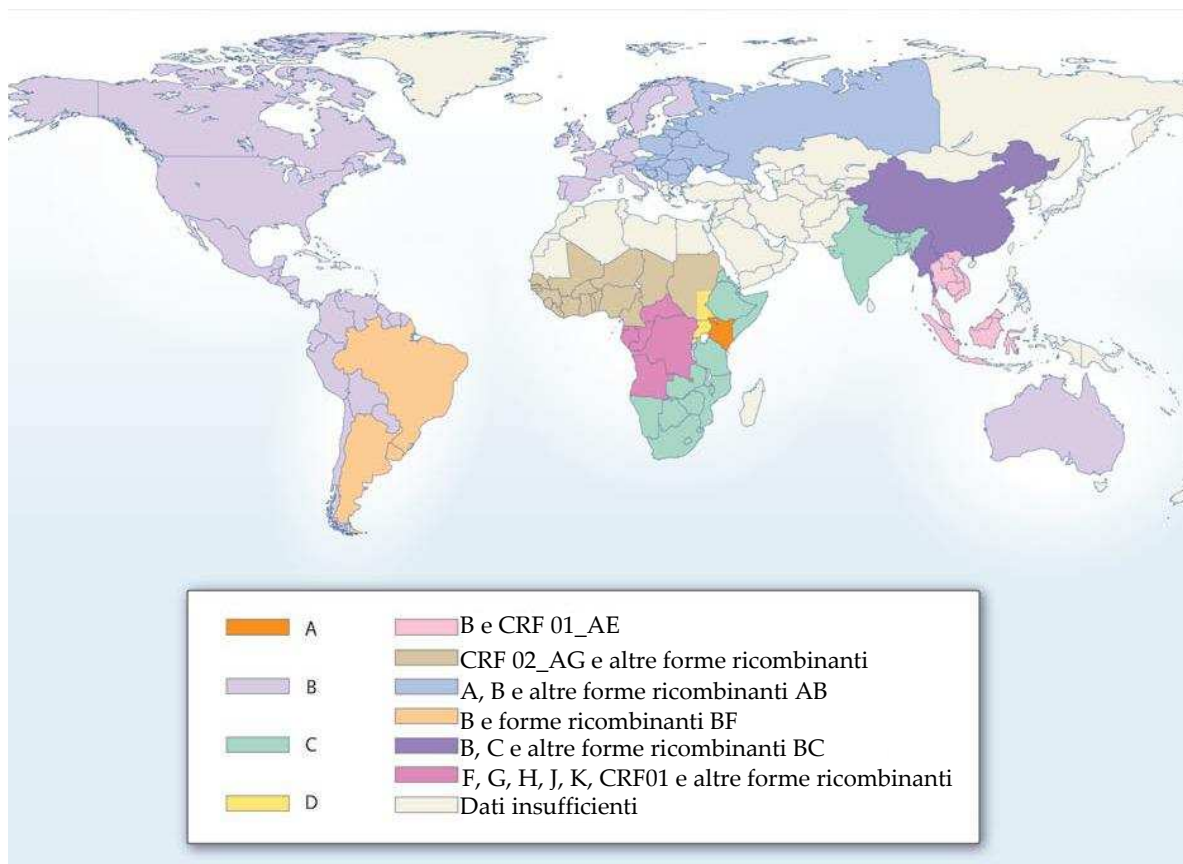


Figura 7. Distribuzione geografica su scala mondiale di sottotipi e CRF di HIV-1.

In accordo con studi più recenti, i ceppi di HIV-1 con maggior prevalenza su scala globale sono i sottotipi C, A (sotto-sottotipo A1), B e la CRF02\_AG.

La maggior parte delle infezioni nei Paesi dell'Africa del Sud, in India e in Etiopia sono sostenute dal sottotipo C, che circola anche in Brasile e in Russia come ceppo minore. Virus ricombinanti con porzioni di C sono comuni in Cina, dove la CRF07\_BC e la CRF08\_BC sono prevalenti tra i tossicodipendenti, e in Tanzania.

Il sottotipo A circola ampiamente nelle aree dell'Africa Centrale e orientale (Kenya, Uganda, Tanzania e Ruanda), nei Paesi europei appartenenti all'ex Unione Sovietica, dove è presente soprattutto tra i tossicodipendenti, e non è rara nell'Africa subsahariana.

In tutta l'Africa Occidentale e in parte di quella Centrale, la forma genetica prevalente è la CRF02\_AG.

Il ceppo B è il principale sottotipo circolante nell'Europa Occidentale e Centrale, negli Stati Uniti e in Australia, ed è comune anche in numerosi Paesi del Sud-Est asiatico, dell'Africa settentrionale, del Medio Oriente e tra gli omosessuali sudafricani e russi.

La forma virale ricombinate CRF01\_AE, ancestrale nell'Africa Centrale, circola estesamente nel Sud-Est asiatico; originaria della Thailandia dove è divenuta prevalente si è estesa anche ai Paesi vicini, tra cui Cina, Giappone e Corea.

Tra le altre forme genetiche, meno prevalenti a livello globale ma comuni su scala più ristretta, vi è il sottotipo D, distribuito soprattutto nell'Africa dell'Est (Uganda, Tanzania e Kenya) e la CRF03\_AB, descritta inizialmente nella città russa di Kaliningrado<sup>51,52</sup> e successivamente ritrovata in numerose città dell'ex Unione Sovietica (San Pietroburgo, Smolensk e Perm).

Un capitolo a parte merita il sottotipo F (sotto-sottotipo F1), predominante in Romania, soprattutto tra bambini ospedalizzati infettatisi attraverso prodotti emoderivati e aghi o siringhe non sterilizzati; cocircola ad alta prevalenza anche in Brasile dove avrebbe originato le numerose CRF BF e rappresenta anche in Italia il sottotipo non-B più diffuso insieme alla CRF02.

Il sottotipo G, che circola nell'Africa Occidentale e Centrale, ha la più alta prevalenza in

Nigeria, ma è diffuso anche in Portogallo e nel Nord-Ovest della Spagna ove ha originato le CRF14.

Gli altri sottotipi non ricombinanti (A2, F2, H, J e K) circolano come forme minori nell'Africa Centrale, mentre le rimanenti CRF hanno una rilevanza epidemiologica minore su scala globale.

In numerosi Paesi dell'Europa Occidentale è stata osservata una percentuale significativa di infezioni sostenute da sottotipi non-B, generalmente trasmesse per via eterosessuale. Molte di queste infezioni sono correlate in maniera diretta a fonti non europee (soprattutto dell'Africa Sub-Sahariana)<sup>53</sup>, ad eccezione del sottotipo G e della CRF14\_BG isolate nelle regioni occidentali della penisola iberica, dove circolano localmente.

Oltre alle CRF, si è osservato un notevole aumento di virus con strutture mosaiche uniche (*Unique Recombinant Form* o URF), rilevati in un unico individuo o in poche persone epidemiologicamente correlate.

In alcune aree è stata riportata un'alta densità di ricombinanti unici (Repubblica Democratica del Congo, Tanzania<sup>54</sup>, Argentina, Cuba e Galizia) tanto che alcune URF sembrano aver preso origine da ricombinazioni secondarie di CRF già note.

In alcuni casi come in Argentina i ricombinanti BF di seconda generazione, derivanti cioè dalla ricombinazione della CRF12\_BF e del sottotipo B, hanno probabilmente una prevalenza maggiore della forma ricombinante originaria. Anche in Spagna sono stati individuati virus ricombinanti di tipo BG che derivano dalla ricombinazione della CRF14\_BG con virus del sottotipo B.

Questi dati suggeriscono la possibilità che possa avvenire una reinfezione con un sottotipo diverso e che questa possibilità possa essere un evento non raro, almeno in alcuni luoghi o con alcuni sottotipi.

## IMPLICAZIONI BIOLOGICHE DELLA VARIABILITÀ DI HIV-1

---

La variabilità è l'arma più potente di HIV, che gli permette di eludere le difese immunitarie dell'ospite e gli effetti di interventi terapeutici (farmaci) e profilattici (vaccini)<sup>55</sup>.

La variabilità di HIV-1 ha due importanti implicazioni a livello biologico:

- la distribuzione nel mondo di distinti sottotipi di HIV-1 ha un impatto importante sull'epidemiologia, la patogenesi, la diagnosi, il trattamento e la prevenzione dell'infezione. Da tempo è stato infatti ipotizzato che la presenza di sottotipi diversi in specifiche aree abbia implicazioni importanti per la patogenesi. La relazione tra sottotipo e decorso della malattia rimane inoltre in larga parte da chiarire.
- i farmaci antiretrovirali (*AntiRetroVirals*, ARV), soprattutto se utilizzati a concentrazioni sub-ottimali, selezionano mutazioni associate a resistenza. Con l'utilizzo sempre più diffuso di questi farmaci, la selezione di ceppi virali resistenti compare in una percentuale significativa degli individui trattati, comportando seri problemi per la loro gestione clinica. Questo rappresenta oggi un serio problema di sanità pubblica, soprattutto nelle aree del mondo dove i farmaci antiretrovirali sono stati largamente utilizzati per molti anni.

## IDENTIFICAZIONE E CIRCOLAZIONE DEI SOTTOTIPI NON-B DI HIV-1

---

La designazione del sottotipo di HIV-1 ha rappresentato un formidabile marcatore molecolare epidemiologico per seguire sia l'evoluzione globale della pandemia da HIV-1, sia l'ingresso e la circolazione di sottotipi diversi da quelli che hanno dato origine all'epidemia nei paesi economicamente avanzati, a seguito dei flussi migratori e degli spostamenti dell'uomo dettati dagli scambi commerciali e dal turismo.

La complessità genetica di HIV-1 ha un impatto importante sull'epidemiologia, la patogenesi, la diagnosi, il trattamento e la prevenzione dell'infezione<sup>56,57,58,59,60,61,62</sup>. Per questo è molto importante studiare e seguire l'evolversi della complessità virale.

In particolare, l'alta variabilità di HIV-1 ha sollevato rilevanti problematiche relative al trattamento dei soggetti HIV positivi. Tra queste, particolare attenzione è stata posta sulla specificità delle mutazioni coinvolte nella resistenza ai farmaci antiretrovirali dei diversi sottotipi di HIV-1. Infatti, inizialmente, i test di resistenza genotipica sono stati sviluppati e condotti su isolati di sottotipo B, in quanto questo era il sottotipo prevalente nei paesi (USA, Europa e Australia) che hanno inizialmente disegnato, testato, e sintetizzato gli agenti anti-HIV-1 e condotto i *trial* clinici. Inoltre, essendo i costi dei farmaci così elevati, il loro uso è stato limitato nei Paesi sottosviluppati dove gli altri sottotipi virali sono prevalenti.

I dati pubblicati riguardanti i sottotipi non-B non sono numerosi ed esaustivi, ma suggeriscono che la loro suscettibilità ai farmaci antiretrovirali può essere differente dalle varianti di sottotipo B. Ad esempio, le varianti di tipo G hanno una moderata suscettibilità ai PI e ceppi di sottotipo F hanno una diminuita sensibilità agli NNRTI<sup>63,64</sup>. Numerosi studi condotti fino ad oggi hanno dimostrato che le sequenze del gene *pol* di HIV-1, generate dai saggi di resistenza, sono affidabili per l'assegnazione del sottotipo virale. Tali studi di tipizzazione virale possono rispondere a molte domande, quali la frequenza di trasmissione di mutazioni associate a resistenza nei ceppi non-B,

l'associazione tra resistenza naturale e polimorfismi e la selezione di mutazioni primarie in corso di terapia antiretrovirale.

Questi studi forniscono, inoltre, informazioni importanti sui profili di resistenza antiretrovirale e sulla loro interpretazione in relazione al sottotipo virale e potrebbero essere utili per indagare il comportamento biologico degli isolati virali in relazione alla patogenicità e trasmissione.

La conoscenza della prevalenza delle distinte varianti di HIV-1 circolanti in specifiche aree è quindi rilevante nella pratica clinica per il *management* del paziente e per assicurare la validità dei *test* diagnostici e dei *marker* di infezione eseguiti in laboratorio durante il *follow-up* clinico.

