



**SAPIENZA**  
UNIVERSITÀ DI ROMA

Facoltà di Medicina e Chirurgia  
Dottorato in Medicina Sperimentale  
XXV ciclo

## **Markers genotipici del tropismo di HIV-1**

DOCENTE GUIDA

Prof. Guido Antonelli

DOCENTE TUTOR

Prof.ssa Ombretta Turriziani

DOTTORANDA

Claudia Montagna

COORDINATRICE DEL DOTTORATO

Prof.ssa Maria Rosaria Torrisi

ANNO ACCADEMICO 2012/2013

# INDICE

Sommario	4
<u>1. INTRODUZIONE</u>	
1.1 <i>Human immunodeficiency virus type 1</i> : genoma e ciclo replicativo	8
1.2 Il gene Env	12
1.3 Ruolo dei corecettori nell'infezione da HIV	14
1.4 Corecettori e tropismo	16
1.4.1 Virus R5 linfocito-tropici ( <i>T cell-tropic</i> )	18
1.4.2 Virus R5 macrofago-tropici ( <i>M-tropic</i> )	19
1.4.3 Virus X4 <i>T-cell tropic</i>	20
1.5 Tropismo e compartimentalizzazione	22
1.5.1 Cellule mononucleate di sangue periferico (PBMCs)	22
1.5.2 Monociti	24
1.5.3 Sistema nervoso centrale	25
1.5.4 Compartimento genitale	26
1.5.5 Tessuto linfoide associato all'intestino (GALT)	27
1.6 Tropismo e stadi di infezione	27
1.7 Test fenotipici per la determinazione del tropismo	31

1.8 Test genotipici per la determinazione del tropismo	34
1.9 Tropismo e carica netta del V3 <i>loop</i>	38
1.9.1 Tropismo e sottotipi non-B	38
<i>Scopo del lavoro</i>	39
<b><u>2. MATERIALI E METODI</u></b>	40
2.1 I pazienti	40
2.2 Raccolta del plasma e separazione dei PBMCs	41
2.3 Separazione delle cellule CD4+ e CD14+	42
2.4 Estrazione dell'RNA virale e del DNA intracellulare	43
2.5 Trascrizione inversa e amplificazione	43
2.6 PCR	45
2.7 Sequenziamento del prodotto di amplificazione	46
2.8 Predizione del tropismo virale mediante <i>Geno2pheno coreceptor</i>	47
2.9 Analisi della carica netta di V3	48
2.9.1 Amplificazione del gene Env	49
2.9.2 Clonaggio di Env	52
2.9.3 Produzione di pseudovirus	55

2.9.4 Saggi luciferasici	56
2.9.5 Analisi statistica	57
<u>3. RISULTATI</u>	57
3.1 Distribuzione del tropismo nelle sottopopolazioni di pazienti	57
3.2. Scelta del <i>cut-off</i> di FPR e concordanza dei risultati	60
3.3 Prevalenza dei ceppi R5 nei compartimenti studiati	62
3.4. Associazione tra carica netta in V3 e FPR	65
3.5 Mutazioni in V3 associate al tropismo	71
3.6 Test fenotipico su campioni concordanti, discordanti e indeterminati	74
3.7 Risultati del test fenotipico	77
3.8 Analisi delle mutazioni in regioni diverse da V3	78
<u>4. DISCUSSIONE</u>	83
<u>5. BIBLIOGRAFIA</u>	91

## SOMMARIO

Lo studio del tropismo di HIV-1 è uno degli aspetti emergenti dello studio del virus per l'interesse che esso ricopre dal punto di vista patogenetico e diagnostico. Al momento dell'infezione della cellula ospite, il virus sfrutta la propria proteina di superficie gp120 per legarsi a un recettore cellulare, il CD4; tale interazione risulta essere necessaria, ma non sufficiente, per consentire l'entrata del patogeno nella cellula. Il legame della gp120 con una seconda molecola, ovvero un corecettore cellulare, è infatti essenziale per l'infezione e per la prosecuzione del ciclo replicativo. L'affinità del ceppo virale ai corecettori cellulari CCR5 e/o CXCR4 viene definita « tropismo ». Una delle armi attualmente a disposizione per impedire al virus di entrare nella cellula è rappresentata dal *Maraviroc*, nuovo farmaco approvato recentemente per l'uso clinico. Tale composto è in grado di agire in maniera competitiva con il corecettore, impedendo in tal modo l'infezione. Tuttavia, date le sue caratteristiche strutturali, il farmaco è in grado di legarsi esclusivamente al corecettore CCR5, rendendo inutile la somministrazione ai pazienti infetti con ceppi virali CXCR4 e DM (*Dual/Mixed* o R5/X4)-tropici e favorendone potenzialmente la selezione. Per tali ragioni, la corretta individuazione del tropismo virale è di cruciale importanza per l'identificazione dei pazienti candidati all'utilizzo degli inibitori del CCR5.

I test *fenotipici* e *genotipici* rappresentano due approcci diversi per la determinazione del tropismo di HIV-1. Il test fenotipico è a tutt'oggi considerato il metodo *gold standard*, in quanto l'intero gene codificante le proteine dell'*envelope* virale contribuisce al risultato del saggio. D'altra

parte, il test fenotipico presenta notevoli svantaggi legati a costi, tempi di esecuzione e difficoltà tecniche che ne limitano l'applicazione nella pratica diagnostica. Diversamente, il test genotipico si basa sul sequenziamento di una specifica regione di gp120 che prende il nome di *V3 loop*, un peptide di 35 amminoacidi essenziale nell'interazione con il corecettore. Tuttavia, l'analisi della sequenza, limitata a una sola regione della proteina, riduce la possibilità di ottenere informazioni, che si ipotizza risiedono nelle altre regioni geniche di Env, utili a una più accurata predizione.

Il lavoro della tesi si è collocato inizialmente nell'ambito dei due progetti nazionali OSCAR e DIVA, svolti rispettivamente nel 2009 e nel 2010, con lo scopo di sviluppare e validare un saggio diagnostico in grado di predire il tropismo di HIV-1 a partire dalla sequenza genetica di V3. Tali progetti hanno coinvolto tutti i principali Centri Virologici italiani e, visto il successo del progetto OSCAR nello sviluppare un saggio genotipico per la predizione del tropismo a partire da RNA virale, lo studio DIVA ha voluto implementare questo metodo applicandolo anche al DNA virale integrato. Le sequenze del *V3 loop* ottenute nel nostro laboratorio sono state quindi analizzate per cercare un'associazione fra alcune caratteristiche del peptide (quali carica netta e mutazioni amminoacidiche) e la predizione ottenuta mediante strumenti bioinformatici. Questa prima parte del lavoro è stata condotta su 144 sequenze di V3 ottenute da pazienti a diversi stadi di infezione e con differente trattamento farmacologico. Per aumentare la variabilità del campionamento sono state analizzate fonti virali diverse, ovvero plasma, cellule mononucleate di sangue periferico (PBMCs), linfociti T CD4+, monociti e biopsie gastriche. I risultati ottenuti hanno

confermato la già nota associazione fra la presenza di amminoacidi carichi positivamente nelle posizioni 11 e 25 di V3, ma anche l'associazione tra la carica netta (NC) del peptide e il tropismo ottenuto tramite il software *Geno2pheno* (G2P) *coreceptor*. E' stata inoltre evidenziata l'associazione fra il risultato di G2P e la presenza di residui positivi nella posizione 32.

Parte dei campioni in esame (rappresentativi della popolazione totale) sono stati in seguito sottoposti a test fenotipico, messo a punto nel Laboratorio di Immunopatologia dello CHUV di Losanna (Svizzera), coordinato dal Prof. G. Pantaleo, dove ho svolto parte di questo progetto nel periodo Gennaio-Settembre 2012 . Lo studio del fenotipo *in vitro*, seppur molto elaborato da un punto di vista tecnico, con lunghe tempistiche e costi sostenuti, si proponeva di verificare l'ipotesi che il gruppo di sequenze con valori di NC in V3 *border line* (compresa tra 4 e 5), fossero effettivamente in grado di utilizzare entrambi i corecettori cellulari. Tale verifica richiedeva l'utilizzo del saggio fenotipico quale metodo di riferimento per la determinazione del tropismo virale. Pertanto, l'intero gene Env è stato amplificato dai campioni selezionati ed utilizzato per la produzione di *pseudovirus* in grado di infettare cellule *reporter* esprimenti specificamente i corecettori CCR5 o CXCR4. In base alla loro capacità di infettare le diverse linee cellulari, gli pseudovirus sono stati classificati come R5, X4 o R5X4-tropici. Una volta stabilito il fenotipo delle varianti virali, esso è stato messo in relazione alla sequenza di Env. Il sequenziamento dell'intero gene ha consentito di aumentare la sensibilità e l'accuratezza della predizione al fine di verificare l'associazione tra il tropismo di G2P e le mutazioni in gp120 che sono state descritte in letteratura. È stato quindi

possibile elaborare un nuovo e sintetico modello di predizione assegnando un diverso punteggio ai marcatori associati al tropismo e lo *score* finale è stato confrontato con il saggio fenotipico quale test di riferimento. In questo modo è emerso che se l'identificazione dei ceppi X4 può essere supportata dalla presenza di chiari marcatori genotipici presenti nella sequenza, la discriminazione tra ceppi R5 puri e *dual*-tropici, in base al modello elaborato, risulta essere meno netta. In particolare, dai risultati si è evidenziato che nel 50% dei campioni con punteggio superiore alla soglia degli X4, ovvero in circa la metà dei campioni con carica netta di V3 compresa tra 4 e 5, le varianti virali erano in grado di utilizzare entrambi i corecettori, seppure con minore efficienza il CXCR4. Dal momento che i ceppi *dual*-tropici sono prevalenti durante la fase di *switch* corecettoriale da R5 a X4, l'identificazione di queste varianti costituirebbe un marcatore diagnostico aggiuntivo per il monitoraggio dell'infezione e la progressione della malattia HIV-associata.



## 1. INTRODUZIONE

### 1.1 *Human Immunodeficiency Virus type 1*: genoma e ciclo replicativo

Il genoma di HIV (Retrovirus associato allo Sindrome da Immunodeficienza Acquisita, AIDS) è costituito approssimativamente da circa 9,8 kB per un peso molecolare di circa 30 kDa. Esso è composto da due filamenti identici di RNA(+) che possiedono un'estremità 5' dotata di gruppo *cap* e un'estremità 3' poliadenilata. I due filamenti di RNA sono tenuti insieme da legami idrogeno che si stabiliscono tra sequenze ripetute invertite presenti alle estremità 5' delle catene. HIV, come gli altri Retrovirus, possiede tre geni strutturali per la replicazione: **Gag** (*group-specific antigen*), **Pol** (*polymerase*), **Env** (*envelope*), organizzati come mostrato in Fig. 1.

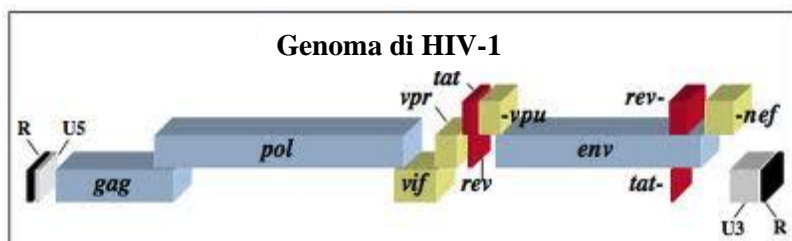
Il gene *Gag* codifica per un precursore proteico (p55) che viene successivamente “clivato”, attraverso meccanismi di proteolisi, in diversi componenti proteici strutturali interni al virione.

Il gene *Pol* codifica anch'esso per un precursore proteico, dal quale si originano successivamente gli enzimi essenziali per il ciclo replicativo: trascrittasi inversa, proteasi ed endonucleasi-integrasi.

Il gene *Env* codifica per un precursore proteico altamente glicosilato, la gp160. Questa glicoproteina, che risulta composta da 850 amminoacidi, viene scissa da un'endopeptidasi cellulare in due glicoproteine a più basso peso molecolare: la proteina extracellulare gp120 e la proteina transmembrana gp41. La gp120 svolge un ruolo prioritario nel

riconoscimento delle cellule CD4, mentre la gp41 è principalmente coinvolta nella fusione del virus con la cellula ospite.

Il genoma di HIV-1, oltre alle proteine codificate dai geni *gag*, *env* e *pol*, codifica per una serie di proteine strutturali con funzione regolatrice: *tat*, *rev*, *nef*, *vif*, *vpr* e *vpu*. La maggior parte di questi geni sono peculiari dei Retrovirus umani.



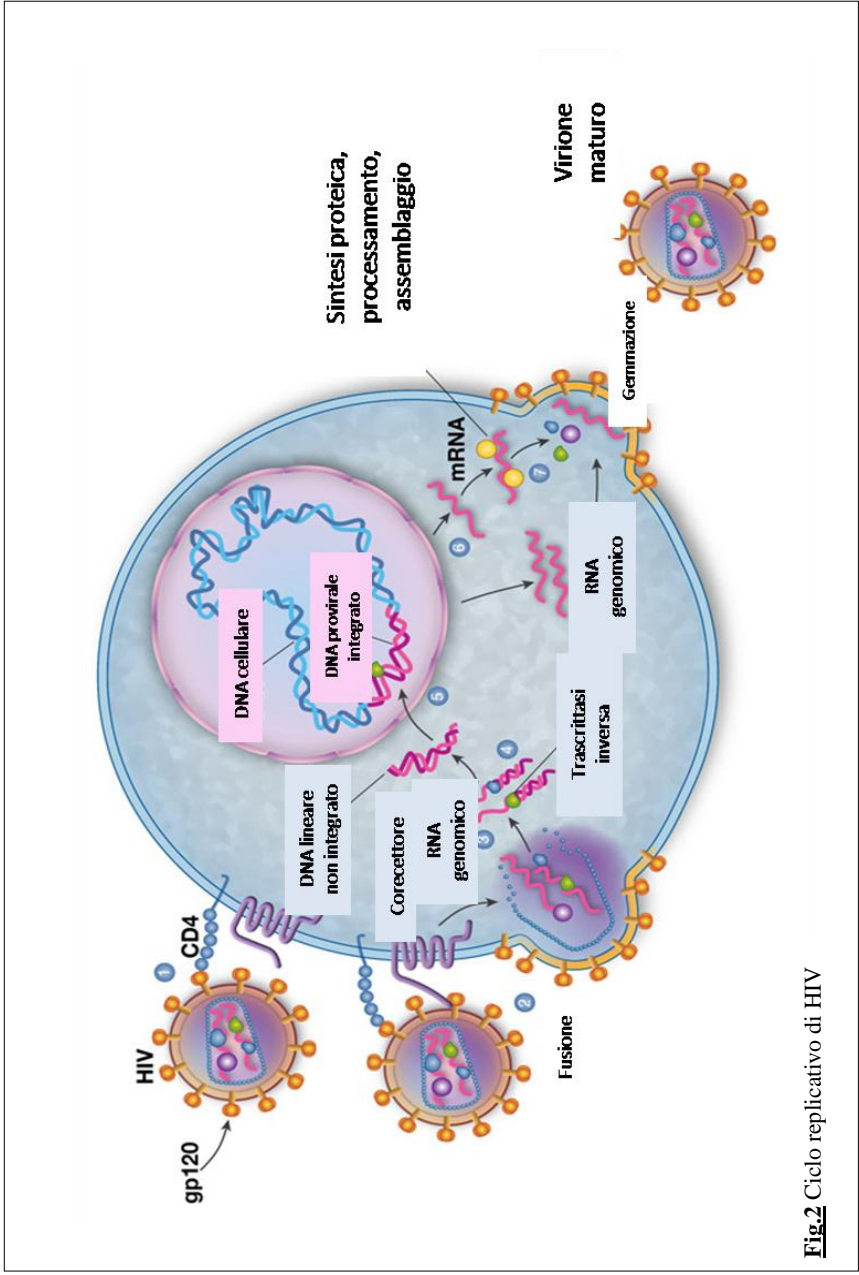
**Fig. 1** Organizzazione del genoma di HIV-1

Analogamente agli altri virus, il ciclo di replicazione di HIV può essere diviso in diverse fasi. Il primo *step* comprende gli eventi iniziali, ovvero adsorbimento del virione, penetrazione nella cellula ospite e scapsidazione. Nella seconda fase il genoma virale a RNA viene retrotrascritto in DNA ad opera della trascrittasi inversa, per poi assumere una forma “intermedia” costituita da una doppia catena lineare libera di DNA non ancora integrata nel genoma della cellula ospite. La successiva circolarizzazione e integrazione nel DNA cellulare si verificano mediante l’azione dell’integrasi e, una volta all’interno del DNA ospite, il genoma virale vi

rimane permanentemente associato, finché la cellula non termina il suo ciclo vitale.

Dal momento che il virus può non uccidere la cellula infettata, è verosimile che HIV possa rimanervi latentemente albergato. Il DNA virale integrato viene trascritto a opera di enzimi cellulari e i prodotti della trascrizione subiscono delle modifiche, attraverso eventi di *splicing*, che li rendono idonei per la traduzione e la sintesi delle proteine virali. Nelle prime fasi della trascrizione, nel nucleo della cellula infetta si ritrovano solo mRNA che codificano per le proteine con funzione regolatrice. L'espressione tardiva porta invece alla trascrizione di mRNA codificanti per le proteine strutturali necessarie alle successive fasi di montaggio e alla produzione di nuove particelle virali (Fig.2). Immediatamente dopo essere state prodotte, tutte le proteine strutturali deputate alla formazione del capsid si riuniscono al di sotto dello strato interno della membrana cellulare dove hanno origine le nuove particelle virali. Le particelle costituenti la progenie virale possono essere rilasciate attraverso una gemmazione lenta, che non comporta il danneggiamento immediato della cellula ospite, oppure attraverso un processo che altera l'intera membrana cellulare provocando danneggiamenti irreversibili nella cellula.

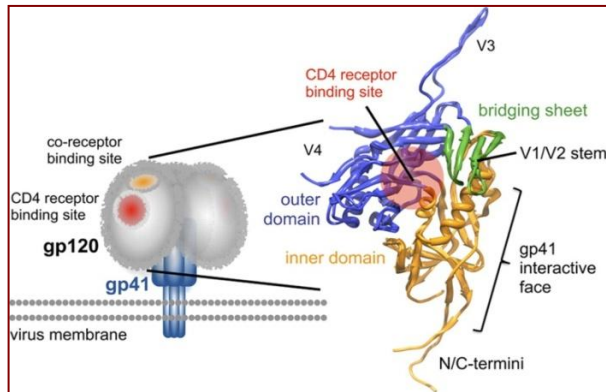
Alcune tappe del ciclo vitale del virus, dunque, dipendono strettamente dal corretto funzionamento della cellula ospite. Altre invece, come la retrotrascrizione del genoma virale o la successiva integrazione, sono regolate da funzioni virus-specifiche.



**Fig.2** Ciclo replicativo di HIV

## 1.2 Il gene Env

Il gene Env di HIV-1 codifica per l'unica proteina espressa dal virus in superficie, ovvero una glicoproteina di 160 kD che, dopo essere stata tradotta, viene processata dando origine ai peptidi gp120 e gp41, che rimangono associati non-covalentemente per formare una singola subunità del trimero esposto sul virione. La porzione C-terminale del trimero è rappresentata dalla gp41, la quale presenta un dominio extracellulare necessario alla fusione fra membrana plasmatica e virale. La subunità N-terminale (gp120) ha una struttura più complessa ed è organizzata in cinque regioni conservate (C1-C5) intersperse in regioni variabili, V1-V5 (Fig.3).



**Fig. 3** Struttura di gp120 [ref.1]

Alcuni residui amminoacidi nelle regioni costanti e nella porzione variabile V4 interagiscono con il recettore CD4 espresso dalle cellule bersaglio del virus, mentre il V3 loop è coinvolto nel legame con il corecettore. Due fondamentali caratteristiche della superficie esposta del trimero, ovvero la

gp120, influenzano fortemente le interazioni con la cellula: la sua variabilità nella sequenza proteica e i siti di glicosilazione che, legando carboidrati, ne mascherano la superficie.

La sequenza codificante della gp120 muta più facilmente (1-2% ogni anno) rispetto a ogni altra regione del genoma di HIV [2]. Tuttavia, tali variazioni sono concentrate in regioni geniche discrete, in modo da proteggere la struttura e la funzionalità proteica. La variabilità è influenzata principalmente dalla pressione selettiva esercitata dal sistema immunitario dell'ospite, ma alcune mutazioni nella sequenza possono anche alterare i meccanismi di interazione con recettori e corecettori cellulari. La pressione selettiva su Env da parte del sistema immunitario è dovuta soprattutto all'azione di anticorpi neutralizzanti. La risposta dell'ospite guida, infatti, guida la selezione di varianti alternative "escape", ovvero invisibili al sistema immunitario. Bersagli fondamentali di tali anticorpi sono le regioni variabili V1, V2, V4 e V5, ma poiché le mutazioni vengono introdotte *random*, l'alto tasso di variabilità di questi *loop* non è dovuta tanto a una maggiore propensione a mutare, quanto piuttosto al fatto che le mutazioni in tali regioni non compromettono la funzionalità delle proteine "escape" codificate.

