



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA
BIOMEDICINA

**RELACIÓN ENTRE LA CARGA VIRAL RESIDUAL EN PLASMA Y EL ESTADO DE
METILACIÓN DEL LTR VIRAL EN INDIVIDUOS INFECTADOS POR EL VIRUS DE
LA INMUNODEFICIENCIA HUMANA BAJO TRATAMIENTO ANTIRRETROVIRAL**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS

PRESENTA:

M. EN C. CORTÉS RUBIO CÉSAR NOÉ

TUTOR PRINCIPAL DE TESIS: DR. SANTIAGO ÁVILA RÍOS
INSTITUTO NACIONAL DE ENFERMEDADES RESPIRATORIAS, INER

COMITÉ TUTOR: DRA. LEONOR HUERTA HERNÁNDEZ
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS, UNAM

COMITÉ TUTOR: DR. ALFONSO DUEÑAS GONZÁLEZ
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS, UNAM

MÉXICO, CD. MX. MARZO, 2020.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE MEDICINA

OFICIO CPCB/119/2020

ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión ordinaria del Subcomité de Biología Experimental y Biomedicina del Posgrado en Ciencias Biológicas, celebrada el día **12 de agosto de 2019** se aprobó el siguiente jurado para el examen de grado de **DOCTOR EN CIENCIAS** del estudiante **CORTÉS RUBIO CÉSAR NOÉ** con número de cuenta **302032345** con la tesis titulada **“RELACIÓN ENTRE LA CARGA VIRAL RESIDUAL EN PLASMA Y EL ESTADO DE METILACIÓN DEL LTR VIRAL EN INDIVIDUOS INFECTADOS POR EL VIRUS DE LA INMUNODEFICIENCIA HUMANA BAJO TRATAMIENTO ANTIRRETROVIRAL”**, realizada bajo la dirección del **DR. SANTIAGO ÁVILA RÍOS**, quedando integrado de la siguiente manera:

Presidente: DRA. MAYRA FURLAN MAGARIL
Vocal: DR. FELIPE VACA PANIAGUA
Secretario: DR. ALFONSO DUEÑAS GONZÁLEZ
Suplente: DRA. MARÍA DE LA PAZ SÁNCHEZ JIMÉNEZ
Suplente: DR. FEDERICO ÁVILA MORENO

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
“POR MI RAZA HABLARA EL ESPÍRITU”
Ciudad Universitaria, Cd. Mx., a 04 de febrero de 2020

COORDINADOR DEL PROGRAMA


DR. ADOLFO GERARDO NAVARRO SIGÜENZA



c. c. p. Expediente del alumno(a)

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS
UNIDAD DE POSGRADO

Edificio D, 1º Piso. Circuito de Posgrados, Ciudad Universitaria
Alcaldía Coyoacán. C. P. 04510 CDMX

Tel. (+5255)5623 7002 <http://pcbiol.posgrado.unam.mx/>

Agradecimientos Institucionales.

Agradezco al Posgrado en Ciencias Biológicas, por su excelente plan de estudios.

Agradezco al Consejo Nacional de Ciencia y Tecnología, por la beca otorgada (CVU 289343).

Al Dr. Santiago Ávila Ríos, por aceptarme en formar parte de su grupo de trabajo, por su tiempo y su confianza. Mil gracias.

Al Dr. Alfonso Dueñas Gonzáles y a la Dra. Leonor Huerta Hernández, quienes fungieron como mi comité tutor. Gracias por sus comentarios y críticas para mejorar este trabajo.

Agradecimientos a Título Personal

A todos los pacientes que gentilmente aceptaron participar en el protocolo B26-09 (sin su donación de sangre, este proyecto no hubiera sido posible).

A todos los profesores que, desde mi etapa básica hasta la profesional, me transmitieron sus conocimientos.

En especial agradezco al Dr. Víctor Valdés López, quien fue mi profesor de Biología Molecular en dos ocasiones y que me hizo interesarme en dicha área.

A muchos ciudadanos mexicanos que con sus impuestos financiaron mi acceso a los estudios de posgrado en la UNAM, la beca que me fue otorgada por el CONACyT y a este proyecto.

Al Dr. Gustavo Reyes Terán, Jefe del CIENI por aceptarme en formar parte de su Centro de Investigación.

Al Biol. Gonzalo Salgado Montes de Oca, por haber concebido este gran proyecto y por creer en que yo podría abordarlo. Mil gracias.

A la QBP. Margarita Matías Florentino por fungir como asesora en secuenciación masiva. Valió la pena el trabajar desde muy temprano hasta la tarde-noche, el acudir incluso sábados, domingos y días festivos. Mil gracias.

A Francisco Javier Prado Galbarro, por haber realizado los análisis de Bioestadística. Mil gracias.

A Robert Edgar, creador del programa USEARCH, por su asesoramiento en el uso de dicho programa. Mil gracias.

A Christopher Ormsby y Ana Paola por diseñar el algoritmo utilizado en el análisis de datos de secuenciación masiva. Mil gracias.

A la Dra. Ma de la Paz Sánchez Jiménez, por su consejo de agregar varios controles en el ensayo de Metilación. Gracias por revisar la tesis a detalle.

Al LIBB Diego Adrian Oliva Rico, por sus consejos acerca del tratamiento con bisulfito de sodio. Mil gracias.

A la Dra. Maribel Soto Nava, por su asesoría y ayuda en varios temas relacionados con este proyecto. Gracias por donarme el sobrenadante del virus NL4-3, requerido para la curva estándar en la medición de la carga viral residual.