## MODIFICAZIONI DELL'EPIDEMIA DA HIV-1 IN ITALIA

---

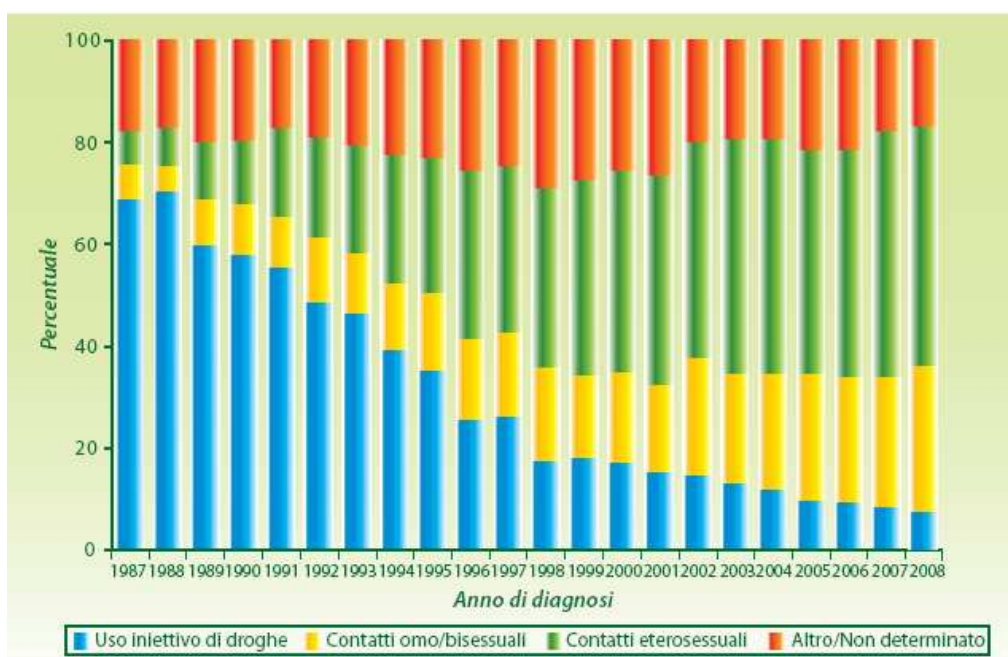
Malgrado il Ministero del Lavoro, della Salute e delle Politiche Sociali, nel 2008 abbia promosso l'attivazione del Sistema di sorveglianza delle nuove diagnosi di infezione da HIV a livello nazionale, provvedendo ad aggiungere l'infezione da HIV all'elenco della Classe III delle malattie infettive sottoposte a notifica obbligatoria, i dati di cui si dispone al 2009 provengono solo da alcune 6 regioni (Lazio, Veneto, Friuli-Venezia Giulia, Piemonte, Liguria e Puglia) e 6 province (Modena, Trento, Bolzano, Sassari, Rimini e Catania), le quali in modo autonomo e hanno raccolto informazioni sulle nuove diagnosi di infezioni da HIV, inviandole periodicamente al Centro Operativo Aids (COA) dell'Istituto Superiore di Sanità. I casi segnalati da queste aree non rappresentano tutti i casi di nuove infezioni da HIV che si verificano in Italia ma, in attesa di dati nazionali, possono fornire un utile indicazione sulla diffusione e sull'andamento temporale dell'infezione da HIV nel nostro Paese<sup>65</sup>.

Nel periodo 1985-2008 sono state riportate 42.747 nuove diagnosi di infezione (30.297 maschi, 12.346 femmine, 14 genere non noto). L'incidenza delle nuove diagnosi ha avuto un picco di segnalazioni nel 1987, per poi diminuire fino al 1998;

successivamente il numero di nuove infezioni si è stabilizzato, anche se negli ultimi anni appare in aumento in alcune zone; questo andamento potrebbe preludere a una possibile riattivazione dell'epidemia in varie aree del nostro Paese. Nel 2008 sono state segnalate 2.012 nuove diagnosi, pari a una incidenza di 6.0 per 100.000 residenti (l'incidenza più bassa è stata osservata in Puglia, quella più alta nella provincia di Rimini). Negli anni si è rilevato un progressivo aumento dei soggetti di sesso femminile, tanto che la proporzione maschi/femmine, che era del 3.5 nel 1985, è diventata 2.9 nel 2008.

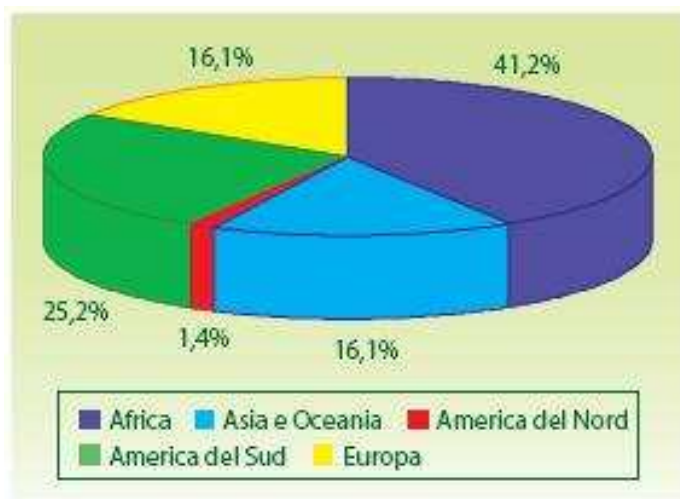
L'età mediana al momento della diagnosi di infezione è aumentata, da 26 anni per i maschi e 24 per le femmine nel 1985 a, rispettivamente 38 e 35 anni nel 2008.

Col passare degli anni si è manifestato un importante cambiamento delle modalità di trasmissione: la proporzione di tossicodipendenti è diminuita dal 74.6% nel 1985 all'7.7% nel 2008, mentre i casi attribuibili a trasmissione sessuale (eterosessuale ed omosessuale) sono aumentati nello stesso periodo dal 7.8% al 75.7%. Per una parte delle persone con nuova diagnosi di infezione, oscillante nei vari anni tra il 20 ed il 30%, non è stato possibile stabilire la modalità di trasmissione; tale percentuale nel 2008 era del 16,6% (Figura 8).



**Figura 8. Distribuzione percentuale delle nuove diagnosi di infezione da HIV-1 per modalità di trasmissione.**

La proporzione di stranieri tra le nuove diagnosi è aumentata dall'11% nel 1992 al 32% nel 2008. La maggior parte degli stranieri proviene dall'Africa (41,2%), America Latina (25,2%), mentre il 16,1% proviene da altri paesi (Figura 9).



**Figura 9. Distribuzione percentuale delle diagnosi di infezione da HIV negli stranieri per area di provenienza.**

Le informazioni raccolte dai sistemi di sorveglianza dimostrano che le caratteristiche di coloro che oggi si infettano con HIV sono radicalmente diverse da quelle di coloro che si infettavano dieci o venti anni fa:

- non si tratta più di persone giovani e prevalentemente tossicodipendenti ma piuttosto di adulti maturi (36-39 anni per i uomini e 33-35 per le donne) che si infettano attraverso rapporti sessuali;
- la maggior parte delle donne acquisisce l'infezione da un partner che sapeva di essere HIV-positivo;
- nonostante il test HIV venga raccomandato a tutte le donne in gravidanza, nel 2008 sono stati segnalati 6 casi di trasmissione da madre a figlio;
- negli anni è aumentata progressivamente la quota di persone che scopre di essere infetta solo in fase avanzata di malattia e che costituiscono quindi, a loro insaputa, una possibile fonte di diffusione dell'infezione;

- si stima che circa un quarto delle persone HIV-positive presenti in Italia non sappia di essere infetta;
- più della metà dei soggetti con una nuova diagnosi di AIDS ignora la propria sieropositività, in modo particolare coloro che hanno acquisito l'infezione attraverso i rapporti sessuali, sia eterosessuali sia omosessuali;
- la proporzione di stranieri con una nuova diagnosi di infezione aumenta costantemente;

L'incidenza di HIV appare particolarmente elevata in alcune province (Rimini e Trento), ma questi dati possono solo essere proiettati nell'ambito di un quadro nazionale non è ancora disponibile. Come riportato, l'incidenza nel 2008 delle nuove diagnosi di infezione da HIV, calcolata nelle aree dove è attivo un sistema di sorveglianza, è stata del 6,0 per 100.000 residenti. Se tale dato fosse generalizzabile a tutto il territorio nazionale, l'Italia si collocherebbe a un livello medio di incidenza rispetto agli altri Paesi europei, con valori simili alla maggior parte delle nazioni dell'Europa Occidentale.

Le stime del numero di soggetti con infezione da HIV o AIDS effettuate finora in Italia e in altri Paesi si basavano principalmente su metodi di *back-calculation* o su estrapolazioni da dati regionali. Recentemente, l'UNAIDS (Joint United Nations against HIV/AIDS ) ha elaborato un metodo nuovo che si basa sulle prevalenze di HIV osservate in vari sottogruppi di popolazione ad alto rischio di infezione; questo metodo stabilendo dei criteri precisi e standardizzati relativamente alle informazioni da utilizzare, ha il vantaggio di consentire un confronto dei risultati ottenuti in nazioni diverse.

La stima per l'Italia è di 165.000 persone viventi con HIV/AIDS (prevalenza sulla popolazione totale di 2,7 per mille abitanti) di cui circa l'82% sono infezioni acquisite per via sessuale e le donne rappresentano il 30% degli adulti.

Le nuove infezioni HIV diagnosticate nel 2008 sono state 3.900-4.100 la principale modalità di trasmissione era rappresentata dai contatti etero/omosessuali (74%), l'età mediana al primo test HIV positive era di 39 anni e la percentuale di stranieri era il 32%.

## 2 SCOPO DELLA RICERCA

---

Durante la pandemia dell'infezione da HIV si sono differenziati 9 differenti sottotipi del gruppo M (A–D, F–H, J and K) a causa della rapidità della replicazione virale, dell'altissima frequenza di errore della trascrittasi inversa (RT) e della pressione selettiva esercitata dal sistema immunitario. L'elevata capacità di ricombinazione dell'RT di HIV-1 ha ulteriormente aumentato la differenziazione di HIV-1 a livello globale con le 43 forme ricombinanti circolanti (CRF) che sono già state caratterizzate e con le numerose forme uniche ricombinanti (URF) che sono state identificate globalmente<sup>66, 67, 68</sup>. La maggior parte dei sottotipi e delle CRF erano inizialmente circoscritte in determinate aree geografiche o popolazioni, ma la loro distribuzione è in continua evoluzione<sup>69</sup>.

Per monitorare l'evoluzione della pandemia è utile l'assegnazione del sottotipo, che permette di valutare l'andamento epidemiologico locale, influenzato dai cambiamenti sociali e dai flussi migratori. Sulla base dei dati a disposizione il sottotipo B è stato il ceppo 'fondatore' dell'epidemia nell'Europa Occidentale, Stati Uniti, in Canada e in Australia, dove è stato il sottotipo maggiormente rappresentato per circa 2 decenni<sup>70</sup>. Tuttavia, negli ultimi anni numerosi studi hanno dimostrato la circolazione di ceppi non-B in aree precedentemente ristrette al sottotipo B<sup>71, 72, 73, 74, 75, 76, 77, 78</sup>. Recentemente l'andamento dell'epidemia da HIV nei Paesi dell'Europa Occidentale, dove si è verificato un ampio flusso immigratorio, è stato caratterizzato da importante incremento dei ceppi non-B tra i soggetti con nuova diagnosi di infezione<sup>79, 80, 81, 82</sup>. Numerosi lavori hanno messo in evidenza che la maggior parte delle infezioni con sottotipi non-B in Europa è legata all'immigrazione da aree ad alta prevalenza di HIV-1 o a viaggi in queste regioni, soprattutto l'Africa Sub-sahariana, il Sud America dove predominano i ceppi non-B. L'immigrazione dal Sud-Est Asiatico, dove cocircolano i ceppi B, C, e la CRF01\_AE, ha riguardato principalmente l'Inghilterra.

La diffusione dei sottotipi non-B in Italia è avvenuta in un arco temporale (anni novanta e duemila) in cui fattori epidemiologici come l'etnia, la modalità di

infezione e il sesso dei soggetti con infezione da HIV subivano profondi cambiamenti. La trasmissione sessuale è diventata la principale modalità di infezione, mentre le nuove infezioni tra i tossicodipendenti sono drasticamente diminuite. Inoltre la trasmissione attraverso i rapporti sessuali è molto aumentata tra gli eterosessuali rispetto agli omosessuali. Di conseguenza la prevalenza della percentuale dei maschi rispetto alle femmine è diminuita nel tempo<sup>83</sup>. Queste considerazioni sono derivate dai dati del Centro Operativo AIDS (COA) dell'Istituto Superiore di Sanità, che indicano che in Italia vivono circa 150.000 individui con infezione da HIV-1. I sistemi di sorveglianza delle nuove diagnosi hanno recentemente mostrato che:

1. i pazienti stranieri (provenienti per il 41.2% dall'Africa, per il 25.2% dal Sud America e per il 16.1% dall'Europa) sono aumentati dall'11% al 32% nel periodo 1992-2007,
2. la proporzione di tossicodipendenti è diminuita dal 69% all'8.6%,
3. la trasmissione sessuale è aumentata dal 13.3% al 73.7%, in particolare la più frequente via di trasmissione è rappresentata dai contatti eterosessuali che sono passati dal 24.6% al 75.9%,
4. il rapporto maschi femmine è diminuito dal 3.5 al 2.5.

Attualmente non sono disponibili stime ufficiali della frequenza di trasmissione dei sottotipi non-B, anche se i dati a disposizione suggeriscono che la trasmissione dell'infezione, sostenuta da questi ceppi sia da attribuire a soggetti di etnia non caucasica. Le informazioni sull'origine delle infezioni da ceppi non-B sono limitate a causa della scarsità e della non accuratezza dei dati.

L'epidemiologia molecolare, studiata e definita attraverso accurati strumenti filogenetici, può contribuire all'acquisizione di importanti informazioni sull'epidemiologia attuale dell'infezione da HIV-1 nel nostro paese i) rilevando la presenza di nuovi sottotipi nella popolazione, ii) indicando l'origine dell'infezione,

iii) evidenziano gli eventi di trasmissione tra le diverse etnie, i gruppi a rischi e i generi, iv) datando i tempi dell'ingresso e della diffusione dei diversi sottotipi fornendo in questo modo gli strumenti per monitorare l'andamento dell'infezione tra i diversi gruppi a rischio<sup>84, 85</sup>.