In aggiunta alla sua variabilità, il tasso di glicosilazione di Env rappresenta un ulteriore fondamentale meccanismo di evasione della risposta anticorpale; la presenza di specifiche triplette amminoacidiche (N-X-S/T) nelle regioni variabili, infatti, ha il duplice effetto di aumentare la probabilità di duplicazioni e delezioni ma anche di codificare per siti di glicosilazione che "nascondono" la proteina al controllo del sistema

immunitario. La sequenza e la struttura di Env mediano, dunque, un equilibrio fondamentale tra l'ingresso del virus nella cellula ospite e l'evasione dalla neutralizzazione anticorpale.

### **1.3 Ruolo dei corecettori nell'infezione da HIV-1**

Alcuni anni prima dell'identificazione dei due corecettori coinvolti nell'entrata di HIV-1 nella cellula ospite, è stato dimostrato che le varianti del virus possono presentare differenti fenotipi *in vitro*. Nel 1987 Sakai et al. [3] isolarono, da uno stesso paziente, cloni virali con diversa citopatogenesi e differente tropismo per cellule T e monociti. Due anni dopo, alcuni ricercatori svedesi mostrarono che le varianti del virus presenti in pazienti HIV+ asintomatici replicavano molto lentamente nelle cellule immunitarie del sangue periferico ed erano incapaci di infettare linfociti T, mentre i virus isolati da pazienti con grave immunodeficienza replicavano rapidamente e con elevata efficienza nelle stesse linee cellulari [4]. Vennero quindi identificati i fenotipi *T cell-tropico* (T-tropico) e *Monocito-tropico* (M-tropico), i quali presentavano diversi effetti patogenetici: l'infezione della linea cellulare linfoblastoide MT-2 con ceppi virali T-tropici si traduceva nella formazione di estesi sincizi (indicatori di una attiva replicazione di HIV-1), mentre l'infezione da parte dei virus M-tropici andava incontro a fallimento. Tali osservazioni furono dunque determinanti per definire fenotipicamente i ceppi T-tropici come « syncitium-inducing » (SI), mentre le varianti M-tropiche vennero identificate come « nonsyncitium-inducing »(NSI). Sebbene i pazienti asintomatici albergassero prevalentemente virus NSI, nel 50% dei casi si osservava uno « switch » verso ceppi T-tropici, che si associava

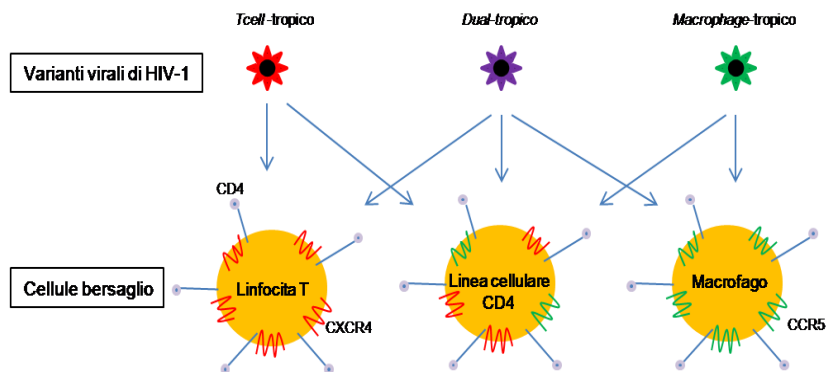
inevitabilmente a una progressione della malattia [5-10]. Per tali ragioni, verificare la capacità degli isolati virali di replicare in cellule MT2 sembrava un metodo estremamente efficace per distinguere le due varianti del virus, al punto che questo test venne introdotto come saggio prognostico per definire lo stadio dell'infezione; tuttavia, con l'introduzione nella pratica diagnostica del monitoraggio della carica virale, molti laboratori persero l'interesse per il saggio MT2 ai fini prognostici.

A metà degli anni '80, il CD4 fu il primo recettore identificato quale strumento chiave per l'ingresso di HIV nella cellula, ma fu subito chiaro che il legame al solo recettore CD4 non era sufficiente all'infezione [11]. Più di dieci anni dopo, nel 1996, venne identificato il corecettore CCR5 quale molecola coinvolta nell'interazione con la gp120 [12-13]. Poco tempo dopo fu scoperto l'altro corecettore, oggi noto come CXCR4, utilizzato dal virus nella fase di ingresso [14-15]. La scoperta dei corecettori fornì la spiegazione molecolare alla base dell'esistenza dei due fenotipi virali. Dal momento che le MT2 esprimevano esclusivamente il CXCR4, i virus T-tropici, fenotipicamente *syncitium-inducing* e MT2(+), erano dunque in grado di sfruttare questa molecola, mentre i virus M-tropici (NSI e MT2(-)) legavano esclusivamente il CCR5. Negli anni successivi, vennero individuati, oltre ai ceppi R5 e X4-tropici, i virus dual/mixed(DM)-tropici (R5X4), ovvero in grado di utilizzare entrambi i corecettori (Fig.4) e che caratterizzavano la fase di *switch* corecettoriale.

In seguito alla scoperta dei corecettori, fu osservata un'associazione tra il tropismo e la sequenza amminoacidica della regione variabile 3 [16-17],



porzione proteica della gp120. Tale osservazione spinse la ricerca verso la formulazione di un farmaco in grado di legare i corecettori e impedirne il legame con la glicoproteina virale. Con l'approvazione dell'inibitore del CCR5 *Maraviroc* per l'uso clinico, si rese necessaria l'introduzione, nella *routine* diagnostica, di test genotipici e fenotipici in grado di predire il tropismo e, di conseguenza, la suscettibilità virale al farmaco.

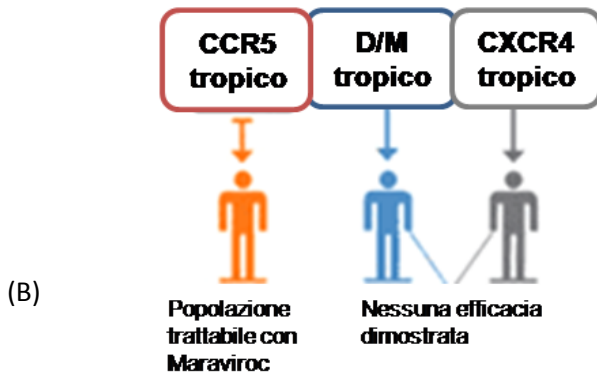
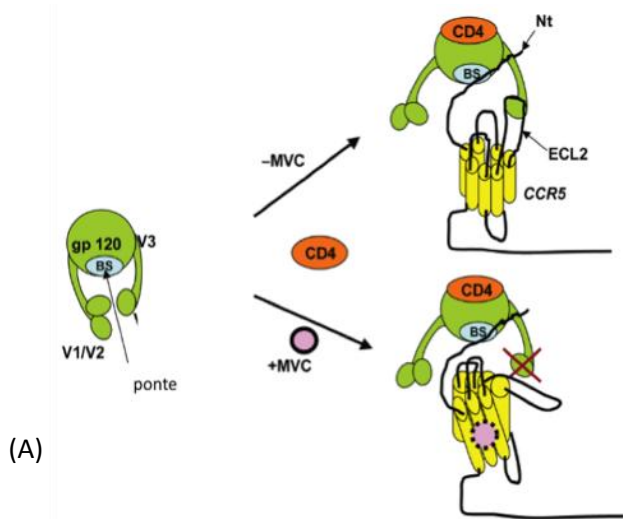


**Fig.4.** Fenotipi di HIV-1

Inoltre, la capacità del Maraviroc di inibire esclusivamente la replicazione di ceppi R5-tropici (Fig.5 A,B) ha aumentato la necessità e l'interesse per lo studio dei corecettori nella patogenesi.

#### 1.4 Corecettori e tropismo

Come descritto, la proteina Env è la molecola deputata a interagire con il CD4 e a subire una serie di cambiamenti conformazionali che consentono



**Fig.5** (A) Meccanismo di azione del Maraviroc: MVC si lega al CCR5 alterandone la conformazione e impedendo al virus R5-tropico di interagire con il corecettore. (B) Gli studi clinici MOTIVATE 1-2 hanno dimostrato che il MVC è efficace esclusivamente nei pazienti infetti con ceppi R5- tropici.

la fusione delle membrane con rilascio del *core* virale nel citoplasma della cellula. Dal momento che la maggior parte degli eventi di trasmissione del virus avvengono attraverso le superfici mucosali, che la replicazione attiva di HIV ha luogo nei tessuti linfoidei e che le manifestazioni cliniche della malattia si osservano in molteplici organi e tessuti, è deducibile che il virus abbia l'opportunità di venire a contatto con differenti tipi cellulari. Pertanto, gran parte della ricerca sulla patogenesi di HIV-1 è centrata su Env e sul suo ruolo nell'ingresso del virione. È dunque necessario, ancora oggi, caratterizzare le diverse varianti in grado di infettare linee cellulari distinte, capirne la prevalenza nel corso dell'infezione e il ruolo chiave nella malattia.

Per 25 anni la teoria della comunità scientifica si è fondata sulla presenza di due fenotipi virali, legati alla capacità degli isolati di crescere in linee di linfociti T e di indurre la formazione di sincizi. La dicotomica classificazione, universalmente accettata, in virus *syncytium-inducing*/X4/linfocito-tropici e virus *non-syncytium-inducing*/R5/macrofago-tropici non palesava però il fatto che soltanto una piccola ma importante proporzione di virus R5 era in grado di infettare i macrofagi; si rese dunque necessaria una ulteriore distinzione fra ceppi CCR5 con tropismo per linfociti T e macrofagi.

#### 1.4.1 Virus R5 linfocito-tropici (*T cell-tropic*)

I linfociti T CD4<sup>+</sup> attivati sono la principale fonte di replicazione di HIV-1 [18]. Le cellule T vengono attivate nell'ambito della risposta immune e il loro rapido metabolismo rende la replicazione virale più efficiente. Queste

cellule sono generalmente infettate dal ceppo virale che viene trasmesso, che persiste nell'ospite per anni e che molto spesso causa immunodeficienza. Il fenotipo di questi virus è CCR5 e l'ingresso nella cellula target è garantito da elevati livelli di espressione di CD4 nei linfociti.

#### 1.4.2 Virus R5 macrofago-tropici (*M-tropic*)

Alcuni isolati virali che replicano nei linfociti T CD4+ possono infettare anche i macrofagi. Un clone virale comunemente utilizzato in laboratorio, Ba-L, è stato il primo isolato M-tropico. I virus M-tropici sono stati spesso isolati nel cervello di pazienti HIV+ con disturbi neurocognitivi, così come nel liquido cerebrospinale; in questi siti, i virus M-tropici giocano un ruolo diretto nella patogenesi [19]. Nel cervello, i ceppi M-tropici replicano nei macrofagi perivascolari o nella microglia, cellule simili ai macrofagi presenti nel parenchima cerebrale. Diversamente, i linfociti T sono pressoché assenti nel sistema nervoso centrale, ma in caso di infiammazione possono essere richiamati ad attraversare la membrana emato-encefalica e diventano fonte di attiva replicazione per i ceppi R5.

Le varianti M-tropiche sono state isolate da autopsie di tessuto cerebrale o dal liquor; nelle fasi avanzate dell'infezione, esse possono essere rilevate come varianti minori nel sangue [20], nonostante la loro origine anatomica non sia stata ancora definita. Sebbene l'infezione dei macrofagi richieda al virus diversi *steps* adattativi, i ceppi virali M-tropici sono in grado di infettare cellule con bassi livelli di CD4, che normalmente viene meno espresso nei macrofagi rispetto ai linfociti T attivati. Alcuni studi hanno

evidenziato, quindi, delle variazioni di sequenza di Env nel sito attivo di legame con il CD4, nelle regioni fiancheggianti V3 ma anche in V1-V2 e gp41, che permetterebbero a queste varianti di legare con alta affinità il CD4 laddove questo venga espresso a bassi livelli; tuttavia, i determinanti genetici per il tropismo macrofagico non sono ancora stati chiaramente definiti.

#### 1.4.3 Virus X4 T-cell tropic

L'evoluzione delle varianti che possono utilizzare il corecettore CXCR4 è un aspetto ancora dibattuto. In particolare, è ancora poco chiaro se queste varianti vengano selezionate al momento della trasmissione, se possano replicare attivamente solo in un ospite con immunodeficienza e se rappresentino la causa o la conseguenza della rapida progressione della malattia. Inoltre, è stato dimostrato in alcuni casi che i differenti sottotipi possono avere una diversa propensione all'evoluzione delle varianti X4. Storicamente, i virus X4 erano concepiti come varianti più patogeniche degli R5 e che comparivano tardivamente nel corso della malattia. Questa idea era principalmente legata alla loro capacità di indurre la formazione di sincizi. Tuttavia, tale affermazione fu parzialmente smentita dall'osservazione che alcuni ceppi R5-tropici (considerati normalmente *non syncitium-inducing*) possono crescere in linee di linfociti T immortalizzati e possono causare la formazione di sincizi se il CCR5 è espresso. Dunque esistono ancora almeno due principali aspetti ancora poco chiari riguardanti i virus X4. Innanzitutto occorre definire, come già accennato, se i virus X4 possano causare una più rapida progressione della malattia o siano un marcatore dell'aumentata compromissione del sistema

immunitario. In secondo luogo, se i virus *dual*-tropici continuano a usare anche il CCR5 *in vivo* o se soltanto il CXCR4 sia biologicamente rilevante. La maggior parte degli isolati X4 sono *dual*-tropici, dunque questo aspetto è fondamentale per una miglior comprensione del loro ruolo nel paziente. Una ulteriore osservazione dalle cause ancora poco chiare riguarda il fatto che gli X4 sono prevalenti nei soggetti trattati che hanno fallito la terapia antiretrovirale (ART).

Ancor prima dell'identificazione dei corecettori di HIV-1, era noto che i cambiamenti nella sequenza del V3 *loop* definivano le caratteristiche fenotipiche dei virus X4-tropici, prevalenti negli stadi tardivi dell'infezione. La maggior osservazione del cambiamento fenotipico da R5 a X4 è la comparsa di amminoacidi basici (carichi positivamente) nelle posizioni 11, 24 e 25 di V3 [21, 22, 23] .

---

21

D'altra parte, la variabilità complessiva di questa regione è maggiore qualora siano presenti le mutazioni nei suddetti residui, suggerendo che esse siano il risultato di un processo evolutivo più articolato. L'introduzione di amminoacidi carichi positivamente e l'aumentata diversità di V3 nei ceppi X4, ha reso possibile lo sviluppo di strumenti bioinformatici in grado di predire il tropismo sulla base della sequenza primaria di V3. Tuttavia, lo *switch* da tropismo R5 a X4 è dovuto a cambiamenti di sequenza anche all'esterno di V3. Virus ricombinanti ottenuti in laboratorio hanno dimostrato il coinvolgimento di altre regioni di Env [24], sebbene l'interpretazione dei risultati ottenuti sia complicata da diversi fattori. In particolare, la capacità di entrare nelle cellule utilizzando il CXCR4 può essere molto variabile, soprattutto nei ceppi

virali in grado di usare entrambi i corecettori; pertanto, la rilevanza biologica *in vivo* dell'uso di CCR5 e/o CXCR4 nei ceppi *dual*-tropici deve ancora essere definita.

## **1.5 Tropismo e compartmentalizzazione**

*In vivo*, HIV-1 è in grado di infettare numerosi tipi cellulari, purché essi esprimano il CD4 e uno dei noti corecettori maggiormente usati dal virus. Il CCR5 è espresso nei linfociti T *memory*, nei linfociti T CD4+ attivati ma anche in macrofagi, cellule dendritiche e microgliali. La frequenza delle cellule T *memory* CCR5+ è relativamente alta nei tessuti linfoidi associati all'intestino (*gut-associated lymphoid tissues*, GALT). L'espressione di CXCR4, invece, è preponderante nei linfociti T *naïve* e si riduce a seguito del loro differenziamento e attivazione. CXCR4 è anche espresso nei linfociti T CD8+, nelle cellule B, nei neutrofili e negli eosinofili. Considerata la diversa espressione dei corecettori nei sottogruppi di cellule T, la replicazione residua di virus R5 e X4-tropici potrebbe avere differente impatto sull'omeostasi dei linfociti T durante la ricostituzione del sistema immunitario.

### **1.5.1 Cellule mononucleate di sangue periferico (PBMCs)**

Nei Paesi industrializzati, la maggior parte dei pazienti trattati con ART ha una carica virale irrilevabile nel plasma e le strategie di gestione degli effetti collaterali della terapia prevedono spesso l'introduzione degli antagonisti del CCR5. Al fine di caratterizzare il tropismo nel DNA provirale, sono stati quindi condotti alcuni studi che hanno dimostrato come la proporzione di ceppi R5 nel DNA possa variare dal 48% nei

pazienti pluritrattati al 92% nei pazienti *naïve* (non sottoposti a trattamento farmacologico) con infezione primaria [25-26]. Questa proporzione è simile a quella stimata nel plasma e, considerato che nei pazienti *naïve* il tasso di discordanza del tropismo tra RNA plasmatico e DNA è del 4.8-10%, si può concludere che vi è notevole corrispondenza nell'uso del corecettore fra i due compartimenti. Alcuni studi, inoltre, hanno osservato che nei pazienti trattati con ART e con carica virale irrilevabile per lungo tempo, c'è un'evoluzione delle quasi-specie virali nel DNA verso tropismo CXCR4.

Differenti meccanismi possono guidare l'evoluzione del tropismo nei *reservoirs* durante la ART e lo *switch* R5-X4 del DNA nei “serbatoi” dell'ospite è un vivo argomento di discussione. Il ripristino del sistema immunitario in risposta alla ART dipende dall'incremento dei linfociti T *naïve*. Queste cellule potrebbero rappresentare un *reservoir* importante nei pazienti infetti con ceppi X4, dal momento che possono esserne facilmente infettate. Il virus presente nei *reservoirs* può comunque replicare a bassi livelli o essere rilasciato sporadicamente; si ipotizza che l'uno, l'altro o entrambi i fenomeni siano all'origine della “viremia residua”. I linfociti T *naïve* contribuirebbero a tale fenomeno mediante due meccanismi: da una parte, l'espansione delle cellule T e il rilascio del virus dagli archivi di ceppi X4 potrebbe portare a una progressiva sostituzione del *reservoir* T *memory*; dall'altra, la replicazione residua dei pazienti sottoposti ad ART potrebbe consentire l'evoluzione genetica delle quasispecie e il cambio dell'uso del corecettore.



### 1.5.2 Monociti

Diversi studi hanno dimostrato chiaramente la capacità del virus di infettare i macrofagi *in vivo*, che rappresentano uno dei principali compartimenti di “riserva” delle varianti virali. Tali evidenze hanno focalizzato l’attenzione sul ruolo dei precursori dei macrofagi, ovvero i monociti, nell’infezione da HIV. I monociti possono essere distinti in tre classi: monociti “classici”, esprimenti alti livelli di CD14 e privi di CD16 (CD14<sup>++</sup>,CD16<sup>-</sup>), monociti “intermedi” con alti livelli di CD14 e ridotta espressione di CD16 (CD14<sup>++</sup>,CD16<sup>+</sup>) e monociti “non-classici” con bassi livelli di CD14 ed elevata espressione di CD16 (CD14<sup>+</sup>,CD16<sup>++</sup>). Nei donatori sani, circa l’85% dei monociti è rappresentato dai “classici” (CD14<sup>++</sup>, CD16<sup>-</sup>). Numerosi studi *in vitro* hanno dimostrato che i linfociti isolati da sangue periferico non vengono infettati produttivamente da HIV-1 ma l’infezione avviene dopo il loro differenziamento in macrofagi, circa 24 ore in seguito all’isolamento [27]. La resistenza dei monociti all’infezione da HIV-1 è dovuta verosimilmente al blocco di diverse fasi del ciclo replicativo virale. Un blocco dell’entrata si potrebbe verificare se la densità del CD4 o di CCR5/CXCR4 fosse molto bassa nei monociti; tuttavia, in queste cellule, il CD4 è espresso a livelli apprezzabili mentre i corecettori sono espressi solo da alcune sottopopolazioni monocitarie e a bassi livelli, che impedirebbero un’infezione produttiva. Un modello alternativo si basa su una teoria secondo la quale l’infezione da HIV-1 nei monociti non è bloccata, in quanto il meccanismo descritto sopra contribuisce soltanto a rallentare il ciclo replicativo del virus. Considerando che i monociti persistono nel sangue soltanto per pochi

giorni prima di morire o migrare nei tessuti e differenziare in cellule dendritiche o macrofagi, potrebbe essere necessario il differenziamento per avere la produzione di virus da parte del monocita precedentemente infettato. A conferma di tale teoria, diversi studi *in vivo* hanno dimostrato la presenza di HIV-DNA nei monociti isolati da sangue periferico [28-29]. Data la scarsa capacità del virus di replicare efficacemente nei monociti, è improbabile che la replicazione in queste cellule contribuisca in modo sostanziale all'aumento della carica virale; tuttavia, considerata la loro capacità di entrare nei tessuti e permanervi a lungo, è plausibile che essi svolgano un ruolo determinante per la disseminazione del virus in siti anatomico-tissutali dove hanno origine popolazioni virali indipendenti.