A Ramón Hernández Juan, por cuantificar los estándares de NL4-3. Mil gracias

A la Dra. Olivia Briseño por darme consejos acerca del análisis de citometría de flujo multiparamétrica. Mil gracias.

A mis amigos Fernando Ocotoxtle Morales, Felipe Arguijo Pérez, Alma Rincón Rubio, Daniela Tapia Trejo, Anahí Sánchez León y Silvia del Arenal por compartir experiencias y por hacer agradable la estancia.

A mi Familia: Luisa, Manuel, Ivette,
Manuel Jr, Javier y Edith por
apoyarme en todos mis proyectos. A
mis sobrinos Karlita y Ricardo para
alentarlos a que se interesen por la
investigación científica.

Índice general

0.1. Resumen	1
0.2. Abstract	4
1. Introducción.	7
1.1. La metilación CpG.	7
1.2. ¿Cómo se puede conocer el estado de metilación en una región del genoma?.	9
1.3. La metilación no CpG.	10
1.4. Antecedentes: El Papel de la metilación en la activación del reservorio del VIH.	11
2. Objetivos.	15
3. Hipótesis.	16
4. Metodología.	17
4.1. Aprobación ética.	17
4.2. Pacientes.	17
4.3. Medición de la carga viral residual en plasma.	18
4.4. Medición de carga proviral a partir de células T CD4+ totales.	18
4.5. Método de análisis de la metilación del LTR viral.	19
4.5.1. Conversión con bisulfito	19
4.6. Amplificación del LTR 5'.	22
4.7. Secuenciación de siguiente generación del LTR	23
4.8. Alineamiento de lecturas (reads) y filtrado.	23
4.9. Análisis de la metilación del LTR.	24
4.10. Inmunofenotipificación.	24
4.11. Análisis estadístico.	25
5. Resultados.	31
5.1. Perfil longitudinal de la metilación del LTR 5' presente en las células T CD4+ totales provenientes de personas bajo TAR a corto plazo.	31
5.2. Patrones temporales en la metilación del LTR 5'	37

5.3. Características basales y de seguimiento en personas con distintos patrones de metilación del LTR 5'.	41
6. Discusión.	61
7. Conclusiones.	70
8. Consideraciones éticas.	72
9. Fondos.	73
10. Contribuciones de los autores.	74
Appendices	75
A. A-E	76
A.1. Dos patrones Globales de Metilación CpG: Metilación Constante e Incrementada.	76
B.2. AFECP: Factores.	77
C.3. AFECP: Carga de Factores	78
D.4. AFECP: Rotación	79
Literatura Citada	80
E.5. Artículo requisito para la obtención de grado.	90

Índice de figuras

1.1. Tratamiento del ADN con bisulfito de sodio	9
4.1. Estrategia de amplificación y de secuenciación del LTR 5' del VIH. . .	27
4.2. Estrategia de compuertas "gating" para definir a las subpoblaciones de células T CD4+.	29
5.1. Perfiles longitudinales de la metilación del LTR 5' del VIH en células T CD4+ totales provenientes de doce participantes bajo TAR a corto plazo	34
5.2. Tendencias de la metilación del LTR 5' del VIH en el tiempo considerando a dos participantes modelo.	44
5.3. Composición detallada de las variantes del LTR 5' del VIH y análisis de metilación del LTR 5' en el resto de los participantes	51
5.4. Cambios en el índice de metilación CpG, la carga proviral y la carga viral residual en plasma a lo largo del seguimiento del estudio.	55
5.5. Cambios en los patrones de metilación del LTR 5' a lo largo del tiempo.	56
5.6. Análisis Factorial Exploratorio utilizando Componentes Principales. . .	60

Índice de cuadros

1.1. Estado del arte en la investigación del nivel de metilación del LTR de VIH.	14
4.1. Historia del tratamiento antirretroviral de los participantes en el estudio.	20
4.2. Primers utilizados para amplificar y secuenciar la región del LTR 5´ del VIH	21
5.1. Características basales de los grupos de metilación constante y metilación incrementada bajo TAR a corto plazo comparados con el grupo de TAR a largo plazo.	36
5.2. Cambios tempranos en los patrones de metilación y su asociación con las variables clínicas.	52
5.3. Variación diferencial de las variables clínicas a lo largo del tiempo entre los grupos de metilación constante y metilación incrementada	59
A.1. Dos patrones Globales de Metilación CpG: Metilación Constante e Incrementada.	76
A.2. AFECP: Factores.	77
A.3. AFECP: Carga de Factores.	78
A.4. AFECP: Rotación	79

Lista de Abreviaturas.