Sulla base di queste considerazioni, gli obiettivi di questo studio erano:

1. valutare la prevalenza e la distribuzione dei sottotipi non-B in Italia in un'ampia coorte di soggetti con infezione da HIV-1 i cui dati di sequenza sono stati generati in un laboratorio referenziato.
2. studiare l'andamento temporale della circolazione dei ceppi non-B e valutare l'associazione tra le infezioni sostenute da sottotipi non-B e le principali variabili demografiche nel periodo 1980-2008.
3. Valutare l'andamento della diffusione dei sottotipi non-B in Italia in relazione all'etnia, alla modalità di infezione e al sesso.

### 3 MATERIALI E METODI

---

## CASISTICA

---

Abbiamo studiato complessivamente 3670 soggetti, seguiti presso 50 Centri Clinici in 13 Regioni italiane nel periodo tra il 1980 e il 2008. I pazienti hanno effettuato un test di resistenza genotipica al momento della diagnosi di infezione da HIV o prima dell'inizio della terapia antiretrovirale o al fallimento virologico. Tutti i test sono stati eseguiti presso il Servizio di Monitoraggio dell'infezione da HIV del Dipartimento di Biologia Molecolare dell'Università di Siena. I pazienti sono stati arruolati nel database dell'Antiretroviral Resistance Cohort Analysis ([www.hivarca.net](http://www.hivarca.net)) dopo aver firmato il consenso informato per l'inserimento dei dati personali in forma anonima. E' stato valutato il primo genotipo disponibile di ciascun paziente incluso nell'analisi. Per 2479 soggetti prelevati nel periodo in studio (1980-2008) era nota la data di diagnosi di HIV, stabilita come la prima sierologia per HIV positiva. Questo gruppo di pazienti è stato denominato HD ('HIV diagnosis').

Le informazioni demografiche (sesso, fattori di rischio, nazionalità, data della diagnosi ed età) sono state raccolte dai medici curanti e inserite in un database insieme ai dati virologici, immunologici, terapeutici e clinici. Complessivamente le informazioni sul Paese di origine, sul sesso, sul fattore di rischio e sulla data di diagnosi di HIV erano disponibili per 1364 individui arruolati nel periodo 1980-2008. L'analisi di correlazione sono state effettuate su questi ultimi pazienti, denominati come gruppo CD ('Complete Data').

## METODI

---

Le sequenze FASTA per l'analisi filogenetica, risultato dei test di resistenza, sono state ottenute impiegando differenti Kit commerciali e saggi *homemade*. ARCA richiede che i laboratori che contribuiscono con il risultato dei test di resistenza abbiano ottenuto la certificazione mediante controlli di qualità.

## ANALISI FILOGENETICA

---

L'analisi filogenetica è volta a studiare le distanze evolutive esistenti tra le sequenze in esame, tali distanze sono rappresentate attraverso alberi filogenetici, che costituiscono dei grafici bidimensionali composti da nodi connessi tra loro tramite rami.

I nodi terminali rappresentano i *taxa* attuali, mentre i nodi interni rappresentano quelli ancestrali.

Le relazioni filogenetiche vengono definite attraverso la topologia dell'albero: la distribuzione dei nodi e la lunghezza dei rami forniscono indicazioni sulla divergenza evolutiva esistente tra i diversi *taxa*, permettendo di individuare gruppi monofiletici e linee evolutive.

L'analisi filogenetica rappresenta, oggi, l'unico metodo affidabile per la tipizzazione delle sequenze, siano esse sottotipi puri o ricombinanti; permette inoltre di esaminare le relazioni genetiche all'interno di un gruppo di sequenze o tra virus provenienti da diverse aree geografiche, fornendo indicazioni epidemiologiche sull'origine e la propagazione dei virus in una popolazione.

Al fine della conduzione di una analisi filogenetica per ceppi di HIV-1 possono essere utilizzate indifferentemente sequenze ottenute dopo estrazione da DNA associato alle cellule (DNA provirale) sia quelle ottenute da virus libero circolante nel plasma dopo RT-PCR.

## DISCRIMINAZIONE TRA SOTTOTIPO B E NON-B MEDIANTE STRUMENTI DISPONIBILI IN RETE

---

Inizialmente ogni sequenza viene sottoposta ad una iniziale genotipizzazione, utilizzando uno strumento disponibile *on line*, il *Rega Institute subtyping tool* versione 2.0 (<http://www.bioafrica.net/virus-genotype/html/subtyping.html>).

Questo sito rispetto ad altri disponibili in rete utilizza l'analisi filogenetica per l'analisi delle sequenze e non solo metodi basati sul confronto e la similarità.

In dettaglio, ogni singola sequenza in formato FASTA (un file di testo formato da una prima riga descrittiva identificata dal carattere iniziale ">", seguita dalla sequenze nucleotidica a partire dalla seconda riga) viene inserita nella finestra di immissione della pagina *web* e sottoposta ad analisi filogenetica; è possibile inoltre visualizzare l'albero generato ed l'immagine ottenuta dall'analisi di *bootscan*.

Nonostante questo sito venga ormai utilizzato per la sottotipizzazione in numerosi articoli, bisogna ricordare che non è possibile pubblicare nel proprio lavoro l'albero ottenuto con questo sito e che l'analisi non tiene conto delle caratteristiche delle proprie sequenze poiché utilizza un *dataset* fisso di sequenze di riferimento. Non sempre inoltre l'assegnazione è certa e supportata.

## ANALISI FILOGENETICA DEI SOTTOTIPI PURI E DELLE CRF

---

Utilizzando il programma BioEdit (ver. 7.0.9.1)<sup>86</sup> le sequenze della regione PR e RT dei campioni vengono allineate tra loro e con un numero variabile di sequenze di riferimento disponibili nel database del LANL<sup>87</sup>, utilizzando, solitamente, da due a cinque ceppi di riferimento per ogni sottotipo e CRF.

Le estremità vengono eliminate al fine di ottenere frammenti della medesima

lunghezza e l'allineamento viene effettuato utilizzando l'algoritmo *Clustal* contenuto nel programma, intervenendo con un *editing* manuale quando necessario. In questo modo vengono costruiti i *dataset* di sequenze utilizzati per le successive analisi di tipizzazione e di ricombinazione.

Ogni *dataset* viene analizzato utilizzando i programmi SEQBOOT, DNADIST, NEIGHBOR e CONSENSE contenuti nel pacchetto *software* PHYLIP<sup>88</sup>(ver. 3.66). Il rapporto transizione/transversione viene calcolato per ogni specifico *dataset* con il programma TreePuzzle (ver. 5.2)<sup>89</sup> e utilizzato per calcolare le distanze genetiche tra le sequenze sottomesse mediante opportuni modelli evolutivisti di differente complessità.

Il programma DNADIST calcola le distanze genetiche tra le sequenze di ogni *dataset*, generando una matrice numerica utilizzata dal programma NEIGHBOR per generare l'albero filogenetico reale, in cui le lunghezze dei rami rispecchiano le effettive distanze evolutive tra i *taxa*. Il programma NEIGHBOR può utilizzare algoritmi diversi per i suoi calcoli.

La funzione del programma SEQBOOT è quella di creare dei replicati delle sequenze in esame, in modo da ottenere un supporto statistico dell'affidabilità e attendibilità dell'albero filogenetico generato. Il programma produce una serie di *dataset* replicati, composti dallo stesso numero di sequenze dell'originale ma in cui le colonne sono una combinazione con ripetizione di quelle iniziali. In questo modo è possibile valutare se l'albero reale generato, rimane invariato anche considerando solo parti delle sequenze in studio.

I replicati sono sottoposti alle stesse analisi del *dataset* iniziale e i dati generati vengono valutati complessivamente dal programma CONSENSE, il quale crea infine un albero filogenetico che rappresenta il consenso di tutti quelli replicati. In quest'albero, sono indicati dei valori, definiti di *bootstrap*, che rappresentano la frequenza con cui ogni diramazione è presente, esattamente nella stessa topologia, in tutti gli alberi replicati. Valori di *bootstrap* uguali o superiori al 70% sono considerati statisticamente significativi e convenzionalmente sono ritenuti indicativi di una ramificazione

attendibile; questo valore soglia è tuttavia arbitrario e può variare a seconda della complessità e numerosità delle sequenze in esame.

## ANALISI DEI CEPPI RICOMBINANTI NON TIPIZZABILI COME CRF MEDIANTE IL PROGRAMMA SIMPLOT

---

Le sequenze che all'analisi filogenetica non risultano tipizzabili, vengono ulteriormente esaminate per cercare eventuali ricombinazioni tra ceppi noti.

A tale scopo viene utilizzato il programma SimPlot (ver. 3.5.1)<sup>90</sup>, che si avvale del metodo per la ricerca delle ricombinazioni genomiche e permette di studiare sequenze con mosaicismi unici.

Nell'analisi di ricombinazione vengono utilizzate unicamente le sequenze di riferimento dei sottotipi puri inclusi nel *dataset* in studio, escludendo le CRF.

Il programma utilizza una finestra scorrevole che si sposta progressivamente lungo l'intera sequenza e genera due distinti *output* grafici: il primo, detto *similarity plot*, è rappresentato da una serie di linee che riflettono la similarità di ogni sequenza (o gruppo di sequenze) di riferimento rispetto a quella da caratterizzare. Il secondo è generato mediante la cosiddetta analisi di *bootscan*, che permette di calcolare e mettere in grafico la percentuale di alberi permutati, sempre utilizzando una finestra scorrevole lungo l'intera sequenza.

In questa tesi, per l'analisi di *bootscan*, abbiamo utilizzato prevalentemente una finestra di 300 bp e 1000 replicati per il confronto con le sequenze di riferimento.

La distanza genetica esistente tra la sequenza esaminata e quelle di riferimento, calcolata in 1000 replicati, è stata valutata utilizzando una matrice di distanza Kimura a 2 parametri, elaborata con il supporto del programma DNADIST e utilizzando l'algoritmo *Neighbor-Joining* per generare gli alberi corrispondenti.

Il processo si ripete per slittamento progressivo della finestra di 20 bp in direzione 5'-

3'.

In questo modo viene analizzata l'intera sequenza nucleotidica in esame. I risultati sono presentati sotto forma di grafico a linee, che riporta i valori di omologia con i diversi sottotipi lungo tutta la sequenza. Sono necessari livelli di omologia superiori al 50% per individuare i sottotipi parentali coinvolti nel mosaicismo della sequenza.

I punti di ricombinazione vengono individuati sia visivamente, osservando l'allineamento, sia ripetendo l'analisi di *bootscan* della sequenza studiata rispetto ad almeno tre sottotipi di riferimento, tra i quali vanno inseriti i ceppi che alla prima analisi risultavano i probabili ricombinanti.

Sulla base dei punti di ricombinazione, le sequenze vengono tagliate e suddivise in frammenti appartenenti ad un solo sottotipo parentale; questi sono sottoposti ad un'ulteriore analisi filogenetica per confermare e discriminare il *pattern* di ricombinazione.

#### CONTRIBUTO ALLA CARATTERIZZAZIONE E TIPIZZAZIONE DEI CEPPI RICOMBINANTI TRAMITE LA RICERCA BLAST

---

Per definire ulteriormente i risultati ottenuti dall'analisi dei ceppi ricombinanti unici e per indagare e approfondire le connessioni epidemiologiche o le eventuali relazioni geografiche, è possibile avvalersi di un altro strumento di supporto costituito dalla ricerca BLAST<sup>91</sup> nelle banche dati.

La sequenza da esaminare viene copiata nell'apposita finestra di immissione che si trova sulla pagina *Web* della banca dati mondiale; il *server* database provvede quindi a confrontarla con tutte le sequenze di HIV sottomesse e fornisce un elenco di quelle con maggiore similarità.

Alle sequenze inserite nella banca dati sono associate schede informative che possono includere note riguardanti la provenienza, l'eventuale caratterizzazione molecolare o altri studi ad esse correlati, patogenetici, farmacologici, immunologici o di altro tipo, la

cui consultazione può contribuire alla caratterizzazione della sequenza in esame.

## ANALISI STATISTICA

---

Per confrontare l'età media, i valori di HIV-RNA e il numero di CD4 nei pazienti con sottotipo B rispetto a quelli non-B è stato utilizzato il test di Wilcoxon. Le variabili categoriche di questi 2 gruppi sono state analizzate mediante il test del chi-quadrato o il test esatto di Fisher.

Gli *odds ratio* (OR) sono stati calcolati mediante un modello logistico univariato e multivariato.

Per l'analisi dei trend temporali è stato impiegato il test Cochrane-Armitage. Per tutte le analisi è stato considerato un errore del 5%.

## 4 RISULTATI

---

## CARATTERISTICHE DELLA POPOLAZIONE

---

Il sottotipo di HIV-1 è stato determinato in tutti i pazienti, rilevando una prevalenza dei ceppi non-B dell'11% (417 su 3670 pazienti). Nel 97% ( $n = 3561$ ) dei casi era noto il Continente di origine (92.2% Europa, 4.5% Africa e 3.3% altri Continenti), nel 53.5% ( $n = 1963$ ) la modalità di infezione (35.6% eterosessuali, 32.9% tossicodipendenti, 26.3% omosessuali e 5.2% altra modalità) e nel 98.1% ( $n = 3602$ ) il sesso (70.4% maschi). Complessivamente l'età media dei soggetti era di 38 anni [intervallo interquartile (IQR) 33-44 anni]. La mediana dei CD4 e della carica virale di HIV erano rispettivamente 305 cell/ $\mu$ L (IQR 150-487 cell/ $\mu$ L) e 4.15 log cp/mL (IQR 3.23-4.89 log cp/mL).