Uno dei principali limiti di tali studi *in vitro* è rappresentato dal fatto che non vengono discriminate le tre diverse popolazioni di monociti. Dal momento che l'85% della popolazione monocitaria *in vivo* è rappresentato da monociti CD14<sup>++</sup>CD16<sup>-</sup>, la maggior parte degli esperimenti viene condotta su queste cellule. Tuttavia i CD14<sup>++</sup>CD16<sup>-</sup> sono per lo più indifferenziati, esprimono i più bassi livelli di CCR5 e sono i meno suscettibili all'infezione *in vitro*. Diversamente, le popolazioni di monociti esprimenti CD16 sono invece più permissive all'infezione e dunque rappresenterebbero un miglior modello da utilizzare.

### 1.5.3. Sistema nervoso centrale

La maggior parte dei ceppi di HIV che infettano i macrofagi usano il corecettore CCR5. Per questo, i virus R5-tropici sono predominanti nei tessuti in cui monociti e macrofagi sono molto diffusi, ad esempio nel

tessuto cerebrale. Come accennato nel paragrafo 1.5.2, i macrofagi rappresentano un importante serbatoio di infezione, teoria supportata dall'evidenza che gli isolati virali di pazienti con encefaliti e demenza da HIV sono R5-tropici. I virus rilevati nel liquido cerebrospinale hanno origine sia dal sistema nervoso centrale che da sangue periferico. Questo compartimento sarebbe quindi un importante sito di incontro e ricombinazione di virus provenienti dal sistema nervoso e circolatorio, talora con tropismo differente. Tale aspetto assume dunque decisiva importanza in caso di somministrazione di antagonisti del CCR5.

#### 1.5.4 Compartimento genitale

Alcuni studi filogenetici hanno riportato che i compartimenti genitali costituiscono anch'essi dei serbatoi di infezione con ceppi virali diversi da quelli presenti nel sangue dei pazienti. Per quanto riguarda i virus presenti nel liquido seminale, non è stata evidenziata una specifica compartimentalizzazione del tropismo, dal momento che i virus presenti in questo sito possono avere indifferentemente tropismo X4 o R5. Analogamente, molti studi hanno riportato che nel tratto genitale femminile esiste una chiara divergenza genetica rispetto ai ceppi virali ematici, ma con una preponderante replicazione in questo sito di virus X4-tropici. Essi deriverebbero da virus in attiva replicazione nel sangue, ma sarebbero poi selezionati positivamente nelle secrezioni femminili. Nonostante tali evidenze, alcuni studi filogenetici sulla trasmissione di HIV hanno riportato che esiste un "collo di bottiglia" che selezionerebbe favorevolmente i ceppi R5, per cui queste varianti dovrebbero essere

prevalenti nei compartimenti genitali. Dal momento che tali risultati sono stati confutati in altri lavori, tale aspetto rimane ancora da chiarire.

#### 1.5.5. Tessuto linfoide associato all'intestino (GALT)

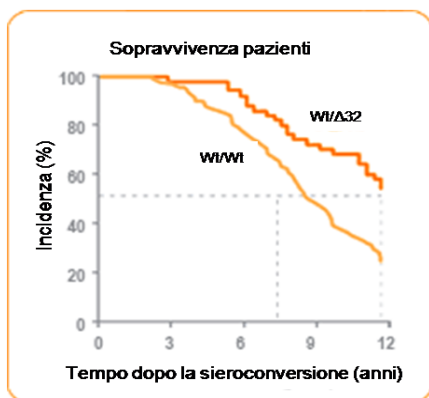
Il GALT è il tessuto linfoide più esteso del corpo umano ed è costituito da un'ampia popolazione cellulare CD4+ esprimente CCR5. È stato riportato che HIV persiste in questo tessuto anche in presenza di una ART efficace. Un recente studio ha chiaramente dimostrato la mancanza di compartimentalizzazione delle quasispecie di HIV-1 tra circolazione ematica e GALT attraverso l'analisi delle sequenze C2-V3. Gli autori dello studio hanno concluso che la mucosa intestinale non rappresenta un sito "santuario" per la replicazione di HIV-1 ma si evidenzia un libero scambio di virus tra sangue ed epitelio intestinale.

### **1.6 Tropicismo e stadi di infezione**

Dal punto di vista della patogenesi, i virus R5 vengono rilevati nella maggior parte delle infezioni primarie, mentre i ceppi X4 emergono tardivamente. Di conseguenza il corecettore CCR5 sembra giocare un ruolo fondamentale nella trasmissione del virus e durante le prime fasi della malattia. Ulteriori dati a supporto di questa teoria derivano dall'identificazione di una mutazione nel gene codificante la molecola CCR5 (CCR5-Δ32) presente nel 10% della popolazione caucasica. L'omozigotità per questa mutazione è estremamente rara (1%) ma è associata alla totale mancanza di espressione della proteina e a resistenza all'infezione da HIV-1. L'eterozigotità della mutazione è invece associata

a bassi livelli di espressione del CCR5 e a una ridotta progressione della malattia (Fig.6).

I primi studi sull'uso del corecettore nelle infezioni recenti (ovvero in cui il paziente era infetto da massimo sei mesi) si basavano sui risultati del saggio sulle MT2 e riportavano quasi esclusivamente la presenza di ceppi NSI, ossia R5-tropici. I successivi saggi fenotipici, basati sulla produzione di pseudovirus ricombinanti, rilevarono invece la presenza di ceppi X4 o *dual*-tropici in circa il 3% dei pazienti con infezione primaria. Diverse percentuali sono state riportate nel caso dei test genotipici, i quali stimano una proporzione di virus X4 in pazienti recentemente infettati pari a circa il 13%. Tuttavia, alcuni studi [30] hanno evidenziato che alcuni metodi di interpretazione della sequenza possono sovrastimare la proporzione di X4 rilevati, pertanto tali percentuali possono variare dal 5 al 20%.



**Fig.6.** Curva di sopravvivenza dei pazienti con CCR5-Δ32

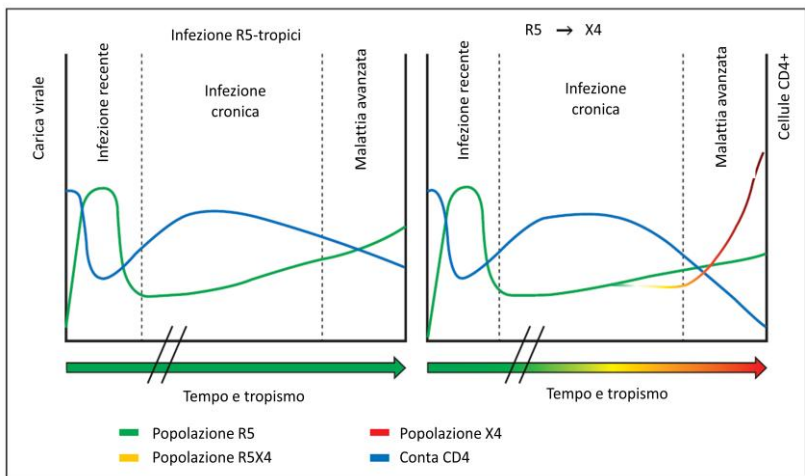
La prevalenza nell'uso del corecettore è stata studiata ampiamente anche nei soggetti *naïve* in procinto di iniziare la terapia. La percentuale di X4

osservata in questo gruppo si aggira intorno al 10% nei pazienti infetti da almeno 4 anni e con un relativamente alto numero di cellule CD4+, mentre nei pazienti in stadio avanzato con CD4<200cell/uL o clinicamente sintomatici, la percentuale di X4 aumenta, oscillando tra il 38% e il 52% [31]. E' interessante sottolineare che fra i pazienti con malattia in fase avanzata, la proporzione di virus X4 supera raramente il 50% ; la metà restante è rappresentata da pazienti infettati con ceppi R5-tropici e quindi candidati all'utilizzo del Maraviroc.

Sulla base di quanto riportato in letteratura si sostiene generalmente che, nel corso dell'infezione, le varianti X4-tropiche si evolvano dai pre-esistenti virus R5. Studi longitudinali con metodiche di sequenziamento ad alta sensibilità (*ultra-deep sequencing*) hanno mostrato che la transizione R5/X4 segue un percorso « multistep », che consiste nell'accumulo di mutazioni in regioni coinvolte nel legame al corecettore (principalmente V3) e che ha come risultato il cambio di tropismo. Altri lavori hanno dimostrato che le varianti X4 possono essere archiviate nel DNA cellulare già nelle prime fasi dell'infezione e che queste possano riemergere in seguito a pressione selettiva. I pazienti HIV+ che sono andati incontro a *switch* del corecettore mostrano un accelerato aumento della carica virale, riduzione delle cellule T e una rapida progressione verso l'AIDS (Fig.7).

I meccanismi alla base dello *switch* sono ancora scarsamente conosciuti e non è del tutto chiaro se l'evoluzione verso il tropismo X4 sia causa o conseguenza della progressione della malattia. Fra le varie teorie, si ipotizza che le varianti X4 emergano a seguito della deplezione di cellule immunitarie CCR5+ o di un'umentata espressione del CXCR4 da parte

dei linfociti ; entrambi i fattori sarebbero dovuti a un avanzamento della malattia. D'altra parte, è anche plausibile che virus X4-tropici siano meglio riconosciuti immunologicamente rispetto agli R5 tropici, a tal punto da potersi replicare attivamente soltanto in seguito a una significativa compromissione del sistema immunitario, che a questo punto fallirebbe nell'identificarli.



**Fig.7.** Rappresentazione schematica delle tre fasi di infezione da HIV-1: recente, cronica e avanzata, in relazione all'uso del coreettore. (A) Pazienti infettati con virus R5-tropici mostrano un costante aumento della carica virale e lenta riduzione dei linfociti T CD4+. (B) Pazienti in cui compaiono varianti X4 mostrano rapido aumento della carica virale e declino dei linfociti T CD4+.

Un ulteriore aspetto da chiarire riguarda il potenziale ruolo del tropismo nella risposta alla terapia antiretrovirale combinata (cART). Alcuni studi in passato avevano dimostrato che i pazienti infetti con ceppi MT2(-) avevano un maggior numero di CD4+ durante i primi 6 mesi di trattamento rispetto ai pazienti MT2(+). La prevalenza di virus X4 è stata associata, dunque, a

una più scarsa risposta alla terapia farmacologica [32]. Contrariamente, altri studi hanno fallito nel riportare qualunque correlazione tra l'uso del corecettore e la risposta alla cART. In uno studio più recente di Waters et al. [33], è stato osservato su 402 pazienti che la presenza di ceppi DM o X4-tropici aveva un effetto negativo sulla conta dei CD4+ e un aumentato rischio di sintomatologia clinica, ma non influenzava la risposta al trattamento.

La determinazione del tropismo è un prerequisito fondamentale per l'utilizzo degli inibitori del CCR5. È dunque fortemente raccomandato che la somministrazione di tali farmaci venga preceduta da un approccio fenotipico o genotipico che possa stabilire l'affinità corecettoriale dei ceppi circolanti nel paziente HIV+.

### **1.7 Test fenotipici per la determinazione del tropismo**

Come accennato precedentemente, il primo saggio fenotipico elaborato per la determinazione del tropismo virale era basato sull'infezione della linea cellulare linfoblastoide MT2, in grado di identificare i virus che inducevano la formazione di sincizi (X4-tropici) e i *nonsyncitium-inducing* (R5). Questo test richiede l'iniziale produzione di virus *in vitro* a partire da cellule mononucleate di sangue periferico (PBMCs) isolate dal paziente; i virioni ottenuti vengono quindi utilizzati per infettare le MT2 ed osservare l'effetto citopatico. In alternativa, i PBMCs del paziente possono essere co-coltivati direttamente con le cellule.

Sebbene tale sistema di determinazione del tropismo sembrava essere estremamente promettente e affidabile in quanto fondato sull'osservazione



pratica dell'effetto citopatico, esso presenta alcuni importanti limiti. Innanzitutto la fase iniziale di coltivazione del virus a partire da cellule immunitarie del paziente potrebbe selezionare dei ceppi adattati alla replicazione *in vitro* e quindi diversi dal virus prevalente *in vivo*, riducendo in tal modo l'affidabilità del risultato. In secondo luogo, l'analisi delle varianti ottenute dai PBMCs potrebbe non riflettere le caratteristiche della popolazione virale presente nel compartimento plasmatico, dando origine a dubbi sulla possibilità di ottenere effetti benefici dall'utilizzo del *Maraviroc*. Inoltre, il saggio sulle MT2 non distingue fra ceppi X4 puri e DM-tropici, in quanto entrambi presentano lo stesso fenotipo *syncytium-inducing*. Infine, i lunghi tempi di incubazione delle MT2 e, di conseguenza, dell'esecuzione del test ne limitano fortemente l'utilizzo nella pratica clinica.

I saggi basati sulla produzione di pseudovirus sono stati sviluppati per evitare la selezione di ceppi indesiderati nelle fasi di coltivazione *in vitro*. Gli pseudovirus vengono ottenuti con le sequenze di Env amplificate dal plasma del paziente e sono utilizzati per infettare cellule *reporter*. La produzione di questi virioni si ottiene trasfettando cellule epiteliali note come HEK293 con il vettore di espressione di Env insieme a un vettore Env-difettivo, ma contenente un gene per la luciferasi. Gli pseudovirus prodotti vengono quindi utilizzati per inoculare cellule *reporter*, quali la linea U87 derivante da cellule di glioma umano.

I trials clinici sugli antagonisti del CCR5 si sono basati sulla determinazione del tropismo di HIV-1 attraverso il test fenotipico *Trofile*, il quale sfrutta la metodologia degli pseudovirus. La soglia di sensibilità

minima della prima versione di questo test (*Original Trofile Assay*, OTA) nella rilevazione delle varianti minori X4 era circa il 10%. In seguito, l'OTA fu sostituito *dall'Enhanced Sensitivity Trofile Assay* (ESTA), una versione la cui sensibilità arriva a 0.3% di rilevazione dei ceppi X4 minoritari. La sensibilità nella rilevazione delle varianti minori è un parametro fondamentale che distingue il saggio fenotipico dai test genotipici, la cui soglia di identificazione delle specie minoritarie non scende al di sotto del 15-20%. L'importanza di rilevare tali varianti è legato alle osservazioni riportate in alcuni studi in cui è stato dimostrato che la replicazione attiva delle specie virali presenti a bassi livelli prima del trattamento può condurre al fallimento virologico. Tuttavia, la soglia al di sopra della quale le varianti X4 minori possono portare al fallimento virologico, in regime di terapia antiretrovirale basata su antagonisti del CCR5, deve ancora essere stabilita. Due diversi studi hanno dimostrato, attraverso *ultra-deep pyrosequencing*, che le varianti X4 presenti oltre il 2% della popolazione totale influenzano negativamente la risposta al trattamento con *Maraviroc*. Tuttavia, non tanto la percentuale di X4 quanto il numero assoluto di virioni X4-tropici potrebbe essere rilevante per predire il fallimento virologico.

Uno dei principali vantaggi del test fenotipico consiste anche nell'essere in grado di determinare il tropismo dei sottotipi non-B. La maggior parte degli algoritmi genotipici, infatti, sono stati sviluppati per il sottotipo B, il ceppo più diffuso in Nord-America ed Europa. Tuttavia il sottotipo C è il più diffuso a livello mondiale e le forme ricombinanti del virus stanno emergendo anche nei Paesi industrializzati. I ceppi non-B rappresentano

circa il 20% delle forme di HIV-1 circolanti in Europa e tale aspetto deve essere considerato nella gestione degli antagonisti del CCR5. Studi contrastanti riportano una diversa affidabilità dei test genotipici nella predizione del tropismo nei sottotipi non-B, rendendo necessarie ulteriori correlazioni fra questi test e il fenotipo *in vitro*.

Sebbene i test genotipi siano ormai entrati a far parte delle *routine* diagnostica grazie ai costi ridotti, ai brevi tempi di esecuzione e alle limitate difficoltà tecniche, il fenotipo sugli pseudovirus ricombinanti rimane una pietra miliare nella ricerca e nella prospettiva di trials clinici per gli antagonisti del CCR5. Il fenotipo consente, inoltre, di avere informazioni sul tropismo derivanti dall'intera sequenza di Env, diversamente dal test genotipico che si basa esclusivamente sulla sequenza di V3. Le regioni esterne al V3 non solo influenzano la predizione del tropismo, ma potrebbero essere di fondamentale importanza nell'individuare i *patterns* di resistenza agli antagonisti del CCR5. Alcune mutazioni esterne al V3 loop, infatti, sembrano essere coinvolte nella resistenza al *Maraviroc*, rendendo il test fenotipico l'unico metodo attualmente in grado di identificare tali *patterns*.

### **1.8 Test genotipici per la determinazione del tropismo**

I saggi genotipici si basano sulla predizione *in silico* dell'uso del corecettore a partire da informazioni contenute nella sequenza virale, mediante sequenziamento del genoma con metodi di clonaggio o di popolazione. Nel primo caso, le singole varianti virali ottenute da paziente vengono clonate, mentre nel sequenziamento di popolazione (*bulk-*

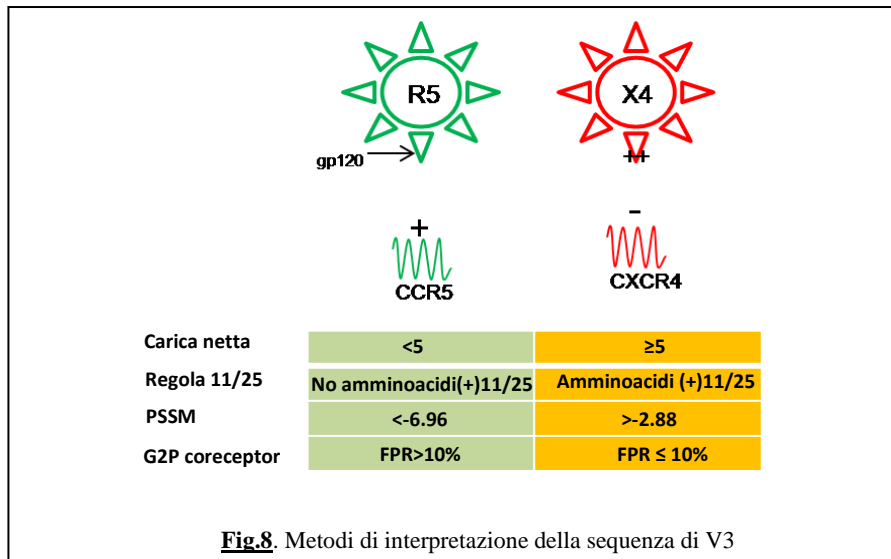
*sequencing*) tutte le quasispecie del virus vengono amplificate ed elaborate in un'unica sequenza. Il *bulk sequencing* è di gran lunga più accessibile per la *routine* diagnostica ma non è da escludere che verrà presto o tardi sostituito da metodi basati sul clonaggio, quali *l'ultra-deep sequencing* o il *single genome sequencing*, che tuttavia rimangono ancora degli strumenti d'élite soprattutto a causa dei costi elevati.

Rispetto ai test fenotipici, i metodi basati sul genotipo sono più semplici dal punto di vista tecnico, più rapidi, più economici e standardizzabili. La predizione genotipica del tropismo tiene conto delle informazioni contenute nella regione di Env maggiormente rilevante nell'uso del corecettore, V3 appunto. Nel caso del clonaggio, sequenze senza ambiguità vengono sottoposte all'analisi, mentre il *bulk sequencing* può introdurre dei nucleotidi "misti", ovvero posizioni in cui due o più basi differenti possono essere rilevate. Le prime mutazioni in V3 associate al tropismo furono amminoacidi basici (carichi positivamente) nelle posizioni 11 e/o 25. A questa semplice regola, si aggiunse l'osservazione che la carica netta del peptide risultava essere più elevata nel caso di virus X4 e che anche il numero dei siti di glicosilazione in gp120 poteva correlare con il tropismo (Fig.8). Sebbene un buon numero di ceppi virali venga correttamente classificato con queste approcci, l'introduzione di strumenti bioinformatici ha reso l'interpretazione della sequenza più immediata.

I metodi bioinformatici più utilizzati per l'interpretazione della sequenza di V3 sono il *position-specific scoring matrix* (PSSM) e il *Geno2Pheno* (*G2P*) *coreceptor*. Entrambi sono definiti *machine learning methods*, ovvero algoritmi che, a partire da un insieme di dati (*database*), ne

estrapolano delle caratteristiche comuni, ossia dei predittori che vengono poi ricercati nella sequenza sottoposta ad analisi. Un principio comune di questi metodi si basa sul “training” del modello su cui si basa il software, eseguito attraverso il confronto dei risultati ottenuti con test genotipico e fenotipico su uno stesso campione. Le banche dati utilizzate per il *training* possono essere il *Los Alamos HIV Sequence Database* o particolari coorti di studi clinici, quali la *HOMER cohort* [34]. Mentre il *Los Alamos Database* fornisce dei dati di elevato standard qualitativo sul genotipo e fenotipo ottenuti con metodi di clonaggio, i risultati della coorte HOMER si basano su *bulk sequencing*, e quindi si applicherebbero meglio alla diagnostica di *routine* (che sfrutta lo stesso metodo). L’analisi con PSSM sfrutta la sequenza primaria del *V3 loop* e calcola, per ogni singolo amminoacido, la probabilità che essa derivi da un virus X4. La predizione basata su algoritmi, quali G2P, invece, è più accurata e la flessibilità di questi sistemi consente di inserire delle informazioni aggiuntive che possono migliorare il modello, quali informazioni strutturali utili alla simulazione funzionale del *V3 loop*. La predizione di PSSM e di G2P è associata a un punteggio, ovvero a un’informazione quantitativa. PSSM dà una somma dei punteggi per tutte le posizioni in cui le sequenze con lo *score* più alto hanno maggior somiglianza. In G2P il risultato R5 o X4 è associato a un *false positive rate value* (FPR, probabilità di falso positivo), che indica la probabilità che un virus R5 sia classificato erroneamente come variante X4. La scelta di un diverso *cut-off* di FPR ha un decisivo impatto sul risultato. Se *cut-off* meno stringenti (ovvero più elevati) riducono infatti la probabilità di classificare erroneamente un virus R5 come X4, aumenta anche la probabilità di ottenere falsi X4 positivi.