- AFCEP/EFAPC: Análisis Factorial Exploratorio utilizando Componentes Principales;
- AID: Desaminasa de Citosina Inducida por Activación;
- APOBEC: Enzima editora de la apolipoproteína B mediante ARNm semejante al polipéptido catalítico;
- DAPK1: Proteína cinasa 1, asociada a muerte;
- EC: Controlador de Élite;
- ESCs: Células troncales embrionarias;
- HDAC2: Desacetilasa de histonas;
- HDACi: Inhibidor de desacetilasa de histona;
- H3K27me3: Marca epigenética en Histonas. Tri-metilación en la Lisina 27 de la Histona 3;
- LINE: Elemento Nucleotídico Interdisperso Largo;
- LTNP: No progresor;
- LTR: Región terminal repetida larga;
- MBD2: Proteínas nucleares que se caracterizan por la presencia de un dominio de unión a sitios CpG metilados (MBD);
- mCpG: Sitios CpG metilados;
- NGS/SSG: Secuenciación de siguiente generación;
- NHEJ: Reparación por recombinación no homóloga;
- NNRTI: Inhibidores de la reverso transcriptasa de tipo no nucleósidos;
- NRTI: Inhibidores de la reverso transcriptasa de tipo nucleósidos;
- PBMC: Células mononucleares de sangre periférica;
- PCR: Reacción en cadena de la polimerasa;
- PI: Inhibidores de proteasa;
- PRC2: Complejo Represor Polycomb 2 (PcG);
- PSSM: Matriz de puntuación específica de posición;
- pVL: Carga Viral Residual en Plasma;
- RIC: Rango intercuartil;
- SINE: Elemento Nucleotídico Interdisperso Corto;
- TAR: Tratamiento antirretroviral;
- T_{CM} : Células T de memoria central;
- T_{EM} : Células T de memoria efectora;
- TN: Naive al tratamiento antirretroviral;
- T_{NEW} : Células T memorias nuevas;
- T_{SCM} : Células T de memoria Stem Cell;
- TSS: Sitio de inicio de la transcripción;
- T_{TE} : Células T terminalmente efectoras;
- T_{TM} : Células T de memoria transicional;
- VIH: Virus de la inmunodeficiencia humana;

-delta % de células T CD4+ = % de células T CD4+ 48 meses - % de células T CD4+ basal.

-delta % de células T CD8+ = % de células T CD8+ 48 meses - % de células T CD8+ basal.

-delta del cociente CD4/CD8 = cociente CD4/CD8 48 meses - cociente CD4/CD8 basal.

Relación entre la Carga Viral Residual en plasma y el estado de Metilación del LTR Viral en individuos infectados por el virus de la inmunodeficiencia humana bajo tratamiento antirretroviral

por

César Noé Cortés Rubio

0.1. Resumen

La persistencia de los provirus competentes para la replicación en estado latente e integrados en las células T CD4+ provenientes de individuos infectados con el virus de la inmunodeficiencia humana (VIH) bajo tratamiento antirretroviral (TAR) es el principal obstáculo para la erradicación del virus. La metilación de la región promotora, llamada terminal repetida larga (específicamente la región del LTR 5'), ha sido propuesta como un posible mecanismo que contribuye a la latencia del VIH; sin embargo, existen observaciones contradictorias respecto a su relevancia. En este trabajo, evaluamos los perfiles de metilación del LTR 5' en células T CD4+ totales provenientes de sangre de 12 participantes bajo TAR a corto plazo (30 meses) seguidos longitudinalmente durante dos años, y un grupo transversal de participantes bajo TAR a largo plazo (6-15 años), utilizando secuenciación de siguiente generación (SSG). Fueron construidas las librerías a partir de los amplicones del LTR (regiones U3-R) tanto de muestras tratadas y no tratadas con bisulfito de sodio. Para cuantificar y evaluar las diferencias en la metilación global, estimamos una métrica llamada "Índice de metilación CpG", promediando la proporción de secuencias metiladas para cada uno de los 9 sitios CpG canónicos presentes en el LTR de la referencia HXB2 en cada muestra.

Para cada punto de tiempo, calculamos el índice de metilación CpG global ponderada y determinamos las cargas provirales, las cargas virales residuales en plasma (pVL) y las distribuciones de las subpoblaciones de células T CD4+.

Posteriormente utilizamos el Análisis Factorial Exploratorio utilizando Componentes Principales (AFECP) como un método de reducción de variables para describir la estructura de datos y resumir la varianza de las variables clínicas basales y de seguimiento en menos dimensiones.

Resultados

Observamos la presencia de metilación en la región del LTR 5' en todos los participantes, la cual se comportó dinámicamente. El número de variantes del LTR 5' encontradas por muestra varió de 1 a 13, con una mediana en la profundidad de secuenciación de 16 270x (RIC 4 107x-46 760x). Se observó un aumento significativo en

la metilación del LTR 5' al mes 42 en comparación con el mes 30 (mediana del índice de metilación CpG: 74.7% vs 0%, $p=0.025$), posteriormente observamos un incremento aún más significativo en la metilación del LTR 5' al mes 48 con respecto al mes 30 (mediana del Índice de Metilación CpG: 81.8% vs. 0%, $p=0.003$).

Posteriormente se observaron dos patrones globales de metilación CpG: 1) la metilación estuvo presente durante todo el período de seguimiento y fue alta ("grupo de metilación constante") o 2) la metilación aumentó significativamente en algún momento (alrededor del mes 42 y 48) a lo largo del período de estudio ("grupo de metilación incrementada"). La separación en estos dos grupos considerando la variable metilación fue significativa a los meses 30, 33, 39 y 54 del seguimiento (en los meses 30 y 54, valores $p<0.01$ y a los meses 33 y 39, valores $p<0.05$).

Mas adelante, nosotros observamos una disminución general en la carga proviral con respecto al tiempo de TAR. Las comparaciones a los meses 42, 48 y 54 vs mes 30 ($p=0.005$, $p=0.00097$ y $p=0.002$ respectivamente). Más específicamente, encontramos que la disminución en la carga proviral (fue significativa únicamente en el grupo de metilación incrementada) al mes 48, 54 vs mes 30 ($p=0.031$ en ambas comparaciones).