Le caratteristiche dei pazienti nel HD subset erano simili. L'etnia era nota nel 99.1% ( $n = 2457$ ) dei soggetti, la modalità di infezione nel 55.1% ( $n = 1365$ ) e il sesso nel 99.2% ( $n = 2461$ ). Il Continente di origine era prevalentemente l'Europa (92.3%), mentre l'Africa rappresentava il 4.6% e gli altri continenti il 3.1% dei pazienti. Per quanto riguarda il fattore di rischio il 35.7% dei pazienti erano tossicodipendenti, il 33.8% eterosessuali il 24.4% omosessuali. Nell'HD subset il 69.3% dei soggetti erano di sesso maschile. Anche l'età media (37 anni, IQR 33-43 anni), la conta dei CD4 (306 cell/ $\mu$ L, IQR 142-488 cell/ $\mu$ L) e la carica virale (4.11 log cp/mL, IQR 3.2-4.9 cp/mL) non erano differenti rispetto a quelle dell'intera popolazione.

Le caratteristiche demografiche e i dati di laboratorio del CD subset, stratificati in base al sottotipo virale, sono mostrati nella Tabella 1.

Tutte le caratteristiche dei pazienti considerate avevano una distribuzione simile nella popolazione generale e nell' HD e CD subset. L'anno di diagnosi dei pazienti era compreso tra il 1980 e il 2006. In 123 soggetti albergava (il 9%) albergava un sottotipo non-B.

	<b>Tutti i pazienti</b>	<b>Pazienti con infezione da sottotipo B</b>	<b>Pazienti con infezione da sottotipo non-B</b>	<i>p</i>
<b>Etnia</b> % (n)				<.0001 <sup>1</sup>
Europea	93.5 (1,275)	97.8 (1,214)	49.6 (61)	
Africana	4.3 (58)	0.2 (3)	44.7 (55)	
Altro <sup>2</sup>	2.3 (31)	2.0 (24)	5.7 (7)	
<b>Fattore di rischio</b> % (n)				<.0001 <sup>1</sup>
Tossicodipendenti	35.8 (488)	39.9 (483)	4.1 (5)	
Eterosessuali	33.7 (460)	29.4 (365)	77.2 (95)	
Omosessuali	24.4 (333)	25.8 (320)	10.6 (13)	
Altro <sup>3</sup>	6.1 (83)	5.8 (73)	8.1 (10)	
<b>Sesso</b> % (n)				<.0001 <sup>1</sup>
Maschi	69.5 (948)	71.2 (884)	52.0 (64)	
Femmine	30.5 (416)	28.8 (357)	48.0 (59)	

<b>Anno di diagnosi</b> % (n)				<.0001 <sup>1</sup>
1980-1992	56.2 (767)	60.3 (748)	15.5 (19)	
1993-2006	43.8 (597)	39.7 (493)	84.5 (104)	
<b>Età (anni), mediana (IQR<sup>4</sup>)</b>	37 (32-42)	37 (33-42)	33 (27-40)	<.0001 <sup>5</sup>
<b>HIV-1 RNA (Log HIV-1 RNA copie/mL), mediana (IQR<sup>4</sup>)</b>	4.1 (3.2-4.8)	4.0 (3.2-4.8)	4.2 (3.2-4.9)	<i>ns</i> <sup>5</sup>
<b>Conta CD4 (cells/<math>\mu</math>L), mediana (IQR<sup>4</sup>)</b>	310 (130-491)	310 (129-486)	324 (197-581)	<i>ns</i> <sup>5</sup>
<b>Totale pazienti,</b> % (n)	100.0 (1,364)	91.0 (1,241)	9.0 (123)	-

<sup>1</sup> Test del chi-quadro

<sup>2</sup> Altro: Latino-Americani, Asiatici

<sup>3</sup> Altro: Trasfusioni, trasmissione verticale, rischio non noto

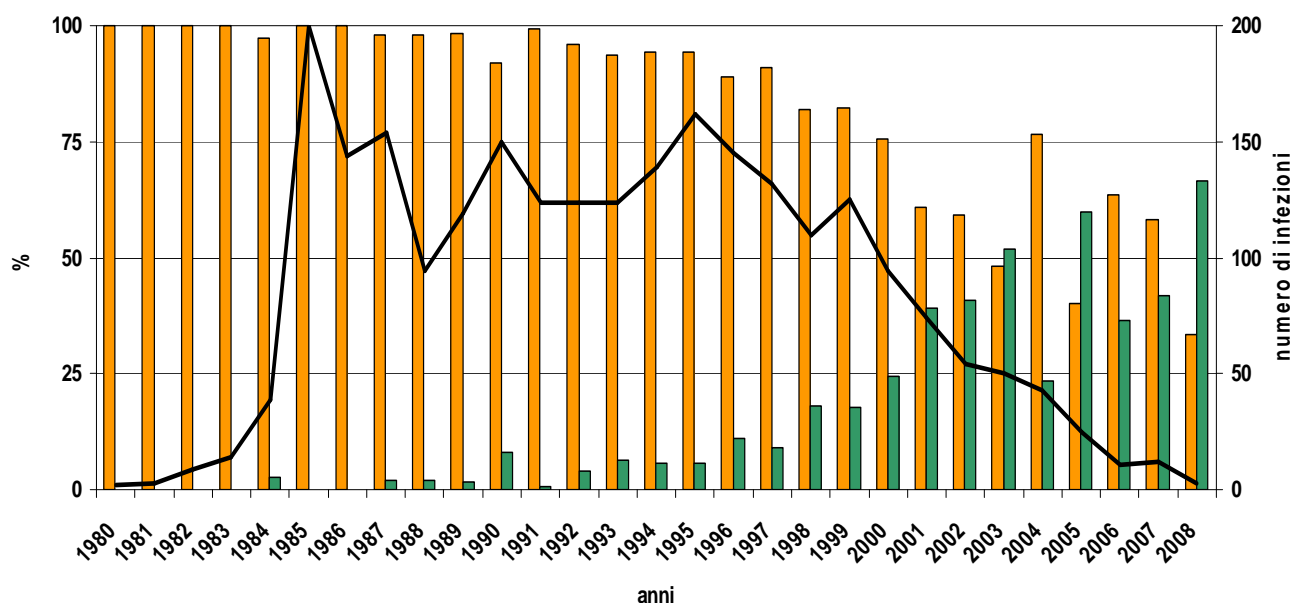
<sup>4</sup> IQR: intervallo interquartile

<sup>5</sup> Test di Wilcoxon

**Tabella 1. Caratteristiche demografiche, immunologiche e virologiche dei 1364 pazienti di cui erano a disposizione i dati completi (CD subset)**

## ANDAMENTO TEMPORALE DEI SOTTOTIPI NON-B NEL PERIODO 1980-2008

Abbiamo valutato la prevalenza nel tempo dell'infezione da HIV-1 con sottotipo B e non-B nei pazienti dell'HD subset che avevano una diagnosi di infezione nel periodo tra il 1980 e il 2008 (Figura 10). Sono stati identificati duecentosettantacinque pazienti (10.4%) che albergavano un ceppo non-B. E' stata riscontrata un'associazione significativa tra l'infezione con sottotipo non-B e l'anno di diagnosi ( $P < 0.0001$ ). Tale associazione era lineare e con un andamento in aumento. L'analisi di regressione, valutando la probabilità di infettarsi con un ceppo non-B anno per anno, supportava questo andamento e indicava che il rischio di infezione con un sottotipo non-B era 1.27 volte più alta per l'anno successivo (95% intervallo di confidenza 1.23-1.31).



**Figura 10. Distribuzione dei pazienti con infezione da HIV con sottotipo B e non-B nell' 'HD subset'. Colonne nere: sottotipo non-B; colonne grigie: sottotipo B. La linea nera indica il numero assoluto di infezioni per anno.**

I primi casi di infezione da sottotipo non-B puro, CRF e URF sono stati identificati in soggetti Africani rispettivamente nel 1984, nel 1990 e nel 1994. Questi pazienti erano giunti in Italia dal Senegal, dal Burkina Faso e dalla Costa d'Avorio e albergavano rispettivamente un sottotipo A1, un ceppo CRF 09\_cpx e un ricombinante CRF02\_AG/A1. Nei primi pazienti di etnia Europea erano presenti un sottotipo puro A1, una CRF01\_AE e una forma ricombinante B/F e hanno ricevuto una diagnosi rispettivamente nel 1987, nel 1996 e nel 1995.

Complessivamente il 52.4% delle nuove diagnosi di HIV-1 risaliva a prima del 1993. Successivamente il numero delle nuove diagnosi è marcatamente diminuito. I sottotipi non-B sono stati riscontrati solo nel 2.6% (34 di 1300) dei soggetti con nuova diagnosi di infezione prima del 1993, ma albergavano ben il 18.9% (223 di 1179) dei pazienti diagnosticati nel periodo 1993-2008 ( $P < 0.0001$ ).

#### DISTRIBUZIONE DEI CEPPI DI HIV-1 NEL TEMPO RISPETTO ALLE VARIABILI DEMOGRAFICHE

---

Abbiamo suddiviso i pazienti del CD subset in due gruppi, quello dei soggetti con diagnosi prima del 1993 e quello dei pazienti con diagnosi dopo il 1993, e abbiamo messo a confronto le caratteristiche demografiche dei due gruppi. I sottotipi non-B rappresentavano il 2.5% (19 su 767) delle nuove diagnosi tra il 1990 e il 1992 e il 17.4% (104 su 597) nel periodo 1993-2006 ( $P < 0.0001$ ). Benchè la percentuale dei pazienti di origine Africana, Latino-Americana e del Sud-Est Asiatico sia significativamente aumentata dall'1.7% ( $n = 14$ ) nel periodo prima del 1993, al 12.6% ( $n = 75$ ) nel periodo successivo ( $P < 0.0001$ ), i ceppi non-B sono marcatamente aumentati tra gli Europei, passando dal 1.9% (13 di 753) nel primo periodo, al 9.2% (48 di 522) nel periodo 1993-2006 ( $P < 0.0001$ ) (Figura 11a).

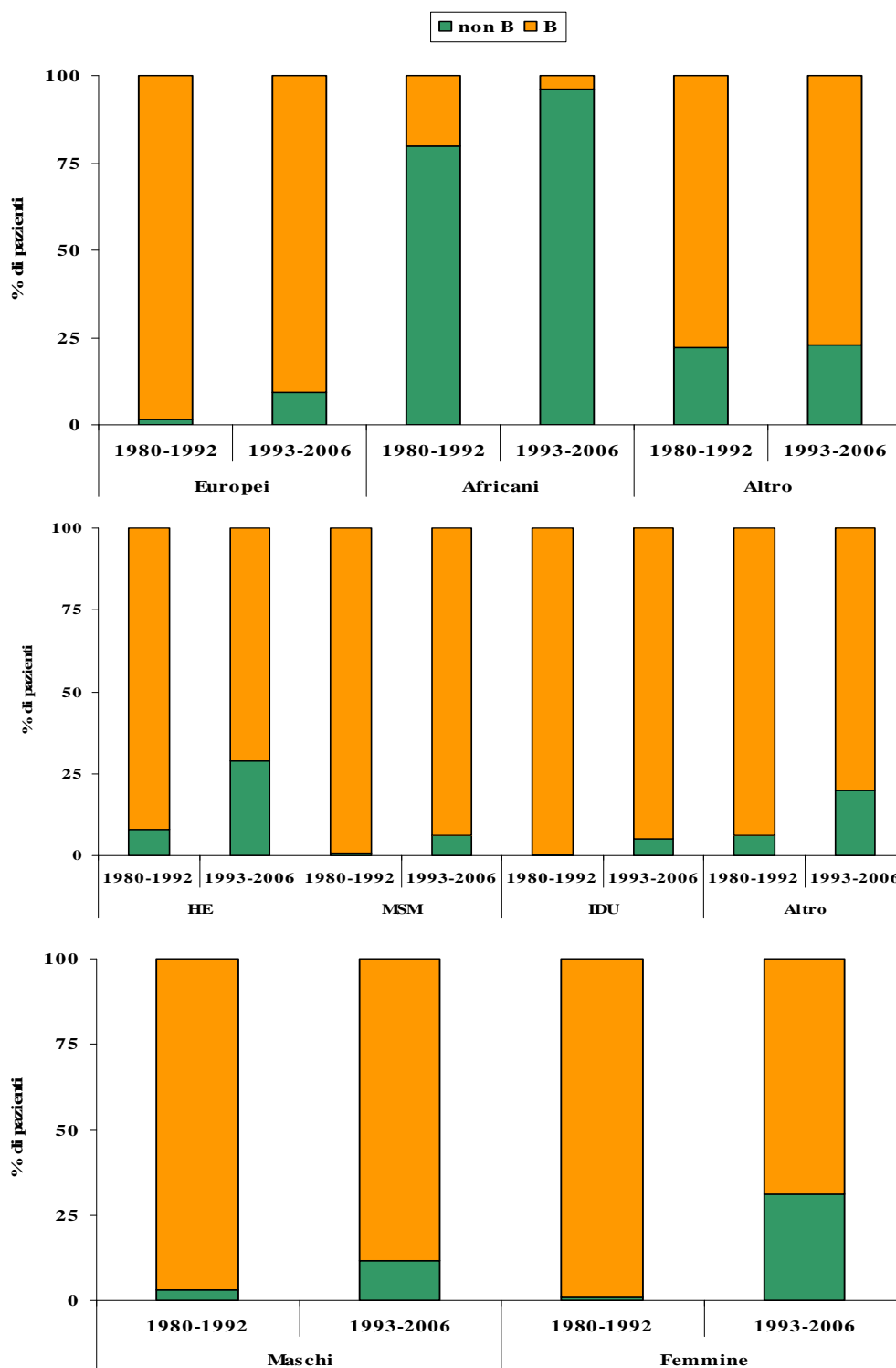


Figura 11. Caratteristiche dei pazienti con infezione da HIV-1 da sottotipo B e non-B nel periodo 1980-1992 e 1993-2006. (a) Etnia, (b) Modalità di infezione e (c) sesso. HE: eterosessuali, IDU: tossicodipendenti, MSM: omosessuali.