Sebbene non sia stato mai eseguito un trial clinico che valuti il corretto *cut-off* di FPR, sarebbe auspicabile un'analisi a posteriori dei dati delle diverse coorti studiate per stabilire il valore soglia ottimale.



Gli algoritmi basati su *machine learning methods* sono certamente caratterizzati da una maggiore sensibilità di predizione (40/60%) rispetto alle tradizionali regole di interpretazione di V3, ma forniscono una predizione binaria, non discriminando le varianti DM-tropiche. Inoltre, il *bulk sequencing* ha la capacità di rilevare le varianti minori in un campione se presenti almeno per il 15-20% all'interno della popolazione. Per aumentare la sensibilità nella rilevazione di queste specie nei campioni di DNA e di plasma con carica virale <1000 copie/mL, alcuni gruppi hanno condotto studi sul tropismo eseguendo il test genotipico in triplicato, ma con risultati non del tutto concordanti. In alcuni casi, l'analisi in triplicato con un FPR al 10%, aumenta l'incidenza di predizione X4 del 4%, mentre

in altri si osserva la stessa proporzione rispetto al test in singolo [35-36]. Per verificare se il test in triplicato sia in grado di aumentare la sensibilità delle varianti minori, sarebbe necessario uno studio prospettico che confronti il risultato del test in singolo/triplicato e *l'outcome* del trattamento con Maraviroc.

## **1.9 Tropismo e carica netta del V3 loop**

L'evidenza che la carica netta del peptide V3 fosse associata al tropismo è riportata in alcuni lavori [37], sebbene la spiegazione biologica di questa relazione non sia ancora completamente nota. Recentemente, Yokoyama et al. [38] hanno osservato, attraverso sofisticate tecniche di simulazione dinamica molecolare, che nei virus in cui la carica netta del V3 è più bassa (NC= +3), la porzione di gp120 bersaglio di anticorpi neutralizzanti è meno esposta rispetto ai virus in cui NC=+7. Di conseguenza, i virus con bassa NC subirebbero una minor pressione selettiva da parte del sistema immunitario e un minor tasso di variabilità in Env.

### **1.9.1 Tropismo e sottotipi non-B**

Un aspetto tanto importante quanto dibattuto riguarda la maggior propensione di alcuni sottotipi non-B di evolvere in ceppi X4. È noto che in circa il 50% dei pazienti infetti con sottotipo B si ha uno *switch* del corecettore da R5 a X4 con tempi variabili tra i soggetti. Il sottotipo D, che filogeneticamente è vicino al ceppo B, sembra avere una maggior propensione allo *switch*, mentre il sottotipo C sarebbe meno incline, nonostante vi siano in proposito degli studi discordanti. Dal momento che tutti i virus possono evolvere in varianti X4 con tempistiche diverse, è

plausibile che sia la distanza evolutiva fra il sottotipo e le caratteristiche genetiche dei virus X4-tropici a determinare la rapidità del cambiamento. Inoltre, l'*HIV GRADE network* ha analizzato i dati ottenuti dalla *routine* diagnostica sottolineando le differenze nella distribuzione del tropismo fra i vari sottotipi. In questo caso l'uso del corecettore è stato predetto con G2P e un *cut-off* di FPR del 10%. Dallo studio è emerso che esiste una significativa differenza fra la prevalenza di X4 nei sottotipi CRF01\_AE e D rispetto ai sottotipi A1 e G. Per i ceppi F e B non è stata invece osservata una differenza significativa nella prevalenza di R5 e X4 rispetto alla distribuzione media dei sottotipi. Per quanto riguarda il sottotipo C vi sono dei dati discordanti, probabilmente dovuti al fatto che i sistemi di interpretazione delle sequenze sono stati sviluppati su *datasets* di campioni B. Un approccio utile potrebbe essere la scelta di un diverso *cut-off* di FPR, in base al singolo sottotipo.

### ***Scopo del progetto***

L'obiettivo iniziale del lavoro è stato lo sviluppo e la validazione di un test diagnostico basato sul sequenziamento di V3 per la predizione del tropismo di HIV-1, indispensabile all'individuazione dei pazienti candidati all'utilizzo del Maraviroc. L'analisi dei campioni ottenuti nel nostro laboratorio è stata quindi il punto di partenza per l'identificazione di marcatori genetici aggiuntivi utili a una più accurata determinazione del tropismo virale. Attraverso il confronto di diversi metodi, è stato possibile ottenere delle ulteriori indicazioni per una migliore interpretazione delle sequenze di V3 ed Env.



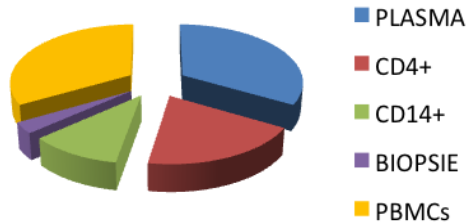
## 2. MATERIALI E METODI

### 2.1 I pazienti

Lo studio è stato eseguito su 144 campioni provenienti da 105 pazienti HIV+, di cui 100 afferenti al “Policlinico Umberto I” di Roma per controlli di *routine* e 5 seguiti presso lo CHUV, Ospedale Universitario Cantonale di Losanna (Svizzera).

I pazienti erano sottoposti a diverso trattamento terapeutico e da essi sono stati prelevati differenti tipi di campione biologico. In particolare, 44 pazienti erano *naïve* al trattamento antiretrovirale, 11 ART- *responder*, 22 ART *non-responder*, 6 in sospensione e 3 non aderenti alla terapia; di 19 pazienti il dato clinico non è stato disponibile. Inoltre, il 34% dei campioni è stato ottenuto da plasma, 35% da PMBCs, 18% da cellule T CD4+, 10% da monociti CD14+ e il 3% da biopsie gastriche (Fig.9). Di 26 pazienti erano disponibili campioni provenienti da diversi compartimenti (quali cellule CD4+, CD14+, tessuto gastrico bioptico) che sono stati tutti inclusi nell’analisi.

Ai fini del nostro studio non si è ritenuto necessario individuare un gruppo omogeneo di pazienti; inoltre, la scelta di considerare diversi tipi di prelievi biologici e pazienti con differente trattamento farmacologico e stadio di infezione ha contribuito ad aumentare l’eterogeneità del campionamento e la significatività dei risultati. La maggior parte delle sequenze (80%) è stata classificata come sottotipo B, sebbene siano stati riscontrati anche i sottotipi C, CRF02\_AG e F1.



**Fig. 9** Tipi di campione biologico inclusi nell'analisi

## 2.2 Raccolta del plasma e separazione dei PBMCs

I campioni di sangue periferico raccolti in tubi con anticoagulante (EDTA) sono stati sottoposti ad una prima centrifugazione di 10 minuti a 1800 rpm. Il plasma è stato prelevato, aliquotato e conservato a  $-80^{\circ}\text{C}$ .

L'isolamento delle cellule mononucleate di sangue periferico (PBMCs) viene eseguito mediante centrifugazione su gradiente di densità, utilizzando del Lympholyte-H secondo le seguenti modalità: dopo aver aspirato il plasma, si aggiunge del PBS in un rapporto di 1:1 alla porzione rimanente di campione e si rende il tutto omogeneo. La soluzione di sangue e PBS viene successivamente stratificata su un volume di Lympholyte-H pari a  $\frac{1}{2}$  di quello della soluzione, cosicché il rapporto finale sarà di due parti di campione su una di Lympholyte-H. Dopo una centrifuga di 20 minuti a 2200 rpm si ottiene un disco ben visibile di PBMCs, che viene prelevato con una pipetta; le cellule

vengono lavate due volte con del PBS e centrifugate in modo da ottenere un pellet secco di circa  $5 \times 10^6$  cellule.

### **2.3 Separazione delle cellule CD4+ e CD14+**

Alcuni esperimenti sono stati condotti sulle popolazioni linfocitarie CD4+ e CD14+. Per separare i due tipi cellulari, i PBMCs ottenuti dopo separazione in base alla densità sono stati separati sfruttando i marcatori di superficie cellulare e per mezzo di appositi kit basati sulla tecnologia del “magnetic cell sorting” (MACS MILTENYI BIOTEC ).

La metodica si basa sull'utilizzo di particelle paramagnetiche coniugate con anticorpi monoclonali (anti CD4 e anti CD14) specifici per le cellule di interesse.

Queste biglie magnetiche hanno una dimensione di circa 50nm, sono biodegradabili e non interferiscono sulla vitalità cellulare grazie alla loro composizione di ossido di ferro e polisaccaride.

La sospensione cellulare viene in un primo tempo incubata per 15 minuti con le biglie “anti CD14”, dal momento che il recettore CD4 è presente su entrambe le popolazioni cellulari CD4+ e CD14+; successivamente la miscela cellule-biglie viene lavata con del buffer, contenente PBS, 0.5% di BSA e EDTA 2mM.

A questo punto la miscela viene fatta passare attraverso una colonna porosa posta all'interno di un magnete; le cellule non marcate attraversano rapidamente la colonna e costituiscono la popolazione di PBMCs-CD4+, mentre rimuovendo la colonna dal magnete è

possibile ottenere le cellule CD14<sup>+</sup> che erano state trattate all'interno della stessa dalla forza del campo magnetico. I CD14<sup>+</sup> vengono successivamente centrifugati in modo da ottenere un pellet secco. La soluzione contenente i CD4<sup>+</sup> viene trattata allo stesso modo. Le popolazioni CD4<sup>+</sup> e CD14<sup>+</sup> così ottenute possono essere conservate a -80°C fino al momento dell'uso.

## **2.4 Estrazione di RNA e DNA virale**

L'RNA virale è stato estratto da 140 µL di plasma utilizzando il metodo "QIAamp Viral RNA Mini Kit" (Qiagen, Milano), in base al protocollo standardizzato. L'RNA trattenuto sulla membrana di gel di silice è stato eluito in 60 µL di buffer AVE.

Il DNA è stato estratto da 5 X 10<sup>6</sup> PBMC utilizzando il "Dneasy Tissue kit" (Qiagen, Milano); dopo un'incubazione *overnight* con il buffer di lisi e la proteinasi K a 56°C, l'estrazione vera e propria viene eseguita seguendo il protocollo descritto nel kit "Nucleo Spin Tissue" (Macherey-Nagel, Milano), che utilizza colonnine contenenti una membrana di silice, per catturare l'acido nucleico. Il DNA viene quindi eluito con 60 µL di buffer BE.

## **2.5 Trascrizione inversa e amplificazione per PCR**

Mediante saggio di genotipizzazione è stato possibile amplificare le sequenze della regione V3, interna al gene Env. A tal scopo sono stati utilizzati dei primers in grado di amplificare V3 nei diversi sottotipi:

V3S2: 5' – CagCACAgTACAATgTACACA – 3'

V3AS5: 5' – CTTCTCCAATTgTCCCTCA – 3'

I campioni di RNA estratti vengono aggiunti alle provette di reazione in cui avviene sia la trascrizione inversa che l'amplificazione mediante PCR. La reazione di trascrizione da HIV-RNA in cDNA avviene ad opera della trascrittasi inversa (RT) del virus della leucemia murina. Una DNA-polimerasi termostabile è stata utilizzata per la reazione di amplificazione.

Il protocollo ha previsto la preparazione di due Master Mix (MM):

- MM I = RT-PCR Primers + dNTP + DTT + RNase-Inhibitor.
- MM II = RT-PCR Buffer + RNase Inhibitor + RT enzyme +DNA Polymerase.

Le provette vengono inserite in un termociclatore dopo aver dispensato la Master Mix I; dopo la denaturazione del campione a 90°C per 2 minuti e dopo 5 minuti a 50°C, si procede con l'aggiunta della Master Mix 2.

La reazione di amplificazione prevede:

- 20 cicli, ciascuno costituito da un passaggio a 94°C per 30 secondi, un passaggio a 57°C per 30 secondi e un passaggio a 94°C per 90 secondi;
- 15 cicli, costituiti da un passaggio a 94°C per 30 secondi, un passaggio a 60°C per 30 secondi, un passaggio a 70°C per 2 minuti, un ciclo a 70°C per 7 minuti e uno a 4°C per 60 minuti.

## 2.6 PCR

Per l'amplificazione della regione V3 da DNA è stato messo a punto un protocollo di Nested-PCR, che ha quindi richiesto l'utilizzo di due coppie di primers: la prima costituita dagli stessi oligo utilizzati nella RT-PCR (V3S2 e V3AS5), mentre la seconda costituita da V3S6 (5'-CTgTTAAATggCAGTCTAgC -3') e V3AS3bis (5'-CAATTTCTgggTCCCCTC -3'). I primers sono stati utilizzati a una concentrazione finale di 0.1uM in un volume di reazione di 50uL. Per la reazione di Nested-PCR è stato utilizzato 1 uL del prodotto di PCR. In Tab.1 sono riportati i profili termici per l'amplificazione.

	PCR	Nested-PCR
Oligo FW	V3S2	V3AS5
Oligo RV	V3S6	V3AS3bis
Profilo termico	94°C x 5min      1 ciclo 93°C x 45 sec } 53°C x 60 sec } 35 cicli 72°C x 60 sec } 72°C x 7 min      1 ciclo	94°Cx5min      1 ciclo 93°C x 45 sec } 55°C x 60 sec } 35 cicli 72°C x 60 sec } 72°C x 7 min      1 ciclo

**Tab.1** Amplificazione della regione V3

Successivamente il prodotto della Nested-PCR è stato corso in gel di agarosio 2% per verificare la riuscita della reazione. Il prodotto della PCR è stato purificato con il kit di purificazione per PCR Microcon (Millipore).

## 2.7 Sequenziamento del prodotto di amplificazione

Dopo aver amplificato il tratto del gene Env contenente la sequenza V3, il prodotto di reazione è stato sequenziato utilizzando l'“Open kit Siemens Healthcare Diagnostics”. Questa reazione di sequenziamento segue i principi della tecnica sviluppata nel 1977 da Fred Sanger, una metodica di sequenziamento di tipo enzimatico, conosciuta come “metodo della terminazione di catena” o più semplicemente *metodo di Sanger*. Il principio della tecnica sviluppata da Sanger si basa sulla sintesi *in vitro* di DNA eseguita in presenza di nucleotidi modificati, ovvero *dideossiribonucleotidi trifosfato* che interrompono la sintesi della catena nel momento in cui vengono aggiunti ad essa. Grazie a questo sistema, si ottengono dei frammenti di dimensioni diverse che consentono di ricostruire la sequenza stampo. Per ogni campione di cui è stato sequenziato il dominio V3 del gene Env sono state preparate 4 provette di reazioni ognuna contenente: 2.8 µl di soluzione tampone CLIP, 8.8 µl di acqua, 2.8 µl di DNA del prodotto di Nested-PCR, 4.4 µl di enzima Thermosequenasi (32 U/µl) e 2.8 µL di primers V3S6 + V3AS3bis 3µM, gli stessi usati per la nested-PCR ma opportunamente marcati con i fluorocromi Cy5.0 e Cy5.5 per consentire la successiva lettura dei frammenti ottenuti con amplificazione. Inoltre, in ognuna delle quattro provette sono stati aggiunti separatamente 3 µl dei quattro *dideossinucleotidi* trifosfato, quantità stechiometricamente inferiori ai deossiribonucleotidi, in modo da permettere una elongazione del filamento sufficiente per l'analisi. La reazione di sequenziamento è stata impostata con un primo ciclo a 94°C per 5 minuti, seguita da 30 cicli comprendenti

uno primo step di denaturazione a 94°C per 20 secondi, un secondo step di *annealing* a 55.5°C per 20 secondi e un terzo step di allungamento a 70°C per 1 min. In seguito ai 30 cicli è stato introdotto uno step di 70°C per 7 min e, dopo l'aggiunta di 6 µl di Stop Loading Dye (buffer essenziale per bloccare la reazione), la temperatura è stata portata a 94°C x 4min per la denaturazione del DNA. I frammenti generati da questa reazione sono stati separati in base alla loro lunghezza in gel di poliacrilammide all'interno di un OpenGene® DNA System che consente la lettura automatizzata e l'elaborazione della sequenza in analisi.

## **2.8 Predizione del tropismo virale mediante *geno2pheno coreceptor***

I test genotipici per la predizione del tropismo di HIV-1 dipendono dall'abilità di interpretare le sequenze del *V3 loop* ottenute da ciascun campione; esse sono state esaminate, nel nostro studio, sfruttando il database e il software di allineamento *G2P coreceptor* (<http://coreceptor.bioinf.mpi-inf.mpg.de/>), algoritmo online che consente di ottenere una predizione del tropismo virale grazie ad allineamento multiplo di sequenze depositate nel database con affinità corecettoriale nota. Dopo aver sottomesso nel sistema la sequenza nucleotidica, il software effettua un allineamento multiplo; la somiglianza tra la *consensus* del database (con tropismo CCR5) e la sequenza *query* (con tropismo ignoto) è espressa con un valore percentuale di FPR (par.1.8). Per poter effettuare l'allineamento, *G2P* richiede all'operatore l'impostazione di uno specifico *cut-off*, ovvero una soglia (che varia da 1% a 20%) in base alla quale le sequenze con FPR inferiore vengono classificate come virus X4-tropici. La scelta del *cut-off* determina la stringenza di analisi, ovvero un aumento di



FPR da 10% a 20% comporta un aumento della sensibilità di identificazione di virus X4, ma anche della percentuale di *falsi* X4-tropici. Alcuni studi riportano che il miglior cut-off di FPR per la predizione del tropismo mediante G2P sia di 5.75% [30,39] in quanto tale soglia permetterebbe di avere un'elevata corrispondenza con il risultato del test fenotipico. D'altra parte, le *European Guidelines on the clinical management of HIV-1 tropism testing* (pubblicate in *The Lancet* a Marzo 2011) mettevano in risalto l'esigenza di discutere il limite di FPR nella *routine* diagnostica per aumentare la sensibilità di individuazione degli X4, in quanto un falso negativo (ovvero un virus X4 identificato come R5) poteva essere causa di fallimento terapeutico in caso di somministrazione non opportuna del *Maraviroc*. Fra le conclusioni delle linee guida veniva suggerito di aumentare il cut-off al 20% in caso di analisi in singolo di DNA provirale o di plasma con carica virale inferiore alle 1000 copie/mL, mentre per campioni con VL>1000 copie/mL si suggeriva un cut-off del 10%. I campioni dello studio qui riportato sono stati analizzati in base alle linee guida europee del 2011, impostando cut-off diversi (10% e 20%), dipendentemente dal campione in esame.

## **2.9 Analisi della carica netta di V3**

In seguito alla scoperta dei corecettori, negli anni '90 vennero pubblicati due studi in cui si ipotizzava l'esistenza di una relazione tra la carica netta (NC) del peptide V3 e il tropismo [40-41]. Sebbene non fosse chiara la spiegazione biologica di tale collegamento e non lo sia tutt'oggi, altri autori hanno confermato che una carica netta "positiva" (>4) sia associata prevalentemente a tropismo X4, mentre cariche inferiori possono essere

ricondotte a tropismo R5 [37]. Alla luce di queste osservazioni, obiettivo principale del nostro studio è stato confrontare le predizioni sul tropismo ottenute con G2P e la carica netta dei 144 campioni in analisi, per definire meglio una associazione fra i due parametri. La NC del V3 è stata calcolata mediante il *Peptide Property Calculator* (PPC, Innovagen), un software online che consente di stabilire la NC di sequenze peptidiche a pH 7. L'algoritmo del PPC Innovagen calcola la NC in base ai valori dei residui carichi come segue:

N-terminus	9.69	C-terminus	2.34
K, Lys	10.5	D, Asp	3.86
R, Arg	12.4	E, Glu	4.25
H, His	6.00	C, Cys	8.33
		Y, Tyr	10.0

e mediante la seguente formula:

$$Z = \sum_i N_i \frac{10^{\text{pKa}_i}}{10^{\text{pH}} + 10^{\text{pKa}_i}} - \sum_j N_j \frac{10^{\text{pH}}}{10^{\text{pH}} + 10^{\text{pKa}_j}}$$

Dove  $N_i$  sono i numeri, la  $\text{pKa}_i$  è la costante di dissociazione della catena N-terminale e delle catene laterali di arginina, lisina e istidina. L'indice "j" si riferisce invece alla porzione C-terminale e ai residui di acido aspartico, acido glutammico, cisteina e tirosina.