Aplicamos el Análisis Factorial Exploratorio utilizando Componentes Principales (AFECP) con rotación varimax con el objetivo de explicar la estructura de los datos. En conjunto, los tres primeros factores explicaron el 50.86% de la varianza común total. De estos tres primeros factores, solo el factor 2 (que explicó el 16.11% de la varianza común total), mostró evidencia de separación a lo largo del tiempo ($p=0.04$). En este análisis, las variables $\%T_{SCM}$, $\%T_{NEW}$, carga proviral e índice de metilación CpG tuvieron una varianza única alta (>0.5 , en cada variable).

Las personas dentro del grupo de metilación incrementada vs metilación constante fueron significativamente más jóvenes ($p<0.001$), tuvieron una mayor ganancia en el recuento de CD4+ a los 48 meses después del inicio del TAR (numérico $p<0.001$ y en porcentaje $p=0.041$), mayor disminución en las cuentas de CD8+ a los 48 meses (numérico $p=0.023$ y en porcentaje $p<0.001$), un mayor incremento en la relación CD4/CD8 a los 48 meses ($p<0.001$), y un mayor número de cambios de régimen de TAR ($p=0.008$).

En conjunto, estas observaciones muestran que las diferencias temporales en los patrones de Metilación en el LTR 5' se asocian con la combinación de características clínicas basales y de seguimiento. Dentro de las características basales destacan: Edad, ganancia de células T CD4+ a los 48M de TAR, disminución de células T CD8+ a los 48M de TAR, el número de cambios en el régimen de TAR, el tiempo transcurrido para lograr una carga viral indetectable, la clase de tercer fármaco en el régimen de TAR y el nadir del recuento de células T CD4+ y dentro de las características del seguimiento destacan: disminución significativa de la carga proviral durante el seguimiento, tiempo bajo TAR, la composición de las subpoblaciones de células T CD4, el porcentaje de células T CD4+ de memoria, naive, T_{CM} y T_{TM} y el porcentaje de células T CD4+ T_{EM} y T_{TE} .

Las personas con TAR a largo plazo mostraron un elevado grado de metilación en

general (mediana del índice de metilación CpG: 78 %; RIC 71-87 %). No se observaron diferencias en la carga viral residual en plasma o en la carga proviral comparando a los individuos con TAR a corto plazo (tanto a los 30 o 42 meses) y el grupo TAR a largo plazo.

Conclusiones

Nuestro estudio muestra evidencia de que la metilación en el LTR 5' del VIH presente en las células T CD4+ totales se comporta de manera dinámica a lo largo del tiempo y que esta variable puede seguir diferentes patrones temporales (a los que denominados metilación constante y metilación incrementada) y que estos patrones están asociados con una combinación de características clínicas basales y de seguimiento. En su conjunto nuestros resultados pueden explicar las diferencias observadas en los estudios previos realizados por otros autores (donde encontraron LTRs hipermetilados, hipometilados o no metilados en individuos VIH+ bajo TAR supresora).

Palabras clave

VIH; LTR 5'; latencia; Metilación del ADN; CpG; longitudinal; Carga Proviral; Carga Viral Residual en Plasma; Metilación no canónica.

Relationship between Residual Viral Load in plasma and Methylation status of Viral LTR in individuals infected with human immunodeficiency virus under antiretroviral treatment

by

César Noé Cortés Rubio

0.2. Abstract

Persistence of latent, replication-competent provirus in CD4+ T cells of human immunodeficiency virus (HIV)-infected individuals on antiretroviral treatment (ART) is the main obstacle for virus eradication. Methylation of the proviral 5' long terminal repeat (LTR) promoter region has been proposed as a possible mechanism contributing to HIV latency; however, conflicting observations exist regarding its relevance. We assessed 5'-LTR methylation profiles in total CD4+ T cells from blood of 12 participants on short-term ART (30 months) followed up for two years, and a cross-sectional group of participants with long-term ART (6-15 years), using next generation sequencing (NGS). The libraries were constructed from purified (U3-R regions) LTR amplicons (both from bisulfite-treated and untreated DNA). To quantify and evaluate differences in global methylation, and estimated a "CpG Methylation Index" metric, averaging the proportion of methylated sequences for each of the 9 canonical HXB2 LTR CpG sites in each sample.

For each time point, we calculate the weighted global CpG methylation index and determine the proviral load, the residual viral load in plasma (pVL) and the distributions of CD4+ T cell subpopulation. Later we used Exploratory Factor Analysis using Principal Components (EFAPC) as a variable reduction method to describe data structure and summarize the variance in baseline and follow-up clinical variables into fewer dimensions.

Results

5'-LTR methylation was observed in all participants and behaved dynamically. The number of 5'-LTR variants found per sample ranged from 1 to 13, with median sequencing depth of 16 270x (IQR 4 107x-46 760x). An overall significant 5'-LTR methylation increase was observed at month 42 compared to month 30 (median CpG Methylation index: 74.7 % vs 0 %, $p=0.025$), subsequently we observed an even more significant increase in 5'-LTR methylation to month 48 with respect to month 30 (median CpG Methylation Index: 81.8 % vs. 0 %, $p = 0.003$).

Subsequently, two global CpG methylation patterns were observed: 1) the methylation was present during all follow-up period and was high ("constant methylation group") or 2) the methylation increased significantly at some time (around month 42 and 48) throughout the study period ("methylation increase group"). The separation in these two groups considering the methylation variable was significant at months 30, 33, 39 and 54 of the follow-up (at months 30 and 54, p values <0.01 and at months 33 and 39, p values<0.05).