Complessivamente la percentuale degli eterosessuali è aumentata dal 23.5% ( $n = 180$ ) nel primo periodo al 46.9% ( $n = 280$ ) nel periodo successivo ( $P < 0.0001$ ), così come per gli omosessuali che sono passati dal 17.3% ( $n = 133$ ) al 33.5% ( $n = 200$ ) ( $P < 0.0001$ ). I tossicodipendenti, al contrario, sono diminuiti dal 52.9% ( $n = 406$ ) al 13.7% ( $n = 82$ ) ( $P < 0.0001$ ). La percentuale degli eterosessuali con infezione da variante non-B nei due gruppi divisi per periodo di studio è aumentata dal 7.8% (14 su 180) al 28.9% (81 su 280) ( $P < 0.0001$ ). Anche nei tossicodipendenti e negli omosessuali è stato rilevato un aumento della prevalenza dei sottotipi non-B passando dal 0.2% (uno di 406) al 4.9% (quattro di 82) ( $P = 0.0003$ ) nei primi e dallo 0.8% (1 su 133) al 6% (12 su 200) nei secondi ( $P = 0.018$ ) (Figura 11b).

Il rapporto uomo-donna non era differente tra i due periodi considerati [rispettivamente il 30.7% ( $n = 236$ ) e 30.2% ( $n = 180$ ) di femmine]. E' stato registrato un numero sproporzionato di donne tra i tossicodipendenti in entrambi i periodi considerati. Tuttavia il numero di pazienti di sesso femminile che albergava ceppi non-B è cresciuta dal 1.3% (3 su 236) nel periodo fino al 1993 al 31.1% (56 su 180) nel periodo successivo ( $P < 0.0001$ ) (Figura 11c).

### FATTORI ASSOCIATI ALL'INFEZIONE CON SOTTOTIPO NON-B

Nei pazienti di cui erano a disposizione i dati demografici completi (CD subset) abbiamo studiato la probabilità di acquisire un'infezione da ceppo non-B (Tabella 2). All'analisi univariata è stata riscontrata una forte associazione tra l'etnia Africana e i ceppi non-B (il 94.8% dei soggetti Africani albergava un ceppo non-B) ( $P < 0.0001$ ), anche se il 49.6% dei pazienti con sottotipo non-B era di origine Europea. La categoria di rischio con maggiore prevalenza nel CD subset erano i tossicodipendenti (35.8%), seguiti dagli eterosessuali (33.7%) e dagli omosessuali (24.4%). Ciononostante un'enorme percentuale di pazienti con infezione da ceppo non-B erano eterosessuali (77.2%;  $P < 0.0001$ ). Nel CD subset il 69.5% dei pazienti erano maschi. E' stata rilevata

una distribuzione del sesso differente nei gruppi di pazienti con infezione da sottotipo non-B e B; tra i pazienti che albergavano un ceppo non-B è stato riscontrato un rapporto maschi femmine 1:1.1 rispetto al rapporto 2.5:1 riscontrato tra i soggetti con infezione da ceppo B. Il sesso femminile è risultato significativamente associato con l'infezione da ceppi non-B ( $P<0.0001$ ), risultando il 14.2% delle donne (59 su 416) con sottotipo non-B rispetto al 6.8% degli uomini (64 di 948).

Il confronto tra i pazienti con ceppo B e non-B ha evidenziato una differenza di età media (rispettivamente 37 *vs* 33 anni) ( $P<0.0001$ ), mentre la conta dei CD4 (rispettivamente 310 *vs* 324 cell/ $\mu$ L) e la carica virale (4.04 *vs* 4.2 log cp/mL) erano comparabili nei due gruppi. L'anno di diagnosi era significativamente associato con la probabilità di infezione da ceppo non-B, essendo i pazienti con sottotipo non-B il 17.4% (104 su 597) nell'intervallo 1993-2006, rispetto al 2.5% (19 di 767) nel periodo prima del 1993 ( $P<0.0001$ ).

L'analisi multivariata ha confermato l'etnia Africana, la modalità di infezione tramite rapporti eterosessuali e una diagnosi di nuova infezione recente quali fattori indipendenti di aumentato rischio di infezione da ceppi non-B (Tabella 2).

Covariata	Analisi univariata			Analisi multivariata		
	OR	95% CI <sup>1</sup>	<i>p</i>	OR	95% CI <sup>1</sup>	<i>p</i>
<b>Etnia</b>						
Europea	-	-	-	-	-	-
Africana	364.9	111.0->999.9	<.0001	149.9	42.5-528.6	<.0001
Latino Americana	3.5	1.0-12.3	0.0275	3.8	1.0-14.6	0.1646
Altro	11.4	3.2-39.9	0.9482	7.9	1.9-32.0	0.9504
<b>Modalità di infezione</b>						
Tossicodipendenti	-	-	-	-	-	-
Eterosessuali	25.1	10.1-62.4	<.0001	9.5	2.8-32.4	0.0001
Omosessuali	3.9	1.4-11.1	0.0963	2.0	0.5-7.6	0.1017
Altro <sup>2</sup>	13.2	4.4-39.8	0.0062	7.0	1.5-32.6	0.0898

<b>Sesso</b>						
Maschi	-	-	-	-	-	-
Femmine	2.3	1.6-3.3	<.0001	0.8	0.4-1.5	0.5055
<b>Anno di diagnosi</b>						
1980-1992	-	-	-	-	-	-
1993-2006	8.3	5.0-13.7	<.0001	4.2	2.1-8.3	<.0001
<b>Età</b>						
per 1-anno in più	0.9	0.8-0.9	0.0003	1.0	0.9-1.0	0.7923

<sup>1</sup> CI: intervallo di confidenza

<sup>2</sup> Altro: Trasfusioni, trasmissione verticale, rischi non noto

**Tabella 2. Rischio (OR) di infezione da ceppo non-B nel gruppo di pazienti con informazioni demografiche complete (CD subset,  $n = 1364$ )**

## ASSEGNAZIONE DEL SOTTOTIPO E DISTRIBUZIONE DEI CEPPI NELLE DIVERSE ETNIE

---

Tra i pazienti è stata riscontrata una grande eterogeneità con 417 ceppi non-B del gruppo M, indipendentemente dai differenti Paesi di origine. Sono stati rilevati tutti i sottotipi puri noti, con l'eccezione del K, più sette distinte CRF (01, 02, 04, 06, 09, 12 e 13). Il sottotipo prevalente era l'F [ $n = 99$  (23.7%); 98 F1 e un F2], seguito dall'A [ $n = 53$  (12.7%); 38 A1, tre A2 e 12 A3], dal C ( $n = 547$ ; 11.3%) e dal G ( $n = 523$ ; 5.5%). Tra le forme ricombinanti circolanti, quelle più diffuse erano la CRF02\_AG e la CRF01\_AE [ $n = 107$  (25.7%) and  $n = 521$  (5.0%), rispettivamente]. Sono state identificate 39 URF (9.3%).

La distribuzione dei sottotipi non-B era molto diversa nei pazienti di origine Europea, rispetto a quelli Africani ( $n = 5192$  e 146, rispettivamente) (Figura 3).

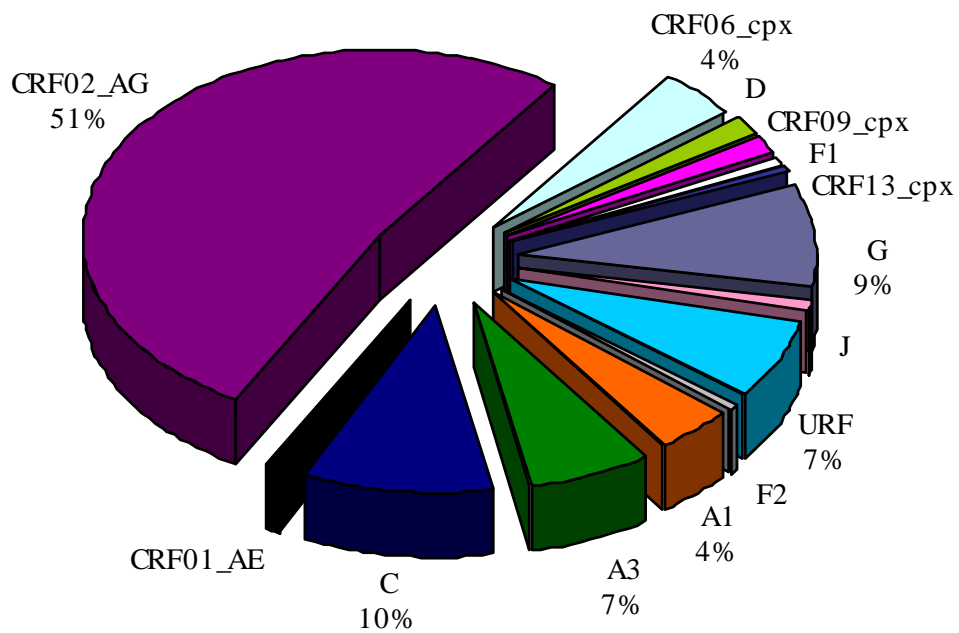
Il sottotipo F1, che era presente in un solo soggetto Africano, era il ceppo più rappresentato nei pazienti di etnia europea con sottotipo non-B (85 su 192; 44.3%), mentre l'A1 ( $n = 524$ ), il C ( $n = 519$ ), la CRF02\_AG ( $n = 59$ ) e le URF ( $n = 519$ )

rappresentavano rispettivamente il 12.5, 9.9, 4.7 e il 9.9%. I pazienti Europei che albergavano il sottotipo F1 erano prevalentemente Italiani ( $n = 568$ ; 82%) e Rumeni ( $n = 513$ ; 15.7%).

Tra i pazienti Europei che albergavano sottotipi non-B, il 64.8% ( $n = 557$ ) erano eterosessuali e il 74.5% (143 su 192) erano maschi. E' stata riscontrata un'associazione tra la modalità di infezione e l'infezione con ceppi non-B ( $P < 0.0001$ ), ma non con il sesso ( $P = 0.46$ ). Abbiamo poi analizzato le differenze tra i pazienti con sottotipo B rispetto ai singoli ceppi non-B per le varianti non-B presenti con una percentuale  $>$  del 5%. Per il sottotipi F1 e C è stata evidenziata un'associazione significativa con la modalità di infezione eterosessuale, con 50% dei pazienti con infezione da F1 eterosessuali (17 di 34), il 100% dei pazienti con sottotipo C (6 su 6) e il 30.6% con ceppo B (528 su 1724) ( $P = 0.006$  per F1 vs. B;  $P < 0.001$  per C vs. B;  $P = 0.026$  per F1 vs. C). Nei pazienti Europei non è stata riscontrata nessuna associazione tra il sesso e i singoli ceppi. La CRF02\_AG è stata riscontrata nel 52.1% dei pazienti Africani che vivono in Italia ( $n = 576$ ), seguita dal sottotipo C ( $n = 515$ ; 10.3%), dall'A [10 A3 (6.9%) and sei A1 (4.1%)], dal G ( $n = 513$ ; 8.9%) e dal B ( $n = 513$ ; 8.2%) e dalle URF ( $n = 510$ ; 6.9%). Per 102 di questi pazienti era noto il Paese di origine. Il 21.6% proveniva dalla Costa d'Avorio, il 21.6% dalla Nigeria, il 12.7% dal Camerun e il 9.9% dal Senegal. La restante parte dei pazienti (34.3%) era migrata dal Nord-Africa ( $n = 59$ ), dall'Africa Occidentale ( $n = 59$ ), Orientale ( $n = 510$ ), Centrale ( $n = 55$ ) e dal Sud-Africa ( $n = 52$ ). Novantasei di questi pazienti (93.2%) erano eterosessuali e il rapporto maschi-femmine era circa 0.5:1 (36:65).

Venti dei 98 (20.4%) pazienti Latino-Americani (52.9% dal Brasile, 15.7% dall'Argentina e il 5% sia dalla Colombia, che dall'Ecuador, dal Perù, dalla Repubblica Dominicana, da Cuba, dal Messico e dal Venezuela) albergavano ceppi non-B (11 F1, sette URF e due C). Nove dei 18 soggetti provenienti dal Sud-Est Asiatico (soprattutto dalle Filippine, Thailandia e India) albergavano sottotipi non-B (sei CRF01\_AE e tre C).