### 2.9.1 Amplificazione del gene Env

Alcuni campioni all'interno della coorte in esame sono stati analizzati anche a partire dalla sequenza di Env. La scelta è stata condizionata dalla

disponibilità del materiale biologico di partenza ma determinata soprattutto dalla concordanza dei risultati fra carica netta e G2P. Sedici campioni (11 DNA e 5 RNA) sono stati amplificati effettuando una trascrizione inversa+PCR per i campioni di RNA e una PCR (1<sup>st</sup> e 2<sup>nd</sup> round) per i campioni di DNA.

La trascrizione inversa è stata eseguita con *random primers* e *Superscript III Reverse Transcriptase* (Invitrogen) in base alle indicazioni riportate nel kit:

- 1) Preparazione della Mix 1
- 2) 5' a 90°C
- 3) 1' in ghiaccio
- 4) Aggiunta della Mix 2
- 5) 5' a 25°C
- 6) 60' a 50 °C
- 7) 15' a 70 °C

Mix 1

Random primers	1 uL
dNTPs mix 10mM	1 uL
H2O sterile	6 uL
RNA campione	5 uL

Mix 2

Buffer 5x	4 uL
DTT 0.1 m	1 uL
RNase Inhibitor	1 uL
Superscript III RT	1 uL

In seguito all'inattivazione della RT, è stato aggiunto 1 uL di *E.coli* RNase H per rimuovere l'RNA complementare dal cDNA neosintetizzato.

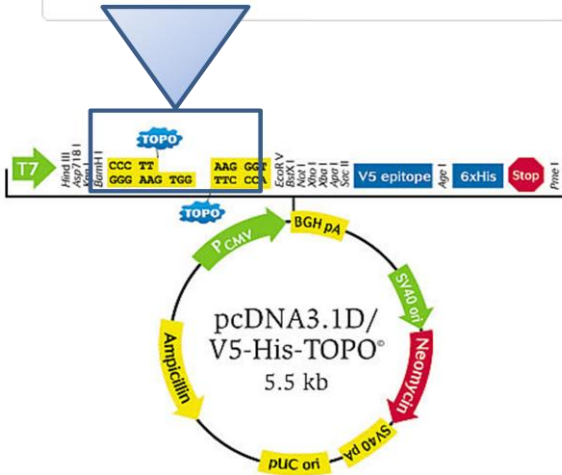
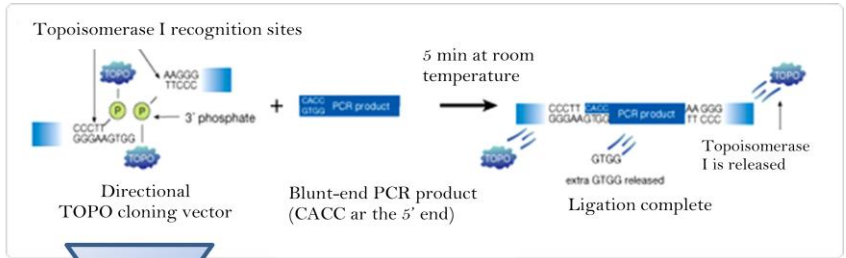
Per la reazione di PCR (sia a partire da DNA estratto che da cDNA) sono stati opportunamente disegnati dei primers degenerati in grado di amplificare la regione di interesse anche dei sottotipi non B (ENV1Adeg 5'-CACCGGCTTAGGCATYTCCTATGGCAGGAAGAA-3'; ENV1Mdeg 5'-TADCCCwTCCAGTCCCCCCTTTTCTYTTA-3'). Per ciascuna reazione sono stati aggiunti controlli negativi e i prodotti di amplificazione di 3200 bp venivano visualizzati in gel di agarosio 0.5%. Qualora il primo step di PCR non consentisse di amplificare il campione, è stato effettuato un 2<sup>nd</sup> round con dei primers interni (ENV1A 5'CACCGGCTTAGGCATCTCCTATGGCAGGAAGAA-3'; ENV2 5'-GCTCCATGTTTTTCCAGGTC -3'). I profili termici utilizzati sono riportati in 51 Tab.2.

1 <sup>st</sup> round		2 <sup>nd</sup> round	
94°C x 3'	1 ciclo	94°C x 3'	1 ciclo
94 °C x30''		94 °C x30''	
57 °C x15''	10 cicli	55 °C x15''	10 cicli
68 °C x3'		68 °C x3'	
94 °C x30''		94 °C x30''	
57 °C x15''	35 cicli	55 °C x15''	35 cicli
68 °C x3.20'		68 °C x3.20'	
68°C x30'	1 ciclo	68°C x30'	1 ciclo

**Tab.2.** Amplificazione del gene Env

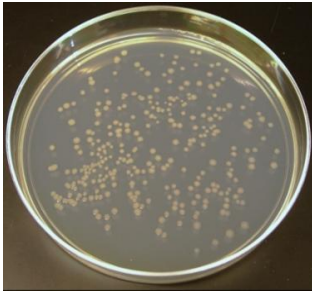
## 2.9.2 Clonaggio di Env

I prodotti di PCR sono stati purificati per il clonaggio in un plasmide di espressione. I prodotti di PCR con una singola banda (in gel di agarosio 0.5%) sono stati purificati mediante il kit commerciale “StrataPrep PCC Purification Kit” (Stratagene), mentre negli amplificati con bande multiple, il prodotto di interesse è stato estratto da gel mediante “PCR clean-up gel extraction” (Macherey-Nagel). La purificazione è stata necessaria per ottenere un prodotto sufficientemente puro (spettrofotometro,  $A_{260}/A_{280}=1.8$ ) da poter essere clonato all’interno di un plasmide di espressione. A tal fine, nel nostro studio è stato utilizzato il “pcDNA 3.1 TOPO expression vector” (Invitrogen). Il clonaggio all’interno di questo vettore è basato sul sistema della topoisomerasi, covalentemente legata all’estremità del plasmide e che provvede alla ligazione qualora l’insero abbia un’estremità “CACC” al 5’ (Fig.10). La ligazione avviene a 23°C x 30’ all’interno di una miscela costituita da 25 ng di inserto Env, 15 ng di pcDNA 3.1 TOPO expression vector, 1µL di soluzione salina e 1 µl di H<sub>2</sub>O (fornite nel kit – pcDNA 3.1 Directional TOPO Expression Kit). Due microlitri della ligazione si aggiungono a 50µL di E.coli chimicamente competenti che, dopo essere stati trasformati con il plasmide legato all’insero, vengono seminati in una piastra con ampicillina per la selezione delle colonie che hanno inserito il plasmide (codificante un gene per la resistenza a questo antibiotico, Fig. 11A). In seguito alla crescita *overnight* a 37°C delle colonie, esse sono state nuovamente poste in coltura in 3mL di LB 4X diluito in H<sub>2</sub>O, in modo tale da poter proseguire con l’estrazione di plasmide e verificarne la positività

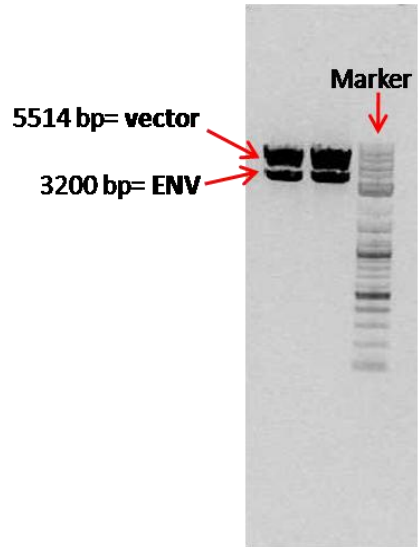


**Fig. 10.** pcDNA 3.1 TOPO expression vector

mediante restrizione. Il sistema di clonaggio utilizzato, infatti, non permette di discriminare tra colonie che hanno inserito il plasmide vuoto dalle colonie con plasmide+inserto. Pertanto si è resa necessaria l'estrazione dei plasmidi dalle singole colonie e la successiva restrizione con enzimi specifici (BamHI e XhoI); ciascun prodotto di restrizione è stato corso in gel di agarosio 0.5% per verificare la presenza della banda di interesse corrispondente all'amplificato di Env. (Fig.11B).



(A)



(B)

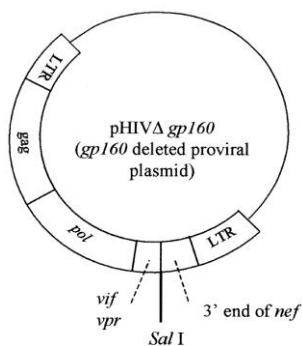
**Fig.11 (A)** Selezione delle colonie con il plasmide mediante ampicillina: solo le colonie che sono state trasformate sono in grado di crescere in presenza dell'antibiotico.

**Fig 11 (B)** Prodotti di restrizione positivi con BamHI e XhoI: le colonie positive presentano una banda a circa 5514 bp, che rappresenta il plasmide originario, e una a circa 3200 bp che rappresenta l'inserto, ovvero il prodotto di Env.

I plasmidi contenenti l'inserto sono stati sequenziati in modo da ottenere le singole varianti virali presenti nel campione. Per il sequenziamento sono stati utilizzati i primers T7 e BGH specifici per il plasmide e forniti nel kit di clonaggio (T7 5'-TAATACGACTCACTATAGGG-3' ; BGH Reverse 5'-TAGAAGGCACAGTCGAGG-3').

### 2.9.3 Produzione di pseudovirus

Le colonie positive alla restrizione sono state coltivate in 100 mL di LB 4x diluito in H<sub>2</sub>O e il DNA plasmidico è stato estratto con “EndoFree Plasmid Maxi Kit” (Qiagen). In seguito alla quantificazione con spettrofotometro, 6 ng di plasmide Env sono stati miscelati in 150µL di DMEM e con 12 ng di vettore pNL43 ΔGP160 (Fig.12). Tale plasmide codifica per i geni necessari alla produzione di pseudovirus tranne Env, che viene fornito dal vettore TOPO; la miscela ottenuta è stata purificata in filtri da 0.22 µm e aggiunta a 652 µL di DMEM e 48 µL di “FuGENE® 6 Transfection Reagent” (Roche), composto in grado di legare il DNA e trasfettarlo all’interno di cellule di mammifero competenti. La linea epiteliale tumorale HEK293T è stata utilizzata a tal scopo: dopo 48h e 72h dalla trasfezione, è stato raccolto il supernatante e utilizzato per l’infezione su cellule *reporter* per verificare la presenza di pseudovirus infettivo, ovvero in grado di esprimere una gp120 funzionale.



**Fig 12** Plasmide pNL43ΔGP160 per la produzione di pseudovirus.



## 2.9.4 Saggi luciferasici

Per provare l'effettiva produzione di virioni, il supernatante delle 293T è stato utilizzato per l'infezione della linea TZM-bl, ovvero cellule HeLA (CXCR4+) modificate per l'espressione aggiuntiva di CD4+ e CCR5+ e di un gene per la luciferasi LTR-inducibile. Le cellule sono state tenute in coltura con DMEM al 10% di FBS. Ciascun campione è stato eseguito in quadruplicato con diluizioni 1:2 e per ogni replicato è stato effettuato un controllo negativo con sole cellule. L'infezione è stata effettuata in piastre da 96 pozzetti con 10000 cellule/well e dopo 48 ore di infezione in incubatore a 37°C, si è proceduto alla lisi delle cellule svuotando ciascun pozzetto del terreno e aggiungendo 50uL di Luciferase Assay reagent (Promega) diluito 1:5. La lettura della luminescenza, in presenza dell'opportuno substrato, è stata eseguita mediante "Luminoskan Ascent Luminometer". Poiché le cellule TZM-bl esprimono comunque basali livelli di luciferasi in assenza di virus, sono stati considerati positivi i pozzetti con una luminescenza (espressa in RLU=relative luminescence unit) superiore a 2.5 volte la deviazione standard della media dei controlli. Mediante calcolo del "Reed and Muench" è stata ottenuta la TCID50 di ciascun pseudovirus, in modo da poter effettuare successivamente le opportune diluizioni e utilizzare la stessa TCID50 (pari a 100/mL) per l'infezione sulla linea cellulare U87. Le U87 sono cellule umane di glioblastoma ingegnerizzate per l'espressione di CD4, CXCR4 e CCR5. Negli esperimenti qui riportati sono state utilizzate separatamente le linee U87\_CD4+CXCR4+ e U87\_CD4+CCR5+ ; ciascun pseudovirus positivo al test su TZM-bl è stato diluito 2, 4 e 8 volte per l'infezione sulle U87

dopo 24h di coltura in piastra 96-wells. Tutti i campioni sono stati eseguiti in quadruplicato e la lettura al luminometro effettuata dopo 48h di incubazione a 37°C. Per ciascun campione è stata calcolata la media corrispondente alle singoli diluizioni ed è stata considerata positiva la luminescenza 10 volte superiore alla media dei controlli.

### **2.9.5 Analisi statistica**

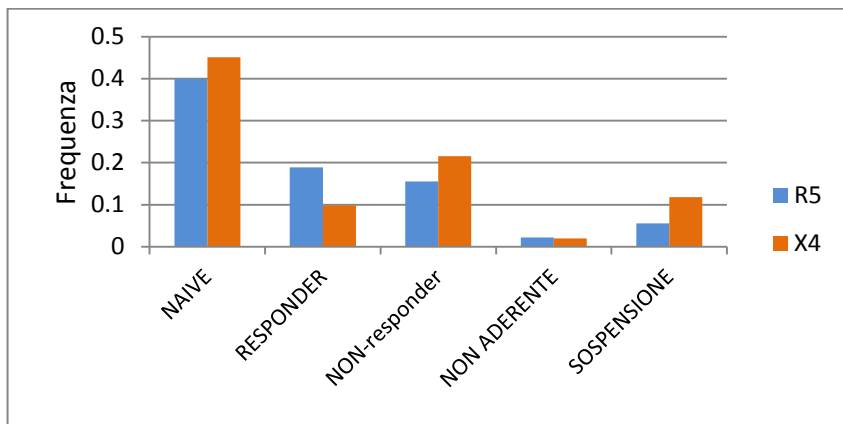
Lo studio delle variabili categoriche nominali (mutazioni), in associazione al tropismo, è stato condotto mediante analisi univariata utilizzando il software *IBM SPSS Statistics 19*. In particolare, dal momento che la variabile dipendente (tropismo) era rappresentata da solo due categorie (R5/X4), è stata utilizzata la *regressione logistica binaria* per valutare la significatività delle associazioni.

## **3. RISULTATI**

### **3.1 Distribuzione del tropismo nelle sottopopolazioni di pazienti**

Dall'analisi delle sequenze con G2P, utilizzando un cut-off di FPR pari a 10% per i campioni di RNA (plasma) e 20% per i campioni di DNA (PBMCs, CD4+, CD14+, biopsie), è emerso che il 64% delle sequenze sono rappresentate da ceppi virali con tropismo R5, mentre il 36% è rappresentato da virus X4. All'interno dei due gruppi di sequenze, è stata rilevata una simile proporzione di campioni derivanti da pazienti *naïve* (0.4 vs 0.45), ovvero il gruppo maggiormente rappresentato della popolazione in esame (42%), mentre fra gli X4 la frequenza di *non-responder* (NR) e

pazienti in sospensione risultava essere maggiore rispetto a quella osservata fra gli R5 (*non-responder*: 0.22 vs 0.16; sosp:0.12 vs 0.06, Fig.13).



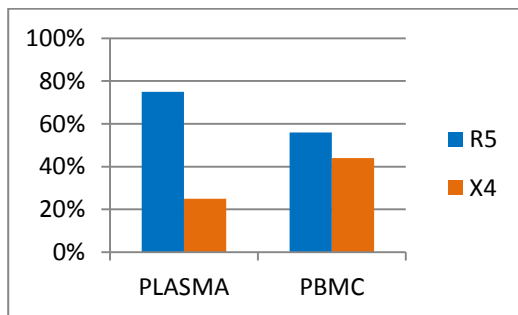
**Fig. 13** Distribuzione di ceppi R5 e X4-tropici nelle sottopopolazioni di pazienti

Recentemente, numerosi studi si sono focalizzati sulla distribuzione del tropismo nella popolazione di pazienti naïve, sia nel plasma che nei PBMCs. Tale aspetto riveste un'importanza notevole per un futuro utilizzo degli inibitori del CCR5 in questi pazienti, sebbene la somministrazione del *Maraviroc*, attualmente, sia indicata nelle linee guida solo per i pazienti con infezione avanzata. Tale scelta è dovuta anche agli elevati costi che annoverano questo farmaco ancora tra le terapie di salvataggio. Molti lavori hanno riportato una prevalenza di virus R5 nel plasma di pazienti naïve, sia con infezione recente (MOTIVATE studies) che avanzata [42]. È stato stimato che la prevalenza sia di circa 80% nei pazienti con infezione primaria e del 62% tra i pazienti in fase avanzata. Tali osservazioni

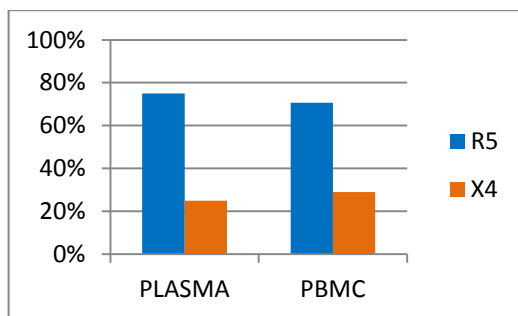
sottolineano la potenziale efficacia degli inibitori del CCR5 in un ampio target di soggetti HIV+, a diversi stadi di infezioni.

Il 41% dei pazienti inclusi in questo studio è rappresentato da soggetti HIV+ naïve al trattamento antiretrovirale. Come riportato in Fig.14 (A,B), il 75% dei virus presenti nei campioni di plasma di pazienti naïve hanno tropismo R5 (FPR cut-off=10%), dato che conferma quanto osservato in precedenti lavori. Diversamente, nei PBMCs, la predizione del tropismo con G2P utilizzando un cut-off di FPR pari al 20% mostra una simile proporzione tra ceppi R5 e X4 (56% vs 44%). Tale risultato sembra essere una sovrastima rispetto a quanto emerso in altre analisi che riportano una percentuale di X4 nel plasma di pazienti con infezione recente pari al 21% [43]. Tuttavia, la frequenza di ceppi X4-tropici e DM nelle infezioni primarie varia notevolmente fra i diversi lavori e in base al metodo di predizione utilizzato, riportando stime da 3.2 a 17% nelle varie coorti studiate [44, 45, 25]. La discrepanza di risultato potrebbe essere dovuta all'utilizzo del metodo genotipico con un cut-off al 20% che aumenterebbe la sensibilità nell'identificazione di X4. La predizione del tropismo, negli stessi campioni, con un cut-off di FPR pari al 10% ridurrebbe la proporzione di X4 dal 44% al 29%, con una distribuzione del tropismo più simile a quanto osservato nel plasma (Fig.14B). Inoltre, è importante considerare che gli studi citati, in cui la proporzione di X4 è del 20%, analizzano pazienti con infezione primaria, mentre nel nostro caso la popolazione naïve è piuttosto eterogenea, ovvero include pazienti recentemente infetti e *late presenters*, ovvero pazienti che effettuano il primo controllo clinico in fase avanzata della malattia. A conferma di quanto riportato nel lavoro di Parisi, la quota di X4 osservata nei pazienti

con tropismo concordante tra plasma e PBMCs (90% della popolazione), è superiore nei pazienti con infezione tardiva cronica (LCI, 39.3%) rispetto alle infezioni recenti (19.4%) e alle infezioni croniche primarie (9.6%).



(A)



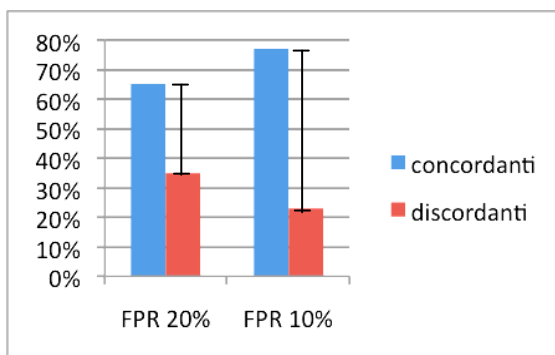
(B)

**Fig.14** Distribuzione del tropismo tra plasma e PBMCs **(A)** con cut-off di FPR per il DNA al 20%; **(B)** con cut-off di FPR per il DNA al 10%

### 3.2 Scelta del *cut-off* di FPR e concordanza dei risultati

L'evidenza che in un paziente naïve possano circolare varianti con diverso tropismo fra i vari compartimenti, è stata riportata nello stesso studio di

Parisi del 2011 nel quale, sebbene in 90/100 pazienti ci fosse concordanza tra plasma e PBMCs, in 21.2% dei casi si osservava una discordanza di tropismo tra plasma, PBMCs e fluido cerebrospinale. All'interno del nostro gruppo di pazienti, 26/105 sono stati analizzati in differenti compartimenti, per cui è stato possibile confrontare i risultati ottenuti sulla predizione del tropismo. In particolare, 17/26 (65%) soggetti presentavano una concordanza nel risultato impostando un cut-off per il DNA pari a 20%. La scelta di un cut-off inferiore, pari a quello utilizzato per l'RNA virale, aumenta la concordanza del 12% fra compartimento cellulare e plasma; di conseguenza, in tale *setting*, i campioni discordanti risultano essere il 23%, a fronte del 77% di concordanti (Fig.15), analogamente a quanto riportato precedentemente.