Subsequently we observed a general decrease in the proviral load with respect to the time of ART. The comparisons at months 42, 48 and 54 vs month 30 (p=0.005, p=0.00097 and p=0.002 respectively). More specifically, we found that the decrease in proviral load (was significant only in the methylation increase group) at month 48, 54 vs. month 30 (p=0.031 in both comparisons).

Using Exploratory Factor Analysis using Principal Components(EFAPC) with varimax rotation in order to explain the structure of the data. Together, the first three factors explained 50.86 % of the total variance. Of these first three factors, only factor 2 (which explained 16.11 % of the total variance), showed evidence of separation over time (p =0.04). In this analysis, the variables % T_{SCM} , % T_{NEW} , proviral load and CpG methylation Index had a high unique variance (>0.5, in each variable).

Persons in the methylation increase group vs constant methylation group were significantly younger (p <0.001), had higher CD4+ T cell gain at 48 months after the start ART (numerical p<0.001 and in percentage p=0.041), greater decrease in CD8+ accounts at 48 months after the start ART (numerical p=0.023 and in percentage p <0.001), a greater change in the CD4/CD8 ratio after 48 months on ART (p<0.001), and a greater number of ART regime changes (p=0.008).

Together, these observations show that temporal differences in methylation patterns in the 5'-LTR are associated with the combination of baseline and follow-up clinical features. Among the baseline characteristics are: Age, gain of CD4+ T cells at 48M on ART, decrease CD8+ T cells at 48M on ART, the number of changes in the ART regimen, the time elapsed to achieve an undetectable viral load, the third drug class in the ART regimen and the nadir of CD4+ T cell count and within the characteristics of the follow-up are: significant decrease of the proviral load during the follow-up, time under ART, the composition of the CD4+ T cell subpopulations, percentage of CD4+ T cells memory, naive, T_{CM} and T_{TM} and the percentage of CD4+ T cells T_{EM} and T_{TE} .

Persons with long-term ART showed overall high methylation (median CpG Methylation Index: 78 %; IQR 71-87 %). No differences were observed in residual plasma viral load or proviral load comparing individuals on short-term (both at 30 or 42 months) and the long-term ART.

Conclusions

Our study shows evidence that HIV 5'-LTR methylation in total CD4+ T cells behaves dynamically along time and that it can follow different temporal patterns (which are called constant methylation and methylation increased) and that these patterns are associated with a combination of baseline and follow-up clinical characteristics. Taken together, our results may explain the differences observed between previous studies by other authors (where they found hypermethylated, hypomethylated or non-methylated LTRs in HIV+ individuals under suppressor ART).

Keywords

HIV; 5'-LTR; latency; DNA methylation; CpG; longitudinal; proviral load; residual plasma viral load; non-canonical methylation.

Capítulo 1

Introducción.

”Todo lo que se constata como veraz para *E. coli* también debe ser cierto para los elefantes”.

Jacques Monod.

1.1. La metilación CpG.

Hace millones de años, en procariontes apareció un proceso epigenético llamado la metilación del ADN que hizo a las bacterias capaces de distinguir entre su propio ADN y el proveniente de un bacteriófago. Este proceso se ha conservado evolutivamente y en eucariontes ejerce diversas funciones entre las que se encuentran: la represión transcripcional, el silenciamiento de secuencias retrovirales endógenas, la participación en la diferenciación celular [1] y su ausencia en sitios metilables lo convierte en un Patrón Molecular asociado a Patógenos que puede ser reconocido por el Sistema Inmune [2].

La metilación en el ADN es un proceso exclusivamente enzimático, donde una enzima llamada ADN metil-transferasa (DNMT) cataliza la transferencia de un grupo metilo desde el cofactor S-adenosil-L-metionina (SAM) al Carbono 5 de una Citosina en un contexto 5´...CpG...3´ (la metilación en este contexto se presenta tanto en bacterias como en mamíferos). Esto da como resultado a una 5-metil-Citosina (5mC). La base modificada no interfiere con el apareamiento de bases del tipo Watson-Crick y ésta puede ser reconocida por proteínas que interactúan con el ADN [3].

Dos mecanismos han sido propuestos para explicar cómo la metilación en el ADN atenúa la expresión: 1) Las proteínas de unión a las CpG metiladas tipo 1 y 2, pueden unirse a los sitios CpG metilados específicamente dentro de una secuencia que es reconocida por los factores de transcripción y por tanto el factor de transcripción, por ejemplo Sp1, es incapaz de unirse al ADN [4, 5] y 2) Los factores de transcripción, como

es el caso de NFKb, pierde su afinidad por los sitios metilados [6].

En mamíferos, en menor medida, puede ocurrir también la metilación no CpG (en sitios como CpA, CpT y CpC). Dentro del genoma de mamíferos, la metilación no CpG ha sido encontrada ya sea independiente pero también coexistiendo con la metilación CpG. La enzima DNMT1, es la enzima responsable de preservar y mantener el patrón de metilación en sitios CpG, actuando después de la replicación del ADN previo al proceso de división celular (esta enzima es conocida como una metil-transferasa de mantenimiento) [3, 7]. La metilación CpG es simétrica, es decir que, si se presenta el dinucleotido CpG metilado en una cadena, también la metilación estará presente en la otra cadena a manera de espejo (mCpG: GpCm). Este tipo de metilación es somáticamente heredable.