### Africani



### Europei

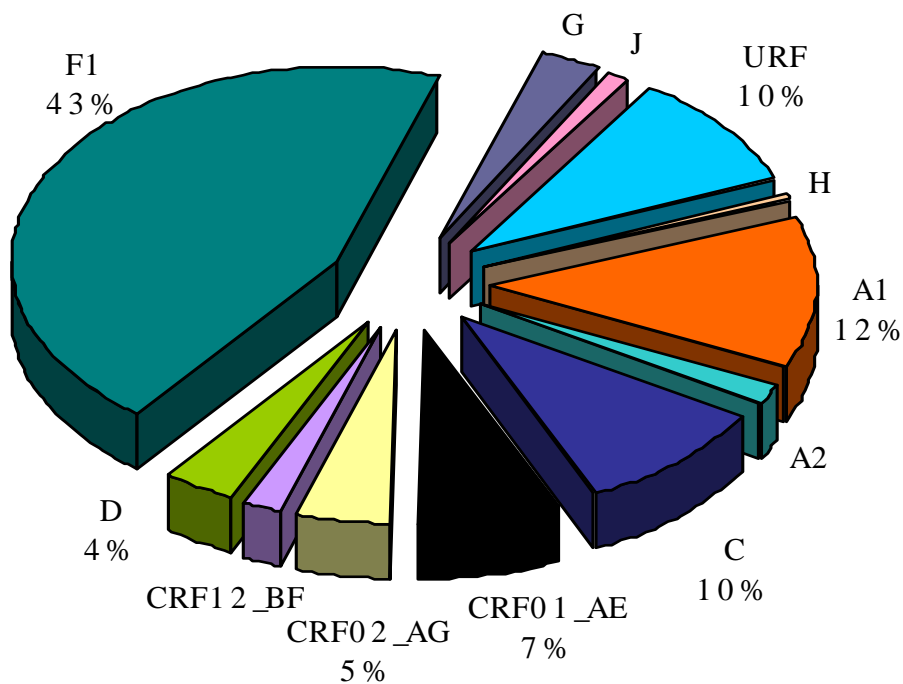


Figura 3. Distribuzione dei sottotipi non-B nei pazienti Africani ed Europei

## ANALISI DI RICOMBINAZIONE DELLE VARIANTI DI HIV-1

---

L'analisi di ricombinazione delle 39 URF ha identificato 13 ricombinanti B/F, sei G/A, quattro D/B, tre A/K, tre G/A/K, tre C/B, due CRF02\_AG/CRF09\_cpx, una CRF02\_AG/B, una CRF06\_cpx/ CRF02\_AG, una CRF18\_cpx/B, una F/C/B e una G/CRF09\_cpx.

La percentuale delle URF era paragonabile nei pazienti con infezione da ceppo non-B Africani (6.8%), Europei (9.3%) e Latino-Americani (7.1%). Come era prevedibile, le URF sono state rilevate nei soggetti Africani provenienti dal Camerun, dalla Repubblica Democratica del Congo, dal Senegal, dalla Nigeria e dalla Costa d'Avorio.

Tutti i ricombinanti B/F sono stati identificati nei pazienti Italiani ( $n = 58$ ) o Brasiliani ( $n = 55$ ).

Un pattern di ricombinazione G/U/F1/B, ottenuto da un paziente cubano, è stato riconosciuto come ricombinante CRF18\_cpx/B, compatibile con il Paese di origine del paziente. Il ricombinante unico CRF06\_cpx/CRF02\_AG è stato correlate con l'isolato 00NE-36 dalla Nigeria, che è stato proposto come sequenza di riferimento per la CRF30\_cpx ([www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html](http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html)).

Una delle due forme mosaico CRF02\_AG/CRF09\_cpx è stata rilevata da un paziente nato in Costa d'Avorio, dove è stata isolata questa ricombinazione di seconda generazione

Da sottolineare che due gruppi di tre sequenze ciascuno erano altamente omologhe con gli isolati (A/K)<sup>92</sup> e 99GR303 (G/A/K)<sup>93</sup>.

## 5 DISCUSSIONE

---

Negli ultimi anni l'epidemia di HIV-1 in Europa è stata caratterizzata dall'importante aumento della penetrazione e della circolazione dei sottotipi non-B in conseguenza dei flussi migratori provenienti dall'Africa sub-Sahariana, dal Sud-Est Asiatico e dal Centro e Sud-America che sono cominciati fin dagli inizi degli anni '90<sup>71-78</sup>. Si pensa inoltre che, oltre ai fenomeni migratori, i viaggi verso aree ad alta prevalenza di infezione da HIV-1, in particolare quelli dove è diffuso il fenomeno del turismo sessuale, siano responsabili dell'ingresso dei diversi sottotipi del gruppo M nei Paesi precedentemente ristretti al sottotipo B. I dati Italiani del Centro Operativo AIDS, che si basano sulle nuove diagnosi di HIV, indicano che la percentuale dei pazienti stranieri (41.2% dall' Africa, 25.2% dall'America Latina e il 16.1% dall'Europa) sono aumentati dall' 11 al 32% dal 1992 al 2007, e che i rapporti eterosessuali sono la principale modalità di infezione essendo aumentati dal 24.6 al 75.9% nello stesso periodo. Complessivamente, tra i pazienti con nuova diagnosi di infezione da HIV nel periodo dal 1985 al 2007, la percentuale dei tossicodipendenti è diminuita dal 69 all'8.6%, mentre la trasmissione sessuale è aumentata dal 13.3 al 73.7% e il rapporto maschi/femmine si è ridotto da 3.5 a 2.5<sup>83</sup>.

La distribuzione dell'etnia e della modalità di infezione della nostra popolazione in studio sono in accordo con questi dati. Inoltre abbiamo potuto studiare le percentuali relative di eterosessuali e omosessuali in un'ampia coorte di pazienti con nuova diagnosi a cui afferivano principalmente pazienti del Centro-Italia. Abbiamo riscontrato che meno del 3% dei pazienti albergava un sottotipo non B prima del 1993, a confronto con il 20% circa negli anni successivi. E' importante sottolineare che la probabilità di acquisire un sottotipo non-B è aumentata di più di 0.25 volte per anno. La prevalenza dei ceppi non-B in Italia è stata valutata per la prima volta in uno studio del 2001 che riportava una prevalenza complessiva del 5.4% tra i pazienti naive per la terapia, con un trend in aumento nel tempo<sup>72</sup>. Due studi successivi hanno riportato una prevalenza più elevata, del 12.6 e 10.7% in regioni con rispettivamente una medio-bassa e alta incidenza di infezione<sup>94,95</sup>. Entrambi, nonostante evidenzino un aumento della prevalenza dei non-B nel tempo, riportano dei valori inferiori rispetto a quelli di

questo studio, come negli studi di sorveglianza condotti negli altri Paesi Europei come la Francia, il Belgio e il Regno Unito<sup>73,74,76</sup>.

In accordo con numerosi studi, la diffusione dei sottotipi non-B è fortemente legata a numerose variabili che definiscono l'epidemia locale dell'infezione da HIV-1, e alla loro evoluzione nel tempo. Nel nostro studio è aumentata nel tempo la percentuale di pazienti di etnia non caucasica e di quelli che hanno contratto l'infezione attraverso rapporti eterosessuali. Tuttavia abbiamo anche riscontrato una prevalenza maggiore di varianti non-B nei pazienti Europei dopo il 1992, avendo riscontrato un aumento di 5 volte della percentuale di soggetti con ceppi non-B rispetto al primo periodo. Come ci aspettavamo, l'analisi di regressione ha indicato una forte associazione tra l'etnia Africana e l'infezione da ceppi non-B. Tuttavia il 50% dei pazienti che albergava una variante diversa dal sottotipo B era di etnia Caucasica, suggerendo che questi ceppi siano stati successivamente trasmessi agli Europei in una percentuale considerevole. L'aumento della prevalenza dei ceppi non-B è stata rilevata complessivamente in tutti i gruppi a rischio; tuttavia l'incremento più evidente è stato evidenziato tra gli eterosessuali. L'analisi multivariata effettuata sui pazienti del CD subset ha mostrato che la modalità di infezione eterosessuale era un fattore di rischio indipendente di infezione da HIV-1 con un ceppo non-B, essendo stato trovato un rischio di 9.5 volte aumentato di albergare un sottotipo non-B per gli eterosessuali. Il rapporto maschi/femmine era paragonabile tra il periodo prima del 1993 e il periodo successivo (rispettivamente 2.25 vs. 2.32), e il sesso femminile non rappresentava un fattore di rischio indipendente di infezione da non-B, probabilmente a causa delle caratteristiche dell'epidemia locale di HIV, come l'alta percentuale di donne fra i tossicodipendenti. Tuttavia le donne con varianti non-B rappresentavano una cospicua percentuale (almeno un terzo) sul numero complessivo di femmine con diagnosi dopo il 1992.

Infine la valutazione del momento della diagnosi di HIV indica chiaramente che il rischio di acquisire un'infezione da non-B è quattro volte superiore per coloro che hanno avuto una diagnosi dopo il 1993 rispetto a quelli che l'hanno avuta negli anni precedenti.

Nel nostro studio abbiamo riscontrato un'ampia eterogeneità di varianti non-B del gruppo M, a dimostrazione che le fonti dell'infezione da non-B erano ubiquitarie. Nonostante la cittadinanza non fosse nota per tutti i pazienti, l'isolamento di sottotipi specifici negli Africani, nei Latino-Americani e nei soggetti del Sud-Est Asiatico rifletteva strettamente i ceppi non-B che circolano ad alta prevalenza nel loro Paese di origine, suggerendo che questi pazienti si erano probabilmente infettati nel loro Paese. Da notare che il nostro studio ha evidenziato che il sottotipo F1 era quello predominante (con una frequenza di più del 50%) tra i pazienti Europei con ceppo non-B, di cui più dell'80% erano Italiani. In particolare è stato riscontrato che questo ceppo, che ha un'alta prevalenza in Sud-America<sup>96</sup> e che ha una certa rilevanza nell'Europa dell'Est<sup>97</sup>, era significativamente associato con la modalità di trasmissione eterosessuale. Questa scoperta necessita di ulteriori approfondimenti, sia attraverso la raccolta di informazioni sulle abitudini sessuali sia utilizzando l'approccio filogenetico per indagare la probabile origine di queste infezioni. Queste informazioni possono essere utili negli interventi di Sanità Pubblica.

Abbiamo isolato un'insolita alta percentuale (circa il 10%) di URF tra i sottotipi non-B nei pazienti Caucasic, Africani e Latino-Americani. Ciò potrebbe essere il risultato di un'analisi filogenetica e di ricombinazione molto accurata. In effetti sono stati rilevati due tipi di URF, ciascuno in tre pazienti, il che potrebbe candidare queste forme a nuove CRF mediante la caratterizzazione e il sequenziamento dell'intero genoma. E' stata inoltre riscontrata un'alta percentuale (> 30%) di URF con pattern di ricombinazione B/F, a conferma dei risultati di precedenti studi che avevano evidenziato che questi due ceppi hanno un'elevata capacità di ricombinazione<sup>96,98</sup>. L'ulteriore diffusione di questi ricombinanti, sovrapponendosi al sottotipo attualmente predominante, potrebbe condizionare l'andamento dell'epidemia da HIV in Italia che in futuro potrebbe differenziarsi dal resto dell'Europa.

Il nostro approccio metodologico aveva alcuni limiti. Innanzitutto non avevamo a disposizione i dati relativi alla durata del soggiorno in Italia dei pazienti di cui conoscevamo il Paese di origine. Di conseguenza questi pazienti potrebbero aver acquisito l'infezione nel loro Paese o successivamente in Italia. Allo stesso modo non

erano state raccolte informazioni relative a possibili viaggi dei soggetti Italiani. Inoltre il numero dei individui di cui era nota la data di sieroconversione era troppo ridotto per poter essere utilizzato per determinare il momento in cui i ceppi non-B sono penetrati in Italia. Pertanto abbiamo utilizzato la data del primo test HIV positivo per stimare la durata dell'infezione, che è un approccio estremamente conservativo; è probabile che l'ingresso dei ceppi non-B in Italia e l'aumento della loro circolazione sia avvenuto prima di quanto suggeriscono i nostri dati. A supportare questa ipotesi potrebbe essere l'elevata percentuale di pazienti "AIDS presenting", che ricevono una diagnosi molti anni dopo aver contratto l'infezione. Il terzo limite consiste nell'uso delle sequenze di pol per l'assegnazione del sottotipo. E' largamente accettato che questa regione virale che comprende più di 1000 nucleotidi sia appropriata per valutare gli andamenti epidemiologici dell'infezione da HIV nelle popolazioni dei pazienti infetti<sup>84,85</sup>. Tuttavia l'impiego del gene pol per individuare il sottotipo non esclude la possibilità che le altre regioni del genoma possano essere attribuibili ad altri sottotipi<sup>99</sup>. L'analisi di una porzione limitata porzione del genoma di HIV-1 come ad esempio pol conferma che gli eventi di ricombinazione si verificano con una frequenza elevate nei ceppi circolanti di HIV-1 e possono influenzare la diffusione dei sottotipi di HIV-1 a livello globale<sup>67-69,82</sup>.