**Fig.15** Concordanza nel tropismo tra plasma e PBMCs con cut-off di FPR per il DNA al 20% e al 10%

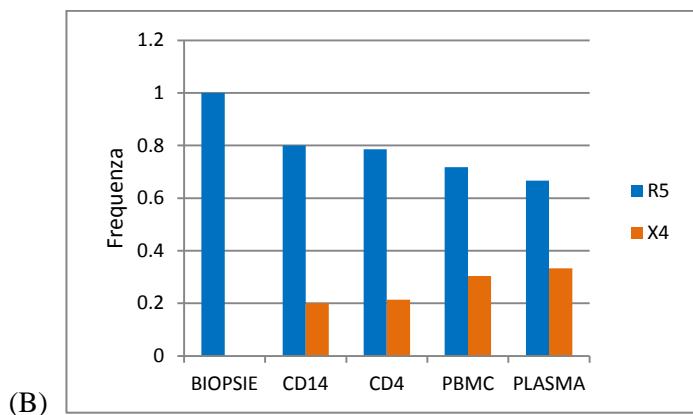
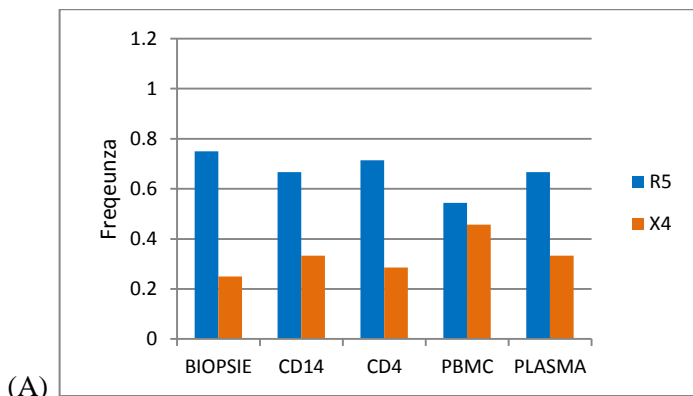
Queste osservazioni sottolineano la necessità di valutare ulteriormente il miglior cut-off di FPR da utilizzare nella pratica clinica: sebbene

l'aumento di questo valore soglia per l'analisi del DNA provirale aumenti la sensibilità nel rilevare gli X4, il tropismo potrebbe in realtà essere predetto erroneamente in più del 10% di questi casi, limitando l'utilizzo degli inibitori del CCR5 anche laddove potrebbero essere somministrati.

### 3.3 Prevalenza di ceppi R5 nei compartimenti studiati

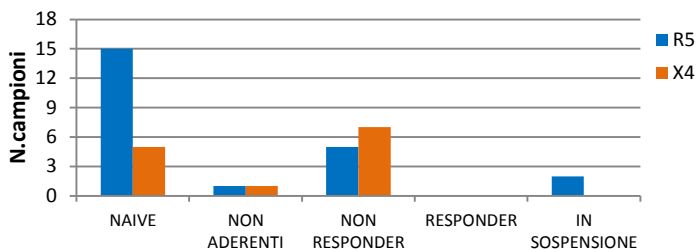
Come specificato nel par. 2.1, la maggior parte dei campioni analizzati in questo studio è rappresentata da plasma (34%) e PBMCs (35%), seguiti da cellule CD4+ (25%), CD14+ (14%) e biopsie (4%). L'istogramma in Fig.16A mostra una generale prevalenza di ceppi R5, sia nel compartimento plasmatico che nei compartimenti cellulari. La differenza meno significativa si osserva nei PBMCs, in cui sembra esservi una maggior incidenza di ceppi X4. Tuttavia, analizzando gli stessi dati con un *cut-off* per il DNA del 10%, oltre ad aumentare la notevole discrepanza nella frequenza di R5 e X4 osservata nelle cellule, nei PBMCs si rileva una distribuzione perfettamente in linea con gli altri compartimenti, nonché una maggior corrispondenza con la distribuzione plasmatica (Fig.16B). Considerando separatamente i campioni di RNA (plasma) e DNA (PBMCs, CD4+, CD14+, biopsie), abbiamo analizzato la distribuzione del tropismo all'interno dei singoli gruppi di campioni naïve, responder, non-responder, in sospensione e non aderenti. Come atteso e alla luce delle precedenti osservazioni, sia nei campioni di RNA che DNA si riscontra una maggior prevalenza dei ceppi R5 e nelle cellule la differenza aumenta con un *cut-off* del 10% (Fig.17). Un dato interessante è emerso relativamente ai campioni derivanti da soggetti *non responder*, ovvero i pazienti che al momento dell'analisi non rispondevano efficacemente alla

terapia antiretrovirale. Nel plasma infatti, sembra esservi una prevalenza di ceppi X4; sebbene l'osservazione sia fondata su un gruppo esiguo di campioni (12) tale risultato potrebbe essere spiegato dal fatto che il virus rilevato nel plasma è in grado di evadere la ART, ovvero presenta delle mutazioni nella regione Pol che consentono l'*escape* della terapia.

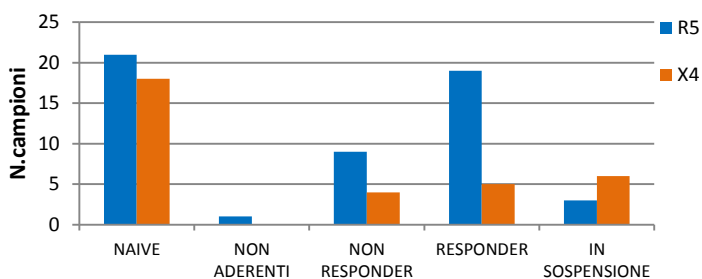


**Fig.16** Distribuzione del tropismo nei diversi compartimenti con diversi cut-off di FPR (A) FPR=20% (B) FPR=10%

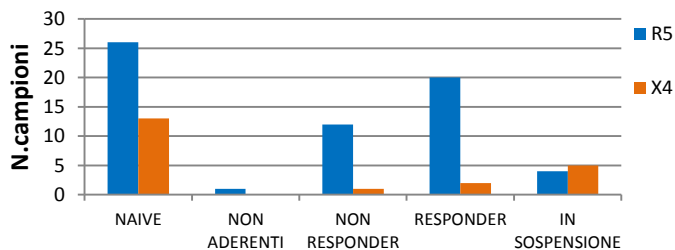




(A) Sottopopolazioni di pazienti



(B) Sottopopolazioni di pazienti



(C) Sottopopolazioni di pazienti

**Fig. 17** Distribuzione del tropismo nelle sottopopolazioni di pazienti (A) nei campioni di RNA (B) nei campioni di DNA, cut-off FPR=20% (C) nei campioni di DNA, cut-off FPR=10%

Nonostante non siano stati ancora riportati dei lavori in cui si dimostra chiaramente che un virus X4 presenta una maggiore diversità in Pol e quindi maggior fitness replicativa in caso di regime terapeutico, nello studio di Kassaye et al. [46] gli autori osservano che esiste una correlata evoluzione dei geni Env e Pol nei pazienti HIV+ sottoposti ad ART, sebbene la regione Env presenti un tasso di mutazione più elevato. È noto, inoltre, che la maggior variabilità di Env sia associata a ceppi X4-tropici; di conseguenza, un virus in grado di sfuggire alla terapia potrebbe essere verosimilmente un virus X4 con aumentata efficienza replicativa. Ciò spiegherebbe anche la differenza osservata nei campioni di DNA, dove si riscontra comunque una prevalenza di virus R5 tropici; la terapia potrebbe favorire fortemente la replicazione dei virus mutati (nel gene Pol), ovvero X4, nonostante il *reservoir* cellulare rimanga sostanzialmente R5-tropico.

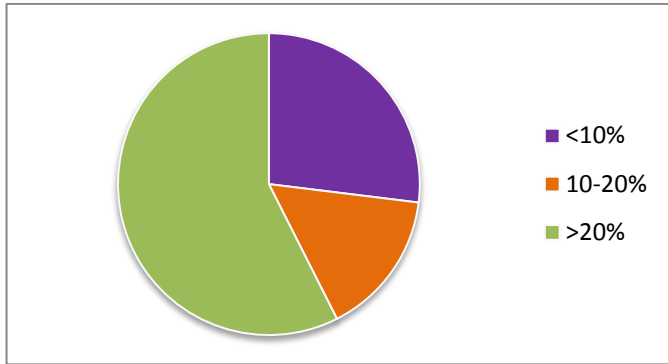
### **3.4 Associazione carica netta in V3 e FPR**

Come descritto nel par. 2.8, la predizione del tropismo mediante geno2pheno si basa sulla interpretazione della sequenza amminoacidica del V3 loop. Il *cut-off* di FPR viene scelto dall'operatore e può variare da 1% a 20%. Valori di FPR superiori al *cut-off* stabilito sono associati a tropismo CCR5, mentre valori inferiori si associano a tropismo X4. In questo studio, il *cut-off* di FPR è stato scelto diversamente in base al tipo di campione analizzato: 10% per l'RNA virale e 20% per il DNA, come suggerito da linee guida del 2011. Di conseguenza è stata effettuata una prima classificazione dei campioni in base al FPR, ottenendo tre gruppi:

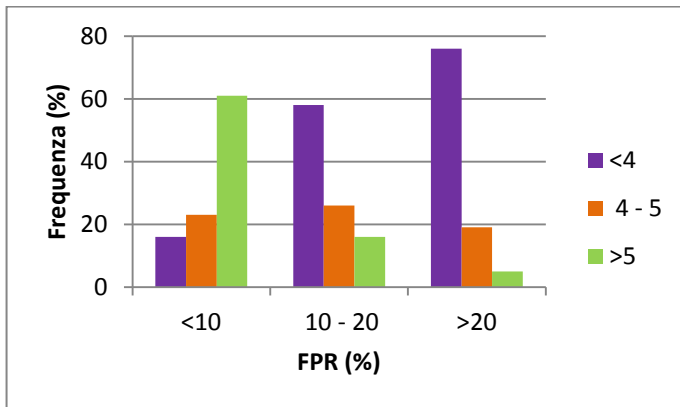
- (i) sequenze con  $FPR > 20\%$
- (ii) sequenze con  $FPR$  compreso tra 10 e 20%
- (iii) sequenze con  $FPR < 10\%$

La maggior parte dei campioni oggetto di studio (57%) presentavano un  $FPR > 20\%$ , ovvero tropismo R5, mentre le sequenze dei gruppi (ii) e (iii) si distribuivano in maniera simile (Fig.18A), rappresentando rispettivamente il 16% e il 27% della popolazione totale. Inoltre, ciascun campione è stato esaminato in base alla NC del V3 *loop* (par.2.9) per cui è stato possibile stratificare le sequenze in base a entrambi i parametri (NC e  $FPR$ ). I gruppi con  $FPR < 10\%$  e  $> 20\%$  mostrano una buona associazione tra NC e  $FPR$ , dal momento che il 76% delle sequenze con  $FPR > 20\%$  (R5) hanno  $NC < 4$  e il 61% dei campioni con  $FPR < 10\%$  (X4) hanno  $NC > 5$ . Il gruppo (ii), ovvero con  $FPR$  compreso tra 10 e 20%, mostra una distribuzione sovrapponibile al gruppo (iii) poiché la maggior parte delle sequenze (58%) hanno  $NC < 4$  (Fig.18B).

Le suddette osservazioni concordano con i risultati riportati da Clevestig et al.[37], i quali mostrano una relazione tra il tropismo e valori estremi di NC in V3 (inferiori a +3 e superiori a +7). Alla luce di questa conferma e dal momento che la distribuzione di NC nei gruppi (ii) e (iii) è estremamente simile, è possibile sostenere che un *cut-off* del 10% può essere tenuto in considerazione sia per la predizione di RNA che di DNA virale.



(A)



(B)

**Fig. 18**

(A) Stratificazione dei campioni in base al valore di FPR

(B) Stratificazione dei campioni in base ai valori di FPR e NC di V3

Lo studio di Clevestig e colleghi aveva anche identificato un *breaking point* di NC in V3 pari a 4.2, al di sotto del quale non si rilevano virus X4 e al di sopra di cui si osserva il 7.5% di sequenze R5. Analogamente, la nostra stratificazione in base ai valori di carica netta in V3, ha consentito di identificare i gruppi:

(j) sequenze con  $NC < 4$

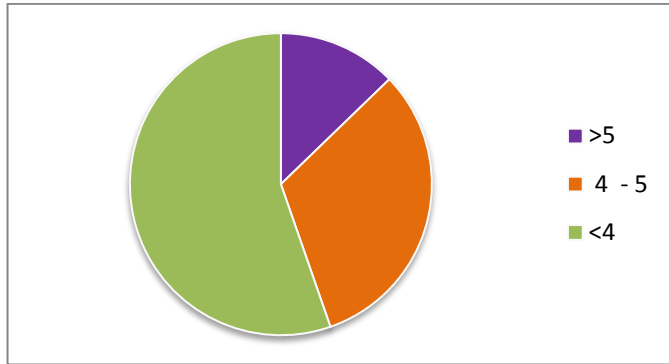
(jj) sequenze con  $4 < NC < 5$

(jjj) sequenze con  $NC > 5$

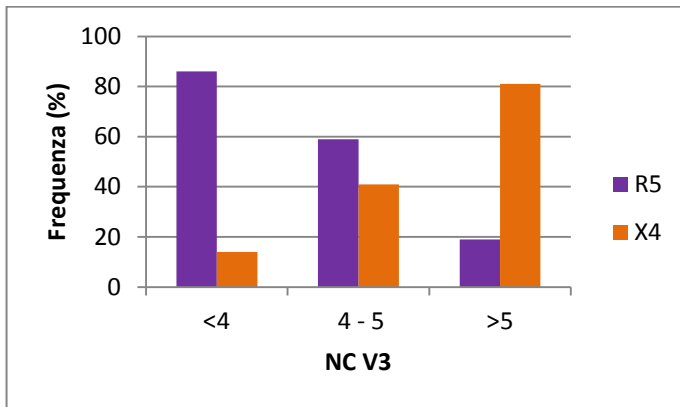
distribuiti come mostrato in Fig.19A.

Analogamente a quanto emerso dall'analisi basata sui valori di FPR, la maggior parte (86%) delle sequenze con carica netta inferiore a 4 sono state predette da G2P come R5, e 81% dei campioni con  $NC > 5$  avevano tropismo X4 (Fig.19B). Un gruppo *outset* è rappresentato dalle sequenze con valori di NC compresi tra 4 e 5, in cui non vi è una chiara associazione tra tropismo e carica netta. In tale gruppo, infatti, è stata osservata una simile proporzione di ceppi R5 e X4 (rispettivamente 59% e 41%).

Complessivamente, il 58% delle sequenze analizzate mostravano piena concordanza tra carica netta e predizione di G2P, mentre il 10% risultavano essere discordanti. Le sequenze con carica netta fra 4 e 5 sono state considerate un gruppo a sé stante, rappresentando il 32% del totale (Fig.20).



(A)

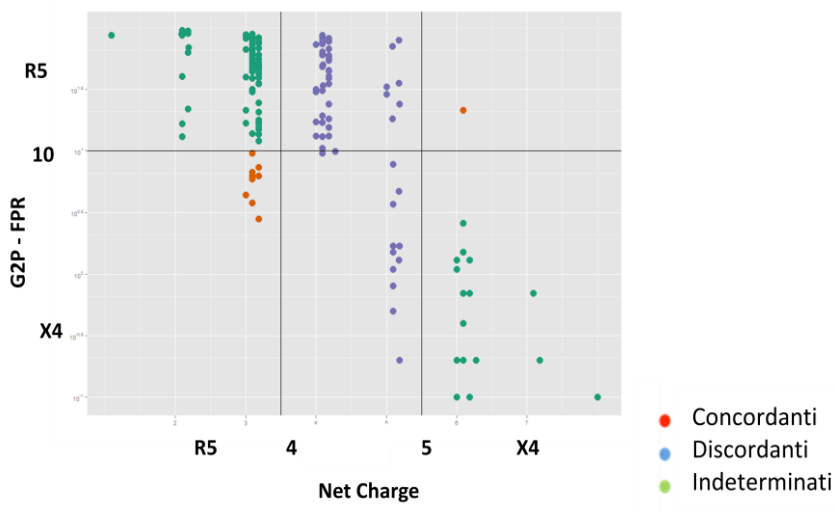


(B)

**Fig. 19**

(A) Stratificazione dei campioni in base alla carica netta di V3

(B) Distribuzione del tropismo analizzato con G2P nei sottogruppi ottenuti in base alla NC di V3

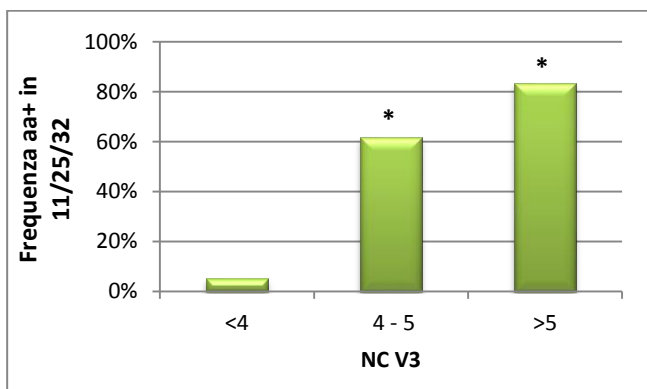


**Fig. 20** Stratificazione dei campioni in base alla concordanza tra il risultato di G2P e la NC di V3.

L'eterogenea distribuzione delle sequenze con NC compresa tra 4 e 5 indurrebbe a pensare alla presenza di varianti Dual o Mixed, che non vengono discriminate da G2P rispetto alle sequenze X4. Sebbene la distinzione dei DM dagli X4 possa non essere fondamentale ai fini di una terapia basata su inibitori del CCR5 (dal momento che *Maraviroc* è inefficace nei pazienti infetti con virus DM), l'identificazione di varianti *dual* potrebbe essere utilizzata come predittore dello *switch* corecettoriale e come marcatore di progressione. Numerosi studi, infatti, riportano che lo *switch* del tropismo si osserva nei pazienti dopo numerosi anni di infezione, ma questo cambiamento è fortemente soggettivo ed ha delle tempistiche variabili nei pazienti (Regoes, TRENDS in Microbiology).

### 3.5 Mutazioni in V3 associate al tropismo

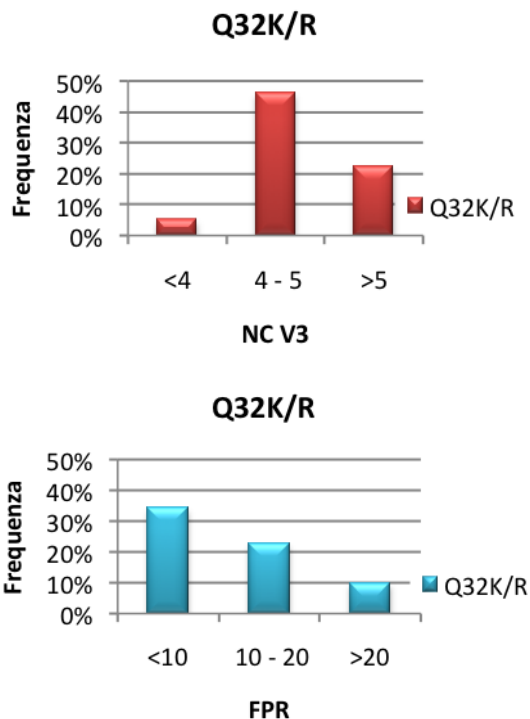
L'associazione tra specifici residui amminoacidici e il tropismo è stata dimostrata in diversi studi; in particolare, la presenza di amminoacidi carichi positivamente (quali lisina,K e arginina,R) nella posizioni 11 e 25 è di solito correlata alla presenza di varianti X4 [47]. Dal momento che la presenza di tali amminoacidi è sufficiente, ma non necessaria, per la predizione del tropismo di HIV, l'identificazione di altri residui coinvolti nell'uso del corecettore riveste ancora notevole interesse. All'interno del nostro gruppo di studio abbiamo dunque calcolato la frequenza di residui positivi nelle posizioni 11 e 25, confermando le osservazioni che ne riportano l'associazione con virus X4-tropici, ma abbiamo rilevato anche una forte associazione tra il tropismo e amminoacidi positivi in posizione Q32 del V3 loop. La presenza di un residuo positivo in almeno una delle tre posizioni (11, 25, 32) si associa significativamente alla carica netta ( $p < 0.05$ ), come mostrato in Fig.21).



**Fig.21** Frequenza di residui amminoacidici positivi (K o R) in almeno una delle posizioni 11/25/32 di V3



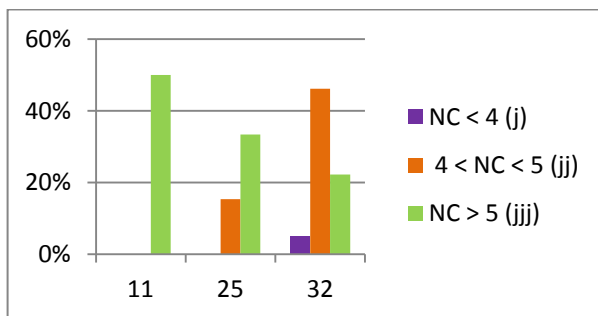
La Q32 si è dimostrato un residuo interessante in quanto mutazioni positive in questa posizione sembrano essere fortemente associate a carica netta compresa tra 4 e 5; si osserva inoltre, una tendenza ad associarsi con  $FPR < 20\%$  (Fig.22).