Dado que únicamente ciertos sitios CpG llegan a ser metilados, esto lleva a la generación de patrones de metilación específicos del linaje celular. A manera de ejemplo, en el genoma humano existen alrededor de 56 millones de sitios CpG y de éstos entre un 60-80 % son metilados y esto corresponde aproximadamente a un 34-45 % de todas las citosinas [8, 9, 10]. Dado que muchos de los sitios CpG se encuentran metilados, se considera que el estatus de expresión de la mayoría de los genes es "off".

En los genomas de mamíferos, los dinucleótidos CpG se encuentran sub-representados en relación a otros dinucleótidos, debido a la propiedad mutagénica de la base 5mC [11, 12]. La 5mC puede desaminarse hidrolíticamente a Timina [13] y la Citosina no metilada puede desaminarse a Uracilo, aproximadamente a una velocidad 24 veces más lenta [14]. El sistema de reparación del ADN es capaz de detectar a un Uracilo (al no ser un componente natural en el ADN) y lo corrige insertando a una Citosina no metilada; en cambio resulta más difícil la detección y la corrección de un producto de la desaminación de 5mC debido a que la Timina, es un componente natural en el ADN [3]. Por tanto, se piensa que, durante el curso de la evolución del genoma humano, el dinucleotido CpG ha sido depletado (subrepresentado estadísticamente) [11, 12, 15, 16], sin embargo, es común que los encontremos formando islas CpGs (que son regiones con alta densidad de sitios CpGs) en las regiones promotoras de cerca de un 70 % de los genes en el genoma, ejecutando una función reguladora muy importante que es la de controlar la transcripción [17]. La metilación del ADN en regiones promotoras de genes permite la represión transcripcional [9, 10].

Dado que cerca de un 45 % del genoma humano es compuesto de transposones, incluidos SINEs, LINEs y LTRs [18], la metilación del ADN además está involucrada en la protección del genoma contra parásitos genéticos [19]. Dado que los transposones y las secuencias repetidas de ADN son ricas en sitios CpG, éstas son altamente metiladas [20].

1.2. ¿Cómo se puede conocer el estado de metilación en una región del genoma?.

Para determinar el estado de metilación de una región específica en el genoma básicamente se trata el ADN con bisulfito de sodio. Si una Citosina no se encuentra metilada, ésta se desamina a Uracilo; en cambio si la Citosina se encuentra metilada, ésta queda protegida del tratamiento y no se desamina, permaneciendo como Citosina (**Figura 1.1**). Este tratamiento, además, hace que las cadenas del ADN ya no sean complementarias al 100 %, por lo que posteriormente se debe seleccionar a una de las cadenas, se diseña el par de primers para amplificar la zona de interés mediante PCR y como paso siguiente se debe realizar la secuenciación de los fragmentos. Finalmente, es necesaria la comparación entre las secuencias tratadas con bisulfito, las no tratadas (control) y una referencia para determinar el estatus de metilación en el ADN.

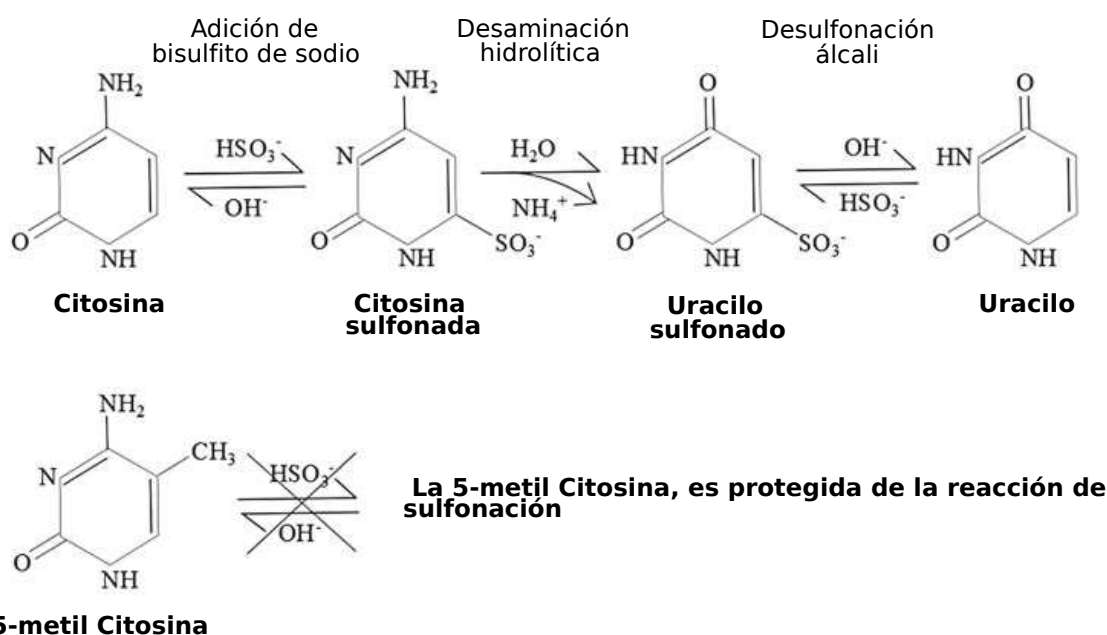


Figura 1.1: El tratamiento del ADN con bisulfito de sodio nos permite conocer el estado de metilación en el ADN. Si la Citosina no se encuentra metilada, el tratamiento con bisulfito de sodio y otros reactivos permiten que se lleven a cabo las reacciones de desaminación y desulfonación, transformando a la Citosina no metilada a Uracilo (**panel superior**), y posteriormente es detectada como Timina en la PCR. En cambio, si la Citosina se encuentra metilada, esta queda protegida del tratamiento (**panel inferior**), por lo que las dos reacciones anteriores no tienen lugar y la Citosina permanece sin cambio.