La variabilità di HIV-1 ha un impatto su diversi aspetti dell'infezione da HIV, che comprendono la diagnosi molecolare<sup>100</sup>, il tropismo cellulare<sup>101,102</sup>, i meccanismi della resistenza ai farmaci<sup>103</sup>, la fitness<sup>104</sup>, la progressione di malattia<sup>105</sup> e la trasmissione madre-bambino<sup>106</sup>. Anche se ha alcuni limiti, l'assegnazione del sottotipo sulle sequenze ottenute dai test di resistenza è uno strumento fondamentale per la sorveglianza della diffusione dei ceppi di HIV-1 e può essere d'ausilio per comprendere le caratteristiche biologiche e cliniche dell'infezione da HIV e incrementare le strategie di prevenzione.

## 6 BIBLIOGRAFIA

---

---

- <sup>1</sup> **Gonda MA, Wong-Staal F, Gallo RC and Chou MJ.** Sequence homology and morphologic similarity of HTLV-III and visna virus, a pathogenic lentivirus. *Science* 1985; 227:173-7.
- <sup>2</sup> **Gonda MA, Braun MJ Clements JE et al.** Human T-cell lymphotropic virus type III shares sequence homology with a family of pathogenic lentivirus. *Science* 1985; 83: 4007-11.
- <sup>3</sup> **Myers G, and Pavlakis GN.** Evolutionary potential of complex retroviruses. In: Levy J.A., ed. *The retroviridae*. New York: Plenum. Press 1992; 51-105.
- <sup>4</sup> **Jones WK, and Curran JW.** Epidemiology of HIV infection in industrialized countries. In: **Broder S, Merigan TC, Bolognesi D.**, eds . *textbook of AIDS medicine*. Baltimore: Williams and Wilkins 1994, 91-108.
- <sup>5</sup> **Quinn TC.** Population migration and the spread of types 1 and 2 human immunodeficiency viruses. *Natl Acad Sci USA* 1994; 91: 2407-2414.
- <sup>6</sup> **Rambaut A, Posada D et al.** The causes and consequences of HIV evolution. *Nature* 2004: 52-60.
- <sup>7</sup> **Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM et al.** Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes* *Nature* 1999; 397: 436-41.
- <sup>8</sup> **Hahn BH, Shaw GM, DeCock KM, Sharp PM.** AIDS as a Zoonosis. *Science* 2000; 287: 607-14.
- <sup>9</sup> **Bailes E, Gao F, Bibollet-Ruche F et al.** Hybrid Origin of SIV in Chimpanzees. *Science* 2003; 300: 1713.
- <sup>10</sup> **Van Heuverswyn F. et al.** Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature* 2006; 444: 166.
- <sup>11</sup> **Korber B, Kuiken C, Foley B et al.** *Human Retroviruses and AIDS 1998: A compilation and Analysis of Nucleic Acid and Amino Acid Sequences*. Los Alamos National Laboratory Los Alamos, New Mexico.
- <sup>12</sup> **Perrin L, Kaiser L, and Yerly S.** Travel and the spread of HIV-1 genetic variants. *The Lancet* 2003; 3: 22-7.

- 
- <sup>13</sup> **Plantier JC., Leoz M.** A new human immunodeficiency virus derived from gorillas. *Nature Medicine* 2009; 871-872.
- <sup>14</sup> **Coffin JM.** HIV population dynamics *in vivo*: implications for genetic variation, pathogenesis, and therapy. *Science* 1995; 267: 483-9.
- <sup>15</sup> **Gallay P, Swingler S, Aiken C et al.** HIV-1 infection of nondividing cells: C-terminal tyrosine phosphorylation of the viral matrix protein is a key regulator. *Cell* 1995; 80: 379-388.
- <sup>16</sup> **Lewis PF, and Emerman M.** Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *J Virol* 1994; 68: 510-516.
- <sup>17</sup> **Stevens SW, and Griffith JD.** Human immunodeficiency virus type 1 may preferentially integrate into chromatin occupied by L1s repetitive elements. *Proc Natl Acad Sci USA* 1994; 91:5557-5561.
- <sup>18</sup> **Contag CH, Dewhurst S, Viglianti GA et al.** Simian immunodeficiency virus (SIV) from Old World monkeys. In: Gallo RC., Jay G, eds. *The human retroviruses*. San Diego: Academic. Press. 245-276. 1991.
- <sup>19</sup> **Delwart EL, Shaper EG, Louwagie J et al.** Genetic relationships determined by a DNA heteroduplex mobility analysis of HIV-1 *env* genes. *Science* 1993; 262:1257-1261.
- <sup>20</sup> **Myers G, Korber BTM, Smith RF et al.** *Human Retroviruses and AIDS. Theoretical Biology and Biophysics*, Los Alamos National Laboratory, Los Alamos, NM, 1994.
- <sup>21</sup> **Korber BTM, Allen EE, Farmer AD et al.** Heterogeneity of HIV-1 and HIV-2. *AIDS* 1995; 9: S5-S18.
- <sup>22</sup> **Hue S, Clewley J, Canea A and Pillay D.** HIV-1 *pol* gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. *AIDS* 2004; 18: 719-728.
- <sup>23</sup> **McCutchan F.** Understanding the genetic diversity of HIV. *AIDS* 2000; 14: S31-S44.
- <sup>24</sup> **Robertson DL, Anderson JP, Bradac JA et al.** HIV-1 nomenclature proposal. In: *Human retroviruses and AIDS 1999: a compilation and analysis of nucleic acid and amino acid sequences*. Kuiken CL, Foley B, Hahn B, Korber B, McCutchan F, Marx PA,

---

*et al*, eds. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.

<sup>25</sup> **Anderson JP, Rodrigo AG, Learn GH, Madan A, Delahunty C, Coon M, Girard M, Osmanov S, Hood L, Mullins JI.** Testing the hypothesis of a recombinant origin of human immunodeficiency virus type 1 subtype E. *J Virol* 2000; 74: 10752-65.

<sup>26</sup> **Paraskevis D, Magiorkinis M, Vandamme AM et al.** Re-analysis of Human Immunodeficiency Virus Type 1 isolates from Cyprus and Greece, initially designated 'subtype I', reveals a unique complex A/G/H/K/? mosaic pattern. *J Gen Virol* 2001; 82:575-580.

<sup>27</sup> **Thomson MM, Delgado E, Manjón N et al.** HIV-1 genetic diversity in Galicia, Spain: BG intersubtype recombinant viruses are circulating among injecting drug users. *AIDS* 2001; 15: 509–16.

<sup>28</sup> **Delgado E, Thomson MM, Villahermosa ML et al.** Identification, by analysis of near full-length sequences, of a newly characterized HIV-1 BG intersubtype circulating recombinant form in Galicia, Spain, exhibiting a pseudotype-like virion structure. *J Acquir Immune Defic Syndr* 2002; 29: 536–43.

<sup>29</sup> **Esteves A, Parreira R, Venenno T et al.** Molecular epidemiology of HIV type 1 infection in Portugal: high prevalence of non-B subtypes. *AIDS Res Hum Retroviruses* 2002; 18: 313–25.

<sup>30</sup> **Mokili J, Rogers M, Carr J et al.** Identification of a novel clade of Human Immunodeficiency Virus Type 1 in Democratic Republic of Congo. *AIDS Res Hum Retroviruses* 2002; 18 (11):817-823.

<sup>31</sup> **Peeters M.** Recombinant HIV sequences: their role in the global epidemic. In: Human retroviruses and AIDS 2000: a compilation and analysis of nucleic acid and amino acid Sequences. Kuiken CL, Foley B, Hahn B, et al, eds. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.

<sup>32</sup> HIV Sequence Database. Los Alamos National Laboratory, NM: <http://hiv-web.lanl.gov>

<sup>33</sup> **Thomson MM, Villahermosa ML, Vázquez-de Parga E et al.** Widespread circulation of a B/F intersubtype recombinant form among HIV-1-infected individuals in Buenos Aires, Argentina. *AIDS* 2000; 14: 897–99.

- 
- <sup>34</sup> **Thomson MM, Delgado E, Herrero I et al.** Diversity of mosaic structures and common ancestry of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Argentina revealed by analysis of near full-length genome sequences. *J Gen Virol* 2002; 83: 107–19.
- <sup>35</sup> **Carr JK, Avila M, Gomez Carrillo M, Salomon H, Hierholzer J, Watanaveeradej V, Pando MA, Negrete M, Russell KL, Sanchez J, Birx DL, Andrade R, Vinales J, McCutchan FE.** Diverse BF recombinants have spread widely since the introduction of HIV-1 into South America. *AIDS* 2001; 15:F41-7.
- <sup>36</sup> **De Sá Filho DJ, Sucupira MC, Caseiro MM, Sabino EC, Diaz RS, Janini LM.** Identification of two HIV type 1 circulating recombinant forms in Brazil. *AIDS Res Hum Retroviruses* 2006; 22: 1-13.
- <sup>37</sup> **Sanabani S, Kleine Neto W, Kalmar EM, Diaz RS, Janini LM, Sabino EC.** Analysis of the near full length genomes of HIV-1 subtypes B, F and BF recombinant from a cohort of 14 patients in São Paulo, Brazil. *Infect Genet Evol* 2006; 6:368-77.
- <sup>38</sup> **Guimarães ML, Eyer-Silva WA, Couto-Fernandez JC, Morgado MG.** Identification of two new CRF\_BF in Rio de Janeiro State, Brazil. *AIDS* 2008; 22:433-5.
- <sup>39</sup> **Thomson MM, Casado G, Posada D, Sierra M, Nájera R.** Identification of a novel HIV-1 complex circulating recombinant form (CRF18\_cpx) of Central African origin in Cuba. *AIDS* 2005; 19:1155-63.
- <sup>40</sup> **Cuevas MT, Ruibal I, Villahermosa ML et al.** High HIV-1 genetic diversity in Cuba. *AIDS* 2002; 16: 1643-53.
- <sup>41</sup> **Pérez L, Thomson MM, Bleda MJ, Aragonés C, González Z, Pérez J, Sierra M, Casado G, Delgado E, Nájera R** HIV Type 1 molecular epidemiology in cuba: high genetic diversity, frequent mosaicism, and recent expansion of BG intersubtype recombinant forms. *AIDS Res Hum Retroviruses* 2006;22:724-33.
- <sup>42</sup> **Osmanov S, Pattou C, Walker N, Schwardlander B, Esparza J.** WHO-UNAIDS Network for HIV Isolation and Characterization. Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. *J Acquir Immune Defic Syndr* 2002; 29: 184–90.
- <sup>43</sup> **Tovanabutra S, Watanaveeradej V, Viputtikul K et al.** A new circulating recombinant form, CRF15\_01B, reinforces the linkage between IDU and heterosexual epidemics in Thailand. *AIDS Res Hum Retroviruses* 2003; 19: 561-567.

- 
- <sup>44</sup> **Tee KK, Li XJ, Nohtomi K, Ng KP, Kamarulzaman A, Takebe Y.** Identification of a novel circulating recombinant form (CRF33\_01B) disseminating widely among various risk populations in Kuala Lumpur, Malaysia. *J Acquir Immune Defic Syndr*; 2006; 43:523-9.
- <sup>45</sup> **Tovanabutra S, Kijak GH, Beyrer C, Gammon-Richardson C, Sakkhachornphop S, Vongchak T, Jittiwutikarn J, Razak MH, Sanders-Buell E, Robb ML, Suriyanon V, Birx DL, Michael NL, Celentano DD, McCutchan FE.** Identification of CRF34\_01B, a second circulating recombinant form unrelated to and more complex than CRF15\_01B, among injecting drug users in northern Thailand. *AIDS Res Hum Retroviruses* 2007; 23: 829-32.
- <sup>46</sup> **Adojaan M, Kivisild T, Männik A, Krispin T, Ustina V, Zilmer K, Liebert E, Jaroslavtsev N, Priimägi L, Tefanova V, Schmidt J, Krohn K, Villems R, Salminen M, Ustav M.** Predominance of a rare type of HIV-1 in Estonia. *J Acquir Immune Defic Syndr* 2005; 39: 598-605.
- <sup>47</sup> **Triques K, Bourgeois A, Vidal N, Mpoudi-Ngole E, Mulanga-Kabeya C, Nzilambi N, Torimiro N, Saman E, Delaporte E, Peeters M.** Near-full-length genome sequencing of divergent African HIV type 1 subtype F viruses leads to the identification of a new HIV type 1 subtype designated K. *AIDS Res Hum Retroviruse* 2000; 16: 139-51.
- <sup>48</sup> **Janssens W, Salminen MO, Laukkanen T, et al.** Near full-length genome analysis of HIV type 1 CRF02\_AG subtype C and CRF02\_AG subtype G recombinants. *AIDS Res Hum Retroviruses* 2000, 16:1183-1189.
- <sup>49</sup> **Vidal N, Peeters M, Mulanga-Kabeya C et al.** Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. *J Virol* 2000; 74: 10498–507.
- <sup>50</sup> **Zhu T, Corber BT et al.** An African HIV-1 sequence from 1959 and implications for the origin of epidemic. *Nature* 1998; 391: 594-7.
- <sup>51</sup> **Liitsola K, Tashkinova I, Laukkanen T, Korovina G, Smolskaja T, Momot O, Mashkilleysen N, Chaplinskas S, Brummer-Korvenkontio H, Vanhatalo J, Leinikki P, Salminen MO.** HIV-1 genetic subtype A/B recombinant strain causing an explosive epidemic in injecting drug users in Kaliningrad. *AIDS* 1998; 12:1907-19.
- <sup>52</sup> **Korovina GI, Smol'skaia TT, Tashkinova IP, Momot OF, Kevlova NA, Liitsola K, Leinikki P, Salminen M.** A recombinant HIV-1 strain causing an epidemic

---

among injection drug addicts in Kaliningrad. *Zh Mikrobiol Epidemiol Immunobiol* 1999; 1: 47-9 (Russo).