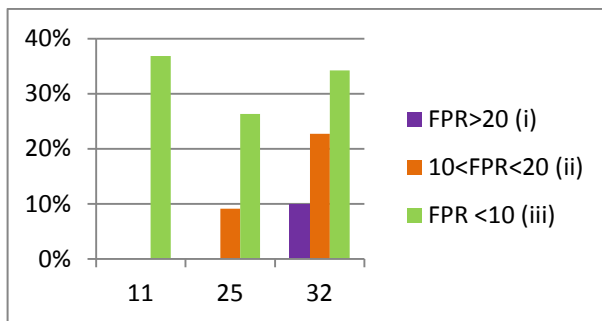
**Fig.22** Frequenza di amminoacidi carichi positivamente in posizione 32

La frequenza dei residui funzionali 11, 25 e 32 è riportata in Fig.23(A,B) che mostra la distribuzione delle cariche positive nelle posizioni interessate stratificando le sequenze in base a NC e FPR. Risulta interessante che la frequenza di cariche positive nelle posizioni 11 e 25 è significativamente

maggiore nelle sequenze dei gruppi (iii) and (jjj) ( $p < 0.01$ ), ulteriore conferma che  $NC > 5$  e  $FPR < 10\%$  si associano a queste mutazioni, ovvero a tropismo X4. D'altra parte, la posizione Q32 presenta delle differenze rispetto alla frequenza di K/R nelle posizioni 11/25: una maggior frequenza di residui positivi in corrispondenza di questa posizione si associa a  $NC > 4$ , osservazione che ci permette di associare Q32K/R alle sequenze del gruppo (jj) ( $4 < NC < 5$ ).



(A)



(B)

**Fig.23** Frequenza di cariche positive nelle posizioni 11,25 e 32 nei campioni suddivisi (A) in base a NC di V3 (B) ai valori di FPR

L'analisi statistica è stata eseguita eliminando 15 sequenze a causa di ambiguità presenti nelle posizioni considerate, ma è concordante con la distribuzione delle mutazioni osservata nei campioni totali. Di conseguenza, Q32 sembra avere un diverso valore soglia di carica netta che sarebbe utile a discriminare tra ceppi DM tropici e X4.

### **3.6 Test fenotipico su campioni concordanti, discordanti e indeterminati**

Per verificare le osservazioni sopra descritte ottenute con il test genotipico, è stato messo a punto un test fenotipico per la determinazione del tropismo. In particolare, sono stati scelti dei campioni rappresentativi per ciascun gruppo descritto nel par. 3.4, ovvero concordanti, discordanti e indeterminati, in base ai risultati di G2P e NC V3, per un totale di 16 pazienti, di cui 8 concordanti (C), 6 indeterminati (I) e 2 discordanti (D), come illustrato in Tab.3.

Ciascun campione è stato quindi amplificato e clonato come descritto in Materiali e Metodi (pag.37), in modo da ottenere singole varianti virali da poter saggiare con il test fenotipico. In seguito al clonaggio e al sequenziamento, infatti, i cloni sono stati utilizzati per la produzione di pseudovirus, ma non tutti risultavano essere infettanti sulle TZM-bl. In alcuni casi ciò era dovuto alla presenza di codoni di stop nella sequenza che impedivano l'espressione di una gp120 funzionante, in altri casi non erano presenti, apparentemente, mutazioni in grado di impedire una corretta espressione di Env.

ID	CAMPIONE	G2P	NC V3	G2P vs NC
A	DNA	R5	4,1	I
BA	DNA	X4	2,0	D
C	DNA	X4	4,1	I
D	DNA	X4	4,1	I
E	DNA	X4	6,0	C
F	DNA	R5	3,0	C
G	DNA	X4	6,0	C
L	DNA	X4	5,0	I
LG	RNA	X4	4,0	I
M	DNA	X4	6,1	C
MA	RNA	R5	3,0	C
N	DNA	X4	6,1	C
O	DNA	X4	8,0	C
ODL	RNA	R5	4,0	I
PB	RNA	X4	3,0	D
VAN	RNA	R5	3,0	C

**Tab.3** Campioni selezionati per la produzione di pseudovirus ricombinanti.

Tuttavia, molti dei campioni analizzati provenivano da PBMCs ed è noto che non tutte le varianti virali archiviate nelle cellule sono in grado di replicare efficientemente, a causa dell'accumulo di mutazioni che ne riducono fortemente la fitness replicativa. Ciò potrebbe spiegare la mancata infettività di alcuni pseudovirus. I campioni analizzati e i cloni ottenuti sono mostrati in Tab.4(A); in corrispondenza dei cloni in grado di replicare nelle TZM-bl, sono riportati i risultati del test genotipico.

ID	CAMPIONE	CLONE	TZM-bl	G2P	FPR (%)	NC V3	G2P vs NC
A	DNA	A2	+	R5	76	4	I
		A4	+	R5	39,6	4,1	I
		A5	+	R5	39,8	4,1	I
		A6	+	R5	81,4	4,1	I
BA	DNA	BA9	STOP CODON	-	-	-	-
		BA22	+	X4	14,7	2	D
C	DNA	C4	-	-	-	-	-
D	DNA	D7	-	-	-	-	-
E	DNA	E9	STOP CODON				
F	DNA	F2	-	R5	20,8	4	I
		F5	-	R5	20,8	4	I
		F6	STOP CODON				
G	DNA	G3	-	X4	6,3	4	I
		G5	-	R5	20,8	4	I
L	DNA	L9	-	X4	7,8	4,1	I
LG	RNA	LG2	+	X4	2,9	5	I
M	DNA	M2	+	X4	0,1	5	I
		M7	-	X4	1,1	5	I
		M8	-	X4	1,1	5,1	C
MA	RNA	MA8	+	R5	89,6	3	C
N	DNA	N3	-	X4	1,1	5	I
		N8	-	X4	1,1	5	I
O	DNA	O5	+	X4	0,1	8	C
		O6	+	X4	0,1	8	C
		O10	+	X4	0,2	8	C
ODL	RNA	ODL1	+	X4	1,7	5	I
		ODL2	STOP CODON				
		ODL3	+	R5	29,2	4	I
		ODL4	STOP CODON				
		ODL17	+	R5	11,7	3	C
PB	RNA	PB6	+	X4	3,2	3	D
VAN	DNA	VAN3	-	R5	43	3,9	C
		VAN16	-	R5	55	2,9	C

**Tab. 4** Pseudovirus testati su TZM-bl

### 3.7 Risultati del test fenotipico

Gli pseudovirus positivi al saggio su TZM-bl sono stati testati su cellule U87\_CD4 esprimenti CCR5 o CXCR4, utilizzando una TCID50 pari a 100/mL. L'unica eccezione è rappresentata dai campioni A2 e A5 che, con una TCID50 pari a 100/mL, mostravano non solo una forte positività sulle U87\_CCR5+, ma anche un debole segnale sulle U87\_CXCR4+; tali campioni sono stati dunque testati nuovamente su cellule esprimenti CXCR4 ma con una TCID50 più elevata (500/mL). Nonostante il segnale di positività fosse nettamente più debole sulle cellule CXCR4+, i due ceppi risultavano essere in grado di infettare entrambe le linee cellulari. Il dato interessante è relativo al fatto che gli altri due ceppi virali isolati dallo stesso paziente (A4 e A6), erano in grado di replicare esclusivamente in cellule R5+, dimostrando ancora una volta l'eterogeneità di *pool* virale potenzialmente replicativo a partire da uno stesso campione. I risultati del test fenotipico riportati in Tab.5 mostrano che, considerando i campioni con NC compresa tra 4 e 5 come *dual*-tropici e che G2P non discrimina tali virus, si è ottenuta una concordanza tra NC e test fenotipico in 11/15 campioni (73%), mentre in 8/15 (53%) si aveva concordanza tra G2P e fenotipo *in vitro*.

Tuttavia, all'interno del gruppo di sequenze con NC compresa tra 4 e 5, la metà dei campioni risultava essere concordante con il saggio su U87, confermando che tali valori di NC sono *border line* per la determinazione del tropismo ma anche che, nel 50% dei casi, i ceppi virali con NC compresa tra 4 e 5 sono in grado di utilizzare entrambi i corecettori, ovvero risultano essere a tutti gli effetti virus *dual*-tropici.

CLONE	TZM-bl	G2P	FPR (%)	NC V3	fenotipo
ODL1	+	X4	1,7	5	DUAL
ODL3	+	R5	29,2	4	DUAL
A4	+	R5	39,6	4,1	R5
A5	+	R5	39,8	4,1	DUAL*
A2	+	R5	76	4	DUAL*
A6	+	R5	81,4	4,1	R5
BA22	+	X4	14,7	2	R5
LG2	+	X4	2,9	5	R5
ODL17	+	R5	11,7	3	R5
PB6	+	X4	3,2	3	R5
M2	+	X4	0,1	5	X4
O5	+	X4	0,1	8	X4
O6	+	X4	0,1	8	X4
O10	+	X4	0,2	8	X4
MA8	+	R5	89,6	3	R5

\*campioni positivi sulle U87\_CD4+CXCR4+ con una TCID50 = 500/mL

**Tab. 5** Risultati del test fenotipico su cellule U87.

Per i valori di NC inferiore a 4 e superiore a 5, dunque, la concordanza tra test fenotipico e risultato della carica netta è pari al 100%, mentre la concordanza tra fenotipo in vitro e G2P, nello stesso *range* di NC, è del 71%. La predizione del tropismo si complica per valori di carica netta 4-5, dove la concordanza tra NC/G2P *versus* fenotipo è pressoché uguale (rispettivamente 4/8 e 3/8).

### 3.8 Analisi delle mutazioni in regioni diverse da V3

L'ipotesi che residui amminoacidici nelle regioni variabili e costanti di Env, e al di fuori della regione V3, fossero coinvolti nella scelta del

corecettore è stata supportata dai risultati ottenuti in seguito agli studi comparativi tra test genotipico e fenotipico. Alcuni autori [48,49] sostengono che i risultati discordanti tra genotipo e fenotipo possono essere dovuti alla presenza di mutazioni nella sequenza di Env non apprezzabili con la sola analisi di V3. L'identificazione di tali mutazioni, tuttavia, è complicata dall'estrema variabilità della gp120, pertanto sono ancora pochi gli studi che identificano un'associazione tra le mutazioni nelle regioni variabili (diverse da V3) e il tropismo. In un lavoro del 2009, L.Monno e colleghi [49] hanno esaminato un gruppo di 66 pazienti infetti con sottotipo B di HIV-1 eseguendo un confronto tra test genotipico e fenotipico da campioni di plasma. Nessun campione analizzato presentava esclusivamente tropismo X4, mentre sono stati identificati pazienti R5 e DM-tropici, *early*-R5 e *late*-R5 (pazienti rispettivamente asintomatici e sintomatici infetti con ceppi R5-tropici). Nei gruppi di pazienti studiati sono state riscontrate, con diversa frequenza statisticamente significativa, mutazioni in C1, C2, C3, C4, V1, V2, V3, V4, V5. Le mutazioni che mostravano significatività anche in seguito ad analisi multivariata sono state prese in considerazione per il nostro studio. A partire dal confronto con il test fenotipico, infatti, abbiamo cercato di elaborare un modello in grado di predire il tropismo con elevata accuratezza e che considerasse anche regioni diverse da V3. In corrispondenza di ogni parametro, considerato a partire dal lavoro di Monno e in base alle osservazioni ottenute dalla prima parte del nostro lavoro, è stato associato un valore compreso fra -1 e +1, in base alle caratteristiche di ciascun parametro in campioni R5, X4, DM, *early*-R5 e *late*-R5, come indicato in Tab.6 A, B.



C1	V84L	DM>R5
V1	N136L	E-R5>L-R5
	N139T	R5>DM
	S144del	E-R5>L-R5
	R166G	R5>DM
V2	V169K	R5>DM
	V181V	R5>X4
	V181I	R5>X4
	D187del	L-R5>E-R5
	D187N	L-R5>E-R5
	S195N	R5>DM
C2	T198A	DM>R5
	T232K	DM>R5
V3	N301glyc	R5>X4
	306K/R	X4>R5
	R306G	E-R5>L-R5
	R315K	R5>DM
	322K/R	X4>R5
	328K/R	X4>R5
	Carica netta	
C3	K343A	DM>R5
	Q344K	R5>DM
	I360K	DM>R5
V4	S398Q	L-R5>E-R5
	T413I	DM>R5
C4	S440K	DM>R5
	S446T	DM>R5
V5	E464S	DM>R5
	E464N	L-R5>E-R5

SCORE	
R5	1
DM	0
X4	-1
Early-R5	1
Late-R5	0,5

(B)

(A)

**Tab. 6 (A)** Mutazioni considerate per il calcolo dello score di *Env* **(B)** Punteggi associati al tropismo

La somma dei punteggi in base ai parametri sopra riportati, ha consentito di ottenere uno *score* totale compreso tra +9 e +18, da cui è stato possibile ricavare una corrispondenza tra *score* inferiore a +12 e ceppi X4-tropici (Tab.7). La presenza dei ceppi *dual*-tropici si associa, invece, a punteggi più elevati, sebbene non sia stato possibile effettuare una chiara distinzione tra DM e R5 puri: i virus DM, infatti, si stratificano eterogeneamente all'interno del gruppo con *score* superiore a +12; ciò suggerisce che esistano altri parametri (non inclusi nel modello) che ne determinerebbero la discriminazione rispetto ai virus R5-tropici. La presenza dei ceppi *dual*-tropici, tuttavia, può essere suggerita da una carica netta di V3 compresa tra 4 e 5.

ID	G2P	FPR	NC V3	SCORE	U87_R5	U87_X4
O10	X4	0,2	8,0	9	-	+
O5	X4	0,1	8,0	9	-	+
O6	X4	0,1	8,0	10	-	+
M2	X4	0,1	5,0	12	-	+
LG2	X4	2,9	5,0	13	+	-
ODL3	R5	29,2	4,0	14	+	+
MA8	R5	89,6	3,0	15	+	-
ODL1	X4	1,7	5,0	15,5	+	+
A6	R5	81,4	4,1	16	+	-
A4	R5	39,6	4,1	16,5	+	-
ODL17	R5	11,7	3,0	16,5	+	-
A2	R5	76	4,0	17	+	+ *
A5	R5	39,8	4,1	17,5	+	+ *
BA22	X4	14,7	2,0	18	+	-
PB6	X4	3,2	3,0	18	+	-

**Tab. 7** Score calcolato considerando tutti i parametri noti delle regioni variabili e costanti di Env (\*TCID50=500/mL)

L'analisi mostra dei risultati simili se si considerano esclusivamente i parametri all'interno delle regioni variabili V1, V2, V3 (Tab.8), supportando l'ipotesi che in tali regioni possono essere ricercate delle caratteristiche discriminanti per l'interazione con il corecettore e che l'analisi delle regioni costanti rappresenta un elemento aggiuntivo, ma non essenziale, ai fini della predizione del tropismo mediante test genotipico.

ID	G2P	FPR	NC	SCORE	U87_R5	U87_X4
M2	X4	0,1	5,0	2,5	-	+
O10	X4	0,2	8,0	2,5	-	+
O5	X4	0,1	8,0	2,5	-	+
O6	X4	0,1	8,0	2,5	-	+
LG2	X4	2,9	5,0	5,5	+	-
ODL1	X4	1,7	5,0	6,5	+	+
ODL3	R5	29,2	4,0	6,5	+	+
A6	R5	81,4	4,1	7,5	+	-
MA8	R5	89,6	3,0	7,5	+	-
ODL17	R5	11,7	3,0	7,5	+	-
A2	R5	76	4,0	8	+	+*
A4	R5	39,6	4,1	8,5	+	-
A5	R5	39,8	4,1	8,5	+	+*
PB6	X4	3,2	3,0	8,5	+	-
BA22	X4	14,7	2,0	11	+	-

\*TCID50=500/mL

**Tab.8** Score calcolato considerando i parametri in V1, V2 e V3

D'altra parte, la debole positività dei ceppi *dual-tropici* con uno score più elevato (A2 e A5) lascia ipotizzare che questi virus, seppur in grado di interagire potenzialmente con il CXCR4, come dimostrato *in vitro*, leghino in maniera preferenziale il CCR5, pertanto il segnale di

chemioluminescenza risulta essere nettamente superiore. Il ruolo *in vivo* di questi virus ambivalenti rimane da chiarire; occorre sottolineare, infatti, che il paziente A, al momento dell'analisi, risultava essere *responder* alla terapia antiretrovirale, pertanto non è stato possibile stabilire se le varianti R5 pure o DM albergate nei PBMCs avessero una diversa fitness replicativa, o se fossero in grado di attivarsi in presenza di una pressione selettiva (rappresentata in tal caso da un inibitore del CCR5). La combinazione di farmaci antiretrovirali che hanno come bersaglio diverse regioni del genoma virale, garantisce un'efficace soppressione della replicazione; di conseguenza, la valutazione del tropismo a livello provirale deve essere contestualizzata considerando, da un lato, che le diverse varianti possono essere presenti nel paziente con proporzioni e caratteristiche differenti; dall'altro, che la determinazione del tropismo a partire da DNA esprime esclusivamente la *potenziale* capacità replicativa dei diversi ceppi, non necessariamente associata a un'effettiva riattivazione del virus.

#### **4. DISCUSSIONE**

L'introduzione degli inibitori del CCR5 nella pratica clinica ha fortemente stimolato la ricerca sullo studio dell'uso dei corecettori da parte di HIV-1. Nonostante non siano stati ancora chiariti i meccanismi molecolari esistenti fra tropismo e patogenesi, molteplici lavori hanno dimostrato l'esistenza di un'evoluzione del tropismo virale nel corso dell'infezione e le conseguenze sul sistema immunitario del paziente, evidenziando come lo *switch* da tropismo da R5 a X4 sia strettamente associato ad un rapido declino dei linfociti T CD4+ . D'altra parte, la somministrazione degli

inibitori del CCR5, e in particolare del *Maraviroc*, non può prescindere dalla determinazione accurata del tropismo nel paziente HIV+; l'utilizzo del farmaco laddove siano presenti varianti minoritarie X4, favorirebbe fortemente la replicazione di questi ceppi, provocando il fallimento della terapia.

Per tali ragioni, negli ultimi anni sono stati condotti numerosi studi mirati a validare un metodi rapido, semplice e sensibile in grado di predire il tropismo virale. La diffusione del test genotipico a partire dalla sequenza V3 è stata supportata dai numerosi vantaggi tra cui i costi contenuti, che lo hanno reso maggiormente appetibile, nella pratica diagnostica, rispetto allo storico saggio fenotipico *in vitro*. Tuttavia, la predizione del tropismo limitata, attualmente, all'interpretazione di una piccola regione di *Env* (V3), fa sì che il fenotipo venga ancora considerato il metodo d'elezione per determinare l'uso del corecettore.

Lo scopo del nostro lavoro è stato quello di studiare in maniera approfondita il rapporto tra la predizione del tropismo mediante *geno2pheno coreceptor* e la carica netta del V3. Nonostante non sia ancora del tutto chiara la motivazione biologica che lega tale parametro all'uso del corecettore, è stata riportata in alcuni lavori una forte associazione. Al fine di identificare dei marcatori in grado di aumentare l'accuratezza della predizione a partire dalla sequenza di V3, abbiamo analizzato in questo studio 144 campioni provenienti da 105 pazienti, che andavano a costituire un gruppo eterogeneo sia in termini di trattamento farmacologico che di fase di infezione. La nostra analisi ha messo in rilievo che esiste una prevalenza di ceppi R5-tropici nella popolazione studiata, prescindendo dai

sottogruppi con diversa terapia e dal compartimento analizzato. La proporzione di R5 osservata è pari al 64%, considerando diversi cut-off di FPR tra campioni di RNA e DNA virale. Poiché numerosi dati in letteratura riportano una notevole prevalenza di ceppi R5 nei pazienti naïve e dal momento che tale osservazione è stata confermata nella nostra casistica, la prevalenza di ceppi R5 nella coorte presa in esame può essere facilmente motivata dalla presenza del 41% di pazienti naïve alla terapia.