1.3. La metilación no CpG.

Un estudio *in vivo* utilizando un knockout para DNMT1 en células troncales embrionarias de ratón demostró que la metilación no-CpG puede ser mantenida independientemente de DNMT1, por lo que debía existir otra enzima capaz de realizarla [21]. En mamíferos, estudios knockdown han mostrado que los sitios no CpG pueden ser metilados mediante la enzima DNMT3a/3b [22], pero los detalles de su establecimiento, mantenimiento y la función biológica de este tipo de metilación no es del todo claro [23, 24].

Las enzimas DNMT3a y DNMT3b son las enzimas responsables de generar el patrón de metilación inicial durante el desarrollo de mamíferos y en las líneas germinales (esta enzima es conocida como metil-transferasa *de novo*). La enzima DNMT3 metila a las Citosinas preferentemente en un contexto CpG, sin embargo se ha demostrado que esta enzima puede también metilar en sitios no CpG [21, 25].

La metilación no CpG (específicamente las del tipo CH y CHH) es asimétrica [26], por ejemplo consideremos a una célula que presenta el dinucleótido metilado 5'...mCpA...3' en una cadena de su ADN, y en la cadena complementaria tendremos 3'...GpT...5' (mCpA:GpT), posteriormente durante la replicación del ADN, las cadenas se abrirán con el fin de que ocurra la síntesis y cada cadena servirá como molde. En el primer caso, la cadena 5'...mCpA...3', servirá como molde y generará una secuencia complementaria con el dinucleótido 3'...GpT...5' (mCpA:GpT) sin embargo en el segundo caso 3'...GpT...5', generará una cadena complementaria 5'...CpA...3' (GpT:CpA) (observar que en este caso la Citosina no estará metilada). Este tipo de metilación ha sido observada en células troncales embrionarias de ratón [26]. Por tanto la asimetría indica que es improbable la existencia de un mecanismo que mantenga a este tipo de metilación y más aún que sea heredable en todas las células hijas (el ejemplo anterior mostró que solo 1 de 2 células conservará el patrón de este tipo de metilación). Por tanto, es probable que la metilación no CpG necesite ser restablecida *de novo* después de cada división celular para que sea mantenida, lo que indica que este tipo de metilación tendría una gran importancia en células de larga vida [27].

Algunos estudios han revelado una sustancial metilación no CpG en algunos tipos celulares provenientes de mamíferos incluyendo células troncales embrionarias humanas (hESC, por sus siglas en inglés) [9, 10, 21, 28, 29], iPSC [29, 30] y en ovocitos [31, 32]. Trabajos más recientes muestran que la metilación no CpG no se restringe a células pluripotentes. Este tipo de metilación es particularmente abundante en el tejido cerebral de humano y de ratón [33, 34, 35]. Por tanto, el asumir que la metilación no CpG disminuye o simplemente es inexistente conforme se da el proceso de diferenciación podría no ser del todo cierto ya que esta marca está presente en tejidos completamente diferenciados donde la metilación no CpG podría jugar un papel importante en la regulación génica [36].

Aunque la metilación no CpG tiene mayor libertad de presentarse en varios con-

textos, es más frecuentemente encontrada en el dinucleótido CpA. Varios estudios de secuenciación por bisulfito del genoma completo han mostrado la frecuencia con la que ocurre la metilación en general: CpG > CpA > CpT > CpC [9, 10, 29, 33, 37]. Mientras que los sitios CpGs son metilados simétricamente en las dos cadenas [10], los sitios no CpGs son asimétricos por lo que pueden ser metilados únicamente en una de las cadenas [9, 38]. En 2014 Guo W., et al encontraron que la metilación no CpG ocurre en cadena específica dentro de intrones, SINEs (particularmente en sitios Alu) y LINEs, mientras que en los LTRs no se presenta esta especificidad de cadena [38]. Además, encontraron que en humanos, la metilación en contexto CHG y CHH (H significa cualquier nucleótido excepto Guanina) tienen una similar secuencia preferencial sugiriendo que ambos patrones de metilación deben ser realizados por la misma ADN metil-transferasa [38].

Ya que la metilación no CpG está asociada con la DNMT3a, DNMT3b y DNMT3L y es independiente de DNMT1, entonces la poca metilación no CpG presente en células somáticas puede ser debida a los bajos niveles de DNMT3L comparado con ESCs [22].

En el 2012 Yu M., et al mostraron que la base hidoximetil-citosina (hmC) está ausente en sitios no CpG [39].

¿Cuál podría ser el papel de la metilación no CpG? Para intentar contestar esta pregunta, en el 2001 Malone C., et al encontraron que la metilación no CpG del tipo CpmCpNpGpG dentro del promotor del gen B29 fue capaz de reprimir la actividad del promotor en células B humanas al bloquear la unión del factor de transcripción EBF [40]. Por tanto en el ejemplo anterior, la metilación no CpG en promotores de genes estaría asociada con la reducción en su expresión.

En la línea de investigación en VIH no hay estudios publicados respecto a la detección y al papel de la metilación no CpG y su probable implicación en regular al LTR viral.