<sup>53</sup> **Thomson MM, Nájera R.** Travel and the introduction of human immunodeficiency virus type1 non-B subtype genetic forms into Western countries. *Clin Infect Dis* 2001; 32: 1732-37.

<sup>54</sup> **Hoelscher M, Kim B, Maboko L, Mhalu F, von Sonnenburg F, Birx DL, McCutchan FE;** The UNAIDS Network for HIV Isolation and Characterization. High proportion of unrelated HIV-1 intersubtype recombinants in the Mbeya region of southwest Tanzania. *AIDS* 2001; 15: 1461-70.

<sup>55</sup> **Menendez-Arias L.** Targeting HIV: antiretroviral therapy and development of drug resistance. *Trends Pharmacol Sci* 2002; 23: 381-388.

<sup>56</sup> **Hu DJ, Dondero TJ, Rayfield MA et al.** The emerging genetic diversity of HIV: The importance of global surveillance for diagnostics, research, and prevention. *JAMA* 1996; 275: 210-216.

<sup>57</sup> **Amornkul PN, Tansuphasawadikul S, Limpakarnjanarat K et al.** Clinical disease associated with HIV-1 subtype B and E infection among 2104 patients in Thailand. *AIDS* 1999; 13: 1963-1969.

<sup>58</sup> **Kilmarx PH, Limpakarnjanarat K, Kaewkungwal J et al.** Disease progression and survival with human immunodeficiency virus type 1 subtype E infection among female sex workers in Thailand. *J Infect Dis* 2000; 181: 1598-1606.

<sup>59</sup> **Neilson JR, John GC, Carr JK et al.** Subtypes of human immunodeficiency virus type and disease stage among women in Nairobi, Kenya. *J Virol* 1999; 73: 4393-4403.

<sup>60</sup> **Kaleebu PFN, Mahe C, Yirrell D et al.** The role of HIV-1 envelope subtypes A and D on disease progression in a large cohort of HIV-1 positive individuals in Uganda. *XIII International Conference on AIDS*. Durban, July 2000 [abstract A3079].

<sup>61</sup> **Chew CB, Herring BL, Zheng F et al.** Comparison of three commercial assay for the quantification of HIV-1 RNA in plasma from individuals infected with different HIV-1 subtypes. *J Clin Virol* 1999; 14: 87-94.

<sup>62</sup> **Burgisser P, Vernazza P, Flepp M et al.** Performance of five different assays for the quantification of viral load in persons infected with various subtypes of HIV-1. Swiss HIV Cohort Study. *J Acquir Immune Defic Syndr* 2000; 23:138-144.

- 
- <sup>63</sup> **Descamps D, Apetrei C, Collin G, Damond D, Brun-Vézinet F.** Naturally occurring decreased susceptibility of HIV-1 subtype G to protease inhibitors. *AIDS* 1998; 12: 1109–11
- <sup>64</sup> **Apetrei C, Descamps D, Collin G et al.** Human immunodeficiency virus type 1 subtype F reverse transcriptase sequence and drug susceptibility. *J Virol* 1998; 72: 3534–38
- <sup>65</sup> **Suligoï B, Boros S. et al.** Aggiornamento delle nuove diagnosi di infezione da HIV al 31 dicembre 2008 e dei casi di AIDS in Italia al 31 dicembre 2009. *Not Ist Super Sanità*, 2010; 23: 3-27.
- <sup>66</sup> **Robertson DL, Sharp PM, McCutchan FE, Hahn BH.** Recombination in HIV-1. *Nature* 1995; 374: 124–126.
- <sup>67</sup> **Leitner T, Korber B, Daniels M, Calef C, Foley B.** HIV-1 Subtype and Circulating Recombinant Form (CRF) Reference Sequences, 2005. Los Alamos National Laboratory, Los Alamos, NM 87545. seq-info@lanl.gov. Available at [www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/2005/partI/leitner.pdf](http://www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/2005/partI/leitner.pdf) (accessed March 2010).
- <sup>68</sup> **Taylor BS, Hammer SM.** The challenge of HIV-1 subtype diversity. *New Engl J Med* 2008; 12: 1590–1602.
- <sup>69</sup> **Hemelaar J, Gouws E, Ghys PD, Osmanov S.** Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 2006; 20: W13–W23.
- <sup>70</sup> **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 Res Hum Retroviruses* 1996; 12: 1179–1183.
- <sup>71</sup> **Sonnerborg A, Durdevic S, Giesecke J, Saïllberg M.** Dynamics of the HIV-1 subtype distribution in the Swedish HIV-1 epidemic during the period 1980 to 1993. *AIDS Res Hum Retroviruses* 1997; 13: 343–345.
- <sup>72</sup> **Balotta C, Facchi G, Violin M et al.** ICONA study group. Increasing prevalence of non clade B HIV-1 strains in heterosexual men and women, as monitored by analysis of reverse transcriptase and protease sequences. *J Acquir Immune Defic Syndr* 2001; 27: 499–505.
- <sup>73</sup> **Chaix ML, Descamps D, Harzic M et al.** Stable prevalence of genotypic drug resistance mutations but increase in non-B virus among patients with primary HIV-1 infection in France. *AIDS* 2003; 17: 2635–2643.
- <sup>74</sup> **Snoeck J, Van Laethem K, Hermans P et al.** Rising prevalence of HIV-1 non-B subtypes in Belgium: 1983–2001. *J Acquir Immune Defic Syndr* 2004; 35: 279–285.

- 
- <sup>75</sup> **Lospitao E, Alvarez A, Soriano V, Holguín A.** HIV-1 subtypes in Spain: a retrospective analysis from 1995 to 2003. *HIV Med* 2005; 6: 313–320.
- <sup>76</sup> **Gifford RJ, de Oliveira T, Rambaut A et al.** Phylogenetic surveillance of viral genetic diversity and the evolving molecular epidemiology of human immunodeficiency virus type 1. *J Virol* 2007; 81: 13050–13056.
- <sup>77</sup> **Palma AC, Araújo F, Duque V, Borges F, Paixão MT, Camacho R.** Portuguese SPREAD network. Molecular epidemiology and prevalence of drug resistance-associated mutations in newly diagnosed HIV-1 patients in Portugal. *Infect Genet Evol* 2007; 7: 391–398.
- <sup>78</sup> **Yerly S, von Wyl V, Ledergerber B et al.** Swiss HIV cohort study. Transmission of HIV-1 drug resistance in Switzerland: a 10-year molecular epidemiology survey. *AIDS* 2007; 21: 2223–2229.
- <sup>79</sup> **Masquelier B, Bhaskaran K, Pillay D et al.** CASCADE collaboration. 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: 505–511.
- <sup>80</sup> **Wensing AM, Vijver DA, Angarano G et al.** Prevalence of drug-resistant HIV-1 variants in untreated individuals in Europe: implications for clinical management. *J Infect Dis* 2005; 192: 958–966.
- <sup>81</sup> **Wensing AMJ, Vercauteren J, Vijver DA et al.** Transmission of drug-resistant HIV-1 in Europe remains limited to single classes. *AIDS* 2008; 22: 625–635.
- <sup>82</sup> **Thomson MM, Pe´rez-Alvarez L, Na´jera R.** Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *Lancet Infect Dis* 2002; 2: 461–471.
- <sup>83</sup> Notiziario COA Marzo 2009. Available at [www.iss.it/binary/publ/cont/COAonline.pdf](http://www.iss.it/binary/publ/cont/COAonline.pdf) (accessed March 2010).
- <sup>84</sup> **Lemey P, Derdelinckx I, Rambaut A et al.** Molecular footprint of drug-selective pressure in a human immunodeficiency virus transmission chain. *J Virol* 2005; 79: 11981–11989.
- <sup>85</sup> **Hue´ S, Gifford RJ, Dunn D, Fernhill E, Pillay D.** UK collaborative group on HIV drug resistance. Demonstration of sustained drug-resistant human immunodeficiency virus type 1 lineages circulating among treatment-naïve individuals. *J Virol* 2009; 83: 2645–2654.

- 
- 86 BioEdit Home Page <http://www.mbio.ncsu.edu/BioEdit/page2.html>
- 87 LANL (Los Alamos National Lab) Home page <http://www.lanl.gov/>
- 88 PHYLIP Home Page <http://evolution.genetics.washington.edu/phylip.html>
- 89 TreePuzzle Home Page <http://www.tree-puzzle.de>
- 90 SimPlot Home Page <http://sray.med.som.jhmi.edu/SCRsoftware/>
- 91 BLAST Home Page <http://www.ncbi.nlm.nih.gov/BLAST/>
- 92 **Oelrichs RB, Workman C, Laukkanen T, McCutchan FE, Deacon NJ.** A novel subtype A/G/J recombinant full-length HIV type 1 genome from Burkina Faso. *AIDS Res Hum Retroviruses* 1998; 14: 1495–1500.
- 93 **Nasioulas G, Paraskevis D, Magiorkinis E, Theodoridou M, Hatzakis A.** Molecular analysis of the full-length genome of HIV type 1 subtype I: evidence of A/G/I recombination. *AIDS Res Hum Retroviruses* 1999; 15: 745–758.
- 94 **Monno L, Brindicci G, Lo Caputo S et al.** HIV-1 subtypes and circulating recombinant forms (CRFs) from HIV-infected patients residing in two regions of central and southern Italy. *J Med Virol* 2005; 75: 483–490.
- 95 **Baldanti F, Paolucci S, Ravasi G et al.** Changes in circulation of B and non-B HIV strains: spotlight on a reference centre for infectious diseases in Northern Italy. *J Med Virol* 2008; 80: 947–952.
- 96 **Thomson MM, Na'jera R.** Molecular epidemiology of HIV-1 variants in the global AIDS pandemic: an update. *AIDS Rev* 2005; 7: 210–224.
- 97 **Op De Coul E, van den Burg R, Asjo" B et al.** Genetic evidence of multiple transmissions of HIV type 1 subtype F within Romania from adult blood donors to children. *AIDS Res Hum Retroviruses* 2000; 16: 327–336.

- 
- <sup>98</sup> **Thomson M, Sierra M, Tanuri A et al.** Analysis of near fulllength genome sequences of HIV type 1 BF intersubtype recombinant viruses from Brazil reveals their independent origins and their lack of relationship to CRF12\_BF. *AIDS Res Hum Retroviruses* 2004; 20: 1126–1133.
- <sup>99</sup> **Razzolini F, Vicenti I, Saladini F et al.** Natural variability in the HR-1 and HR-2 domains of HIV type 1 gp41 from different clades circulating in Italy. *AIDS Res Hum Retroviruses* 2007; 23: 558–563.
- <sup>100</sup> **Holguin A, Lopez M, Molinero M, Soriano V.** Performance of three commercial viral load assays, versant human immunodeficiency virus type 1 (HIV-1) RNA bDNA v3.0, Cobas AmpliPrep/Cobas TaqMan HIV-1, and NucliSens HIV-1 EasyQ v1.2, testing HIV-1 non-B subtypes and recombinant variants. *J Clin Microbiol* 2008; 46: 2918–2923.
- <sup>101</sup> **Lobritz MA, Marozsan AJ, Troyer RM, Arts EJ.** Natural variation in the V3 crown of human immunodeficiency virus type 1 affects replicative fitness and entry inhibitor sensitivity. *J Virol* 2007; 81: 8258–8269.
- <sup>102</sup> **Huang W, Eshleman SH, Toma J et al.** Coreceptor tropism in human immunodeficiency virus type 1 subtype D: high prevalence of CXCR4 tropism and heterogeneous composition of viral populations. *J Virol* 2007; 81: 7885–7893.
- <sup>103</sup> **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). *J Int AIDS Soc* 2009; 12: 1–11.
- <sup>104</sup> **Arie'n KK, Abraha A, Quinones-Mateu ME, Kestens L, Vanham G, Arts EJ.** The replicative fitness of primary human immunodeficiency virus type 1 (HIV-1) group M, HIV-1 group O, and HIV-2 isolates. *J Virol* 2005; 79: 8979–8990.
- <sup>105</sup> **Keller M, Lu Y, Lalonde RG, Klein MB.** Impact of HIV-1 viral subtype on CD4T-cell decline and clinical outcomes in antiretroviral naive patients receiving universal healthcare. *AIDS* 2009; 23: 731–737.

---

<sup>106</sup> **Eshleman SH, Guay LA, Mwatha A et al.** Comparison of mother-to-child transmission rates in Ugandan women with subtype A versus D HIV-1 who received single-dose nevirapine prophylaxis: HIV network for prevention trials 012. *J Acquir Immune Defic Syndr* 2005; 39: 593–597.