I risultati riportati sulla distribuzione del tropismo nei diversi compartimenti sono stati ottenuti impostando diversi *cut-off* di FPR al momento dell'analisi con G2P *coreceptor*. La corretta scelta di questo valore soglia è fondamentale, in quanto determina l'inclusione di un paziente nel regime terapeutico con inibitori del CCR5. La scelta di un cut-off del 20% (superiore a quanto indicato per l'analisi dell'RNA plasmatico) è stata suggerita da Linee guida europee del 2011 per l'interpretazione di sequenze di DNA provirale, al fine di incrementare la sensibilità nella rilevazione di ceppi X4-tropici. Tuttavia, un alto *cut-off* di FPR aumenta anche il rischio di ottenere risultati falso-positivi, ovvero falsi X4, escludendo da una terapia, in molti casi di salvataggio, pazienti che potrebbero beneficiarne. La frequenza di virus X4-tropici si riduce fortemente nel compartimento cellulare dei pazienti da noi studiati nel momento in cui si riduce il *cut-off* al 10%. I risultati ottenuti in tale *setting* trovano conferma in altri lavori [43, 50] in cui si osserva una prevalenza di X4 nella popolazione naïve pari a circa il 20%. La scelta di un *cut-off* di FPR più basso contribuisce anche ad aumentare la concordanza di tropismo tra compartimento plasmatico e cellulare; considerando un FPR del 10%,

la concordanza osservata nel nostro studio è di circa il 77%, analogamente a quanto riportato in alcuni lavori [43] ma inferiore rispetto ad altri [51]. Si può concludere, tuttavia, che la soglia di FPR al 10% per la predizione del tropismo nel compartimento cellulare, rappresenta un buon compromesso tra l'elevata sensibilità nel rilevare i ceppi X4-tropici e il basso rischio di ottenere dei falsi-X4.

I dati in letteratura inerenti il tropismo nei compartimenti cellulari analizzati in questo studio e diversi dai PBMCs sono scarsi, soprattutto per la ridotta reperibilità di tali campioni. Nonostante cellule CD4+, CD14+ e biopsie gastriche costituiscano complessivamente solo il 31% del nostro gruppo, è stato possibile osservare una distribuzione estremamente simile fra tutti i diversi compartimenti, con una comune prevalenza di virus R5. Un dato interessante è emerso dal confronto dei compartimenti monocitario e linfocitario, dove nell'87% dei casi si aveva tropismo concordante, per lo più R5.

Anche l'analisi del tropismo nei sottogruppi di pazienti, classificati in base alla terapia, ha fatto emergere delle osservazioni interessanti. Come già riportato, si è osservata una preponderanza di virus R5-tropici sia nel plasma che nei PBMCs dei pazienti naïve, dove la già dimostrata concordanza del tropismo, in questo e in altri studi, è a sostegno delle medesime proporzioni osservate nei due compartimenti. La maggiore differenza è stata riscontrata invece nella popolazione di pazienti *non-responder* al trattamento antiretrovirale: se nel plasma di questi pazienti la percentuale di X4 sembra essere addirittura maggiore rispetto agli R5, nei PBMCs prevale apparentemente la popolazione R5, in accordo con gli altri

compartimenti cellulari. Il limite maggiore di questo dato risiede nel fatto che non è stato possibile confrontare il tropismo di compartimenti provenienti dallo stesso paziente, ma solo la prevalenza nella popolazione generale. Tuttavia, tale osservazione è a sostegno dell'ipotesi che nonostante l'archivio cellulare possa albergare varianti virali molto diverse fra loro, con mutazioni o con tropismo che rispecchiano la storia pregressa del paziente, la situazione a livello plasmatico può essere molto diversa e la variante virale in attiva replicazione potrebbe derivare da *reservoirs* diversi dalle cellule circolanti oppure essere selezionata favorevolmente, sebbene sia presente nei PBMCs come variante minoritaria.

Indipendentemente dal compartimento in analisi, la predizione del tropismo di HIV-1 a partire dalla sequenza di V3 non può non tener conto di tutte le informazioni derivanti da questa regione del genoma virale. G2P *coreceptor* è, attualmente, il più diffuso metodo di interpretazione della sequenza V3, che tiene conto della similarità della sequenza amminoacidica primaria e di residui carichi in posizioni chiave del *loop*. La presenza di amminoacidi carichi, d'altra parte, è fortemente connessa alla carica netta di V3 ed è per tale ragione che il principale obiettivo del nostro lavoro si è focalizzato sullo studio di questo parametro in associazione a G2P e al tropismo ottenuto con il test fenotipico, notoriamente il metodo di elezione. L'evidenza di una relazione tra tropismo, carica netta e risultato di G2P è risultata chiara a seguito di numerose stratificazioni dei campioni studiati in base ai diversi fattori, consentendo di concludere che soltanto nel 10% dei casi vi era una discordanza tra carica netta e il risultato di G2P. Dato interessante è



scaturito dalla classificazione dei campioni in base a NC di V3, in quanto il 32% dei campioni costituiva un gruppo *outset*, con NC compresa tra 4 e 5. Tali sequenze dal tropismo “indeterminato” hanno destato rilevante interesse, nell’ipotesi che potessero rappresentare i ceppi *dual*-tropici identificabili esclusivamente con l’analisi del fenotipo *in vitro*, e che dal punto di vista patogenetico rappresentano l’elemento chiave dello *switch* corecettoriale. La ricerca di un legame tra mutazioni in V3 associate ai diversi gruppi, ottenuti sulla base della NC di V3 e dei valori di FPR, ha messo in rilievo, come atteso, che residui carichi positivamente nelle note posizioni 11 e 25 si associano a valori di carica netta superiore a 5, ma anche che la posizione 32 è coinvolta nella scelta del corecettore e può essere associata a valori di NC >4 e FPR<20%, ovvero a tropismo R5 e DM. L’insieme di questi fattori, dunque, contribuisce ad aumentare l’accuratezza nella predizione del tropismo a partire esclusivamente dalla sequenza di V3.

La conferma che valori di carica netta dal significato criptico (4-5) fossero associati alla capacità del virus di legare entrambi i corecettori, è stata ottenuta a seguito della messa a punto in un saggio fenotipico *in vitro* su cellule *reporter* U87\_CD4+, esprimenti selettivamente CCR5 o CXCR4. Lo studio del fenotipo è stato condotto su dei sottogruppi campione, ovvero rappresentativi di sequenze concordanti, discordanti e indeterminate descritte dal confronto tra NC e G2P nella prima parte del lavoro. Il clonaggio delle sequenze in un vettore di espressione ha consentito di ottenere singole varianti virali da ciascun campione, con il considerevole vantaggio di eliminare ambiguità nella sequenza che

potessero lasciar spazio a un'analisi inesatta e consentendo, invece, di associare inequivocabilmente le caratteristiche genotipiche al tropismo. Inoltre, l'ottenimento di ceppi virali con diverso tropismo a partire da uno stesso campione, ha confermato la potenziale coesistenza di differenti varianti che costituiscono un *pool* eterogeneo di virus, in particolare nei PBMCs. Tali varianti, seppur con sequenze e NC in V3 estremamente simili, possono utilizzare diversamente i corecettori, dando origine a diverso fenotipo *in vitro*.

Il test fenotipico ha permesso, dunque, di confermare che NC di V3 compresa tra 4 e 5 può essere associata, nel 50% dei casi, a virus *dual-tropici* e che G2P è in grado di identificare questi ceppi come X4 solo in una minoranza di casi (25%), supportando la necessità di combinare i diversi metodi per un'analisi più approfondita, in particolare laddove la NC sia rappresentata da un valore *border line*. Lo studio delle regioni all'esterno di V3 può contribuire ad aumentare la specificità della predizione, sebbene vi siano ancora delle incertezze per la determinazione dei ceppi DM. L'elaborazione di un modello basato su marcatori genetici nelle regioni variabili e costanti di Env rappresenta un strumento alternativo allo studio del fenotipo *in vitro*, ma allo stato delle conoscenze attuali non sono stati ancora identificati tutti i markers determinanti per la discriminazione tra ceppi R5 e DM. I risultati riportati in questo lavoro, infatti, confermano che le sequenze con un punteggio superiore a +12 (considerando tutti i parametri noti nelle regioni variabili e costanti di Env), sono proprie di virus che utilizzano il corecettore CCR5, ma circa la metà sono in grado di sfruttare anche il legame con il CXCR4, seppur in

misura inferiore, in base a quanto emerso dai dati sulla luminescenza. Tuttavia, è da chiarire il ruolo *in vivo* di questi virus, dal momento che la “preferenza” per un corecettore che si leghi alla gp120 con maggior affinità potrebbe favorire nettamente la replicazione esclusiva di un ceppo, rendendo replicative, solo *potenzialmente*, le altre varianti del *pool*. D’altra parte, la sempre più diffusa analisi del tropismo in pazienti *responder* alla ART, e quindi su campioni di DNA provirale, necessita di tener conto di tale aspetto, ancor più dello studio del tropismo su HIV-RNA plasmatico. Da quanto emerso in questo lavoro, varianti R5 e *Dual-tropiche* possono coesistere nello stesso *pool* provirale e possono quindi essere attivate in caso di pressione selettiva, in quanto abili nello sfruttare entrambi i corecettori. D’altra parte, la somministrazione degli inibitori del CCR5 in combinazione ad altri farmaci, garantisce una miglior soppressione della replicazione virale, riducendo fortemente il rischio di selezionare dei virus che sfuggano al controllo farmacologico.

Infine, lo studio dei marcatori genotipici in Env associati al tropismo, è limitato, allo stato attuale delle conoscenze, alle regioni variabili V1, V2 e V3. L’analisi condotta sulle altre regioni variabili e costanti non sembra aumentare l’accuratezza nella predizione, confermando il ruolo chiave delle suddette regioni nella scelta del corecettore. Tuttavia, la distinzione tra ceppi R5 e DM rimane ancora da definire e necessita dell’identificazione di ulteriori fattori chiave che siano in grado di “guidare” lo switch corecettoriale, e che sarebbero pertanto estremamente utili come fattore predittivo per l’avanzamento della malattia.

## 5. BIBLIOGRAFIA

1. Guttman M, Kahn M, Garcia NK, Hu SL, Lee KK. Solution structure, conformational dynamics, and CD4-induced activation in full-length, glycosylated, monomeric HIV gp120. **J Virol.** **2012 Aug;****86(16):8750-64**
2. Leitner T, Kumar S, Albert J. Tempo and mode of nucleotide substitutions in gag and env gene fragments in human immunodeficiency virus type 1 populations with a known transmission history. **J Virol.** **1997;****71:4761-70**
3. Sakai K, Ma XY, Volsky DJ. Low-cytopathic infectious clone of human immunodeficiency virus type-I (HIV-I). **FEBS Lett** **1988;** **238:257-261**
4. Fenyo EM, Albert J, Asjo B. Replicative capacity, cytopathic effect and cell tropism of HIV. **AIDS** **1989;** **3:S5-S12**
5. Koot M, Keet IP, Vos AH, et al. Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4<sup>+</sup> cell depletion and progression to AIDS. **Ann Intern Med** **1993;** **118:681-688**
6. Nielsen C, Pedersen C, Lundgren JD, Gerstoft J. Biological properties of HIV isolates in primary HIV infection: consequences for the subsequent course of infection. **AIDS** **1993;** **7:1035-1040**
7. Fenyo EM, Fiore J, Karlsson A, et al. Biological phenotypes of HIV-1 in pathogenesis and transmission. **Antibiot Chemother** **1994;** **46:18-24**
8. Fiore JR, Bjorndal A, Peipke KA, et al. The biological phenotype of HIV-1 is usually retained during and after sexual transmission. **Virology** **1994;** **204:297- 303**
9. Karlsson A, Parsmyr K, Aperia K, et al. MT-2 cell tropism of human immunodeficiency virus type 1 isolates as a marker for

response to treatment and development of drug resistance. **J Infect Dis** 1994; **170**:1367–1375

10. Miedema F, Meyaard L, Koot M, et al. Changing virus-host interactions in the course of HIV-1 infection. **Immunol Rev** 1994; **140**:35–72
11. Berger EA. HIV entry and tropism: the chemokine receptor connection. **AIDS** 1997; **11 (Suppl A)**:S3–S16
12. Alkhatib G, Combadiere C, Broder CC, et al. CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. **Science** 1996; **272**:1955–1958
13. Dragic T, Litwin V, Allaway GP, et al. HIV-1 entry into CD4<sup>+</sup> cells is mediated by the chemokine receptor CC-CKR-5. **Nature** 1996; **381**:667–673
14. Bleul CC, Farzan M, Choe H, et al. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. **Nature** 1996; **382**:829–833
15. Oberlin E, Amara A, Bachelier F, et al. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. **Nature** 1996; **382**:833–835
16. Shioda T, Levy JA, Chengmayer C. Macrophage and T-cell line tropisms of HIV-1 are determined by specific regions of the envelope Gp120 gene. **Nature** 1991; **349**:167–169
17. Fouchier RA, Groenink M, Kootstra NA, et al. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. **J Virol** 1992; **66**:3183–3187
18. Douek DC, Brenchley JM, Betts MR, et al. HIV preferentially infects HIV-specific CD4<sup>+</sup> T cells. **Nature**. 2002;**417**:95–8

19. Dunfee RL, Thomas ER, Wang J, et al. Loss of the N-linked glycosylation site at position 386 in the HIV envelope V4 region enhances macrophage tropism and is associated with dementia. **Virology. 2007;367:222–34**
20. Li S, Juarez J, Alali M, et al. Persistent CCR5 utilization and enhanced macrophage tropism by primary blood human immunodeficiency virus type 1 isolates from advanced stages of disease and comparison to tissue-derived isolates. **J Virol. 1999;73:9741–55**
21. Hwang SS, Boyle TJ, Lyerly HK, Cullen BR. Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. **Science. 1991;253:71–4**
22. de Jong JJ, Goudsmit J, Keulen W, et al. Human immunodeficiency virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity. **J Virol. 1992;66:757–65**
23. Milich L, Margolin BH, Swanstrom R. Patterns of amino acid variability in NSI-like and SI-like V3 sequences and a linked change in the CD4-binding domain of the HIV-1 Env protein. **Virology. 1997;239:108–18**
24. Huang W, Toma J, Fransen S, et al. Coreceptor tropism can be influenced by amino acid substitutions in the gp41 transmembrane subunit of human immunodeficiency virus type 1 envelope protein. **J Virol. 2008;82:5584–93**
25. Frange P, Galimand J, Goujard C, Deveau C, Ghosn J, Rouzioux C, Meyer L, Chaix ML. High frequency of X4/DM-tropic viruses in PBMC samples from patients with primary HIV-1 subtype-B infection in 1996-2007: the French ANRS CO6 PRIMO Cohort Study. **J Antimicrob Chemother. 2009;64(1):135-41.**

26. Verhofstede C, Brudney D, Reynaerts J, Vaira D, Fransen K, De Bel A, Seguin-Devaux C, De Wit S, Vandekerckhove L, Geretti AM. Concordance between HIV-1 genotypic coreceptor tropism predictions based on plasma RNA and proviral DNA. **HIV Med.** **2011;12(9):544-52**
27. Naif HM, Li S, Alali M, et al. CCR5 expression correlates with susceptibility of maturing monocytes to human immunodeficiency virus type 1 infection. **J Virol.** **1998;72:830-6**
28. Zhu TF, Muthui D, Holte S, et al. Evidence for human immunodeficiency virus type 1 replication in vivo in CD14(+) monocytes and its potential role as a source of virus in patients on highly active antiretroviral therapy. **J Virol.** **2002;76:707-16**
29. Sonza S, Mutimer HP, Oelrichs R, et al. Monocytes harbour replication-competent, non-latent HIV1 in patients on highly active antiretroviral therapy. **AIDS.** **2001;15:17-22**
30. Chalmet K, Dauwe K, Foquet L, et al. Presence of CXCR4-using HIV-1 in patients with recently diagnosed infection: correlates and evidence for transmission. **J Infect Dis** **2012; 205:174-184**
31. Shepherd JC, Jacobson LP, Qiao W, et al. Emergence and persistence of CXCR4-tropic HIV-1 in a population of men from the Multicenter AIDS Cohort Study. **J Infect Dis** **2008; 198:1104-1112**
32. Karlsson A, Parsmyr K, Aperia K, et al. MT-2 cell tropism of human immunodeficiency virus type 1 isolates as a marker for response to treatment and development of drug resistance. **J Infect Dis** **1994; 170:1367-1375**
33. Waters L, Mandalia S, Randell P, et al. The impact of HIV tropism on decreases in CD4 cell count, clinical progression, and

- subsequent response to a first antiretroviral therapy regimen. **Clin Infect Dis** 2008; 46:1617–1623
34. Hogg RS, Yip B, Chan KJ, et al. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. **J Am Med Assoc** 2001; 286:2568–2577
  35. Symons J, Vandekerckhove L, Paredes R, et al. Impact of triplicate testing on HIV genotypic tropism prediction in routine clinical practice. **Clin Microbiol Infect** 2012; 18:606–612
  36. De Luca A. HIV-1 co-receptor tropism evolution in naive patients undergoing successful ART: concordance of DNA vs. RNA and triplicate versus singlicate population sequencing. **Rev Antivir Ther Infect Dis. Barcelona; 2012**
  37. Clevestig P, Pramanik L, Leitner T, Ehrnst A. CCR5 use by human immunodeficiency virus type 1 is associated closely with the gp120 V3 loop N-linked glycosylation site. **J Gen Virol.**2006 Mar;87(Pt 3):607-12.
  38. Yokoyama M, Naganawa S, Yoshimura K, Matsushita S, Sato H. Structural dynamics of HIV-1 envelope Gp120 outer domain with V3 loop. **PLoS One.** 2012;7(5):e37530. **Epub 2012 May 18.**
  39. Swenson LC, Mo T, Dong WW, Zhong X, Woods CK, Jensen MA, Thielen A, Chapman D, Lewis M, James I, Heera J, Valdez H, Harrigan PR. Deep sequencing to infer HIV-1 co-receptor usage: application to three clinical trials of maraviroc in treatment-experienced patients. **J Infect Dis.** 2011 Jan 15;203(2):237-45.
  40. De Jong JJ, De Ronde A, Keulen W, Tersmette M, Goudsmit J. Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution. **J Virol.** 1992 Nov;66(11):6777-80



41. De Wolf F, Hogervorst E, Goudsmit J, Fenyö EM, Rübsamen-Waigmann H, Holmes H, Galvao-Castro B, Karita E, Wasi C, Sempala SD, et al. Syncytium-inducing and non-syncytium-inducing capacity of human immunodeficiency virus type 1 subtypes other than B: phenotypic and genotypic characteristics. WHO Network for HIV Isolation and Characterization. **AIDS Res Hum Retroviruses**. 1994 Nov;**10(11):1387-400**.
42. Simon B, Grabmeier-Pfistershammer K, Rieger A, Sarcletti M, Schmied B, Puchhammer-Stöckl E. HIV coreceptor tropism in antiretroviral treatment-naïve patients newly diagnosed at a late stage of HIV infection. **AIDS**. 2010 Aug **24;24(13):2051-8**.
43. Parisi SG, Andreoni C, Sarmati L, Boldrin C, Buonomini AR, Andreis S, Scaggiante R, Cruciani M, Bosco O, Manfrin V, d'Ettorre G, Mengoli C, Vullo V, Palù G, Andreoni M. HIV coreceptor tropism in paired plasma, peripheral blood mononuclear cell, and cerebrospinal fluid isolates from antiretroviral-naïve subjects. **J Clin Microbiol**. 2011 Apr;**49(4):1441-5**. Epub 2011 Mar 2.
44. de Mendoza C, Rodriguez C, García F, Eiros JM, Ruíz L, Caballero E, Aguilera A, Leiva P, Colomina J, Gutierrez F, del Romero J, Agüero J, Soriano V; Spanish HIV Seroconverter Study Group. Prevalence of X4 tropic viruses in patients recently infected with HIV-1 and lack of association with transmission of drug resistance. **J Antimicrob Chemother**. 2007 Apr;**59(4):698-704**. Epub 2007 Feb 27
45. Eshleman SH, Husnik M, Hudelson S, Donnell D, Huang Y, Huang W, Hart S, Jackson B, Coates T, Chesney M, Koblin B. Antiretroviral drug resistance, HIV-1 tropism, and HIV-1 subtype among men who have sex with men with recent HIV-1 infection. **AIDS**. 2007 May **31;21(9):1165-74**.

46. Kassaye S, Johnston E, McColgan B, Kantor R, Zijenah L, Katzenstein D. Envelope coreceptor tropism, drug resistance, and viral evolution among subtype C HIV-1-infected individuals receiving nonsuppressive antiretroviral therapy. **J Acquir Immune Defic Syndr.** 2009 Jan 1;50(1):9-18.
47. Richman DD, Bozzette SA. The impact of the syncytium-inducing phenotype of human immunodeficiency virus on disease progression. **J Infect Dis** 1994; 169:968–974
48. Saracino, A., et al. 2010. Are the proposed env mutations actually associated with resistance to maraviroc? **J. Acquir. Immune Defic. Syndr.** 53:550–552.
49. Monno L, Saracino A, Scudeller L, Punzi G, Brindicci G, Altamura M, Lagioia A, Ladisa N, Angarano G. Impact of mutations outside the V3 region on coreceptor tropism phenotypically assessed in patients infected with HIV-1 subtype B. **Antimicrob Agents Chemother.** 2011 Nov;55(11):5078-84. **Epub 2011 Aug 29.**
50. Brumme ZL, Goodrich J, Mayer HB, Brumme CJ, Henrick BM, Wynhoven B, Asselin JJ, Cheung PK, Hogg RS, Montaner JS, Harrigan PR. Molecular and clinical epidemiology of CXCR4-using HIV-1 in a large population of antiretroviral-naive individuals. **J Infect Dis.** 2005 Aug 1;192(3):466-74. **Epub 2005 Jun 23.**
51. Raymond, S., et al. 2010. CXCR4-using viruses in plasma and peripheral blood mononuclear cells during primary HIV-1 infection and impact on disease progression. **AIDS** 24:2305–2312.