1.4. Antecedentes: El Papel de la metilación en la activación del reservorio del VIH.

La combinación moderna del tratamiento antirretroviral (TAR) es capaz de reducir la carga viral en plasma (pVL) del virus de inmunodeficiencia humana (VIH) a niveles indetectables (<40 copias de ARN de VIH/mL), logrando así una reducción de 4-5 log₁₀ durante el primer año de tratamiento [41]. Sin embargo, la carga proviral del VIH es más estable con solo una reducción de 1 log₁₀ en el mismo período de tratamiento [42].

La persistencia de provirus infecciosos en las células T CD4+ (y tal vez en otros linajes celulares) en individuos infectados por el VIH bajo TAR es el principal obstáculo para lograr la erradicación del virus [43]. El principal reservorio del VIH-1 es un pequeño grupo de células T CD4+ de memoria en reposo con infección latente [44], que están presentes en una frecuencia de 1/10⁶ de células T CD4+ en reposo en personas bajo TAR supresor [45].

Las estrategias de cura del VIH del tipo "patear y matar" ("kick and kill") tienen como objetivo el lograr una activación completa de las células con infección latente para que puedan ser eliminadas mediante diversas estrategias inmunológicas y farmacológicas. Hasta el día de hoy, estas estrategias han mostrado un éxito limitado [46].

Para complicar los esfuerzos de cura del VIH, se estima que el 88 % de los provirus son defectuosos; mientras que solo el 12 % constituyen a los provirus intactos [47], que, en teoría, tendrían la capacidad de activarse y producir viriones. Además, se ha logrado la reactivación *ex vivo* de solo una pequeña fracción de los provirus latentes (0.12 %) a partir de células T CD4+ en reposo provenientes de individuos bajo TAR, utilizando estrategias basadas en vorinostat, un inhibidor de desacetilasas de histonas (HDACi) [46]. *In vivo*, el tratamiento con panobinostat, otro HDACi, no logró una reducción en toda la cohorte respecto al ADN total del VIH, el ADN del VIH integrado, la proporción de células portadoras de virus competentes para la replicación, ni en el número de células con infección latente [48].

Los resultados obtenidos en estos y en otros trabajos hacen necesario replantear la estrategia seguida debido básicamente a dos aspectos: 1) la baja eficiencia en la reactivación viral y 2) la alta toxicidad sistémica, al emplear agentes activadores inespecíficos.

Una explicación plausible de por qué ciertos provirus no se reactivan después de la estimulación puede deberse al hecho de que el silenciamiento epigenético del VIH podría ser el producto de un patrón complejo y heterogéneo de marcas de histonas [49, 50, 51, 52], de metilación del ADN [53, 54, 55, 56, 57, 58, 59], del efecto de los complejos remodeladores de la cromatina como Polycomb [52, 60, 61], de los factores asociados con BRG1 (BAF) [62, 63], de los ARN no codificantes [64, 65, 66, 67, 68], sin mencionar a los factores genéticos y microambientales asociados con el sitio de integración viral [47]. De hecho, se ha observado que las estrategias basadas en HDACi para activar al reservorio latente de VIH, podrían, paradójicamente, incrementar el grado de metilación del ADN de la región terminal repetida larga (LTR) 5' del VIH y por lo tanto, esto podría aumentar la estabilidad del reservorio [59]. Por lo tanto, se justifica un conocimiento más profundo sobre los mecanismos individuales de latencia del VIH, así como sobre sus interrelaciones.

Los patrones de metilación del ADN son definidos tempranamente en el desarrollo y están estrechamente relacionados tanto con la diferenciación celular, como con la expresión génica, siendo éstos modificados en muchos estados patológicos, como el cáncer, la autoinmunidad y en enfermedades infecciosas [69]. Existe evidencia de que el ADN de los retrovirus integrados en los genomas de los mamíferos puede llegar a ser metilado *de novo* [70]. Además, la metilación CpG de los promotores y de enhancers retrovirales localizados en el LTR 5' ha sido correlacionada con el silenciamiento de varios retrovirus tales como el virus de la leucemia de células T humanas tipo 1 [71], el virus de la leucemia murina Moloney [72], el virus del sarcoma de Rous [73] y los retrovirus endógenos humanos [74]. Estas observaciones han inspirado a varios grupos de investigación a explorar el papel de la metilación del LTR en la infección por el VIH.

Entre los primeros estudios que analizaron el estado de la metilación de la región promotora LTR 5' del VIH utilizaron como modelo a la línea celular Jurkat J-Lat (línea celular infectada con VIH; modelo de latencia). Encontraron que la línea celular, poseía un LTR 5' de VIH, hipermetilado, específicamente en dos islas CpG cercanas al sitio de inicio de la transcripción [54] (**Cuadro 1.1**).

Posteriormente se encontraron resultados contrastantes con respecto al papel de la metilación de la región promotora LTR 5' del VIH en la expresión de los provirus provenientes individuos sin y bajo TAR supresor (**Cuadro 1.1**). Se sugirió que las células T CD4+ de memoria en reposo provenientes de individuos bajo TAR a largo plazo (mediana de 11.5 años), sin viremia detectable, a menudo contenían LTR 5' hipermetilados y resistentes a la activación, en oposición a LTR 5' hipometilados en individuos virémicos [53] (**Cuadro 1.1**). Sin embargo, otro estudio cuestionó estas observaciones, y descubrió que la metilación del LTR 5' es rara en células T CD4+ en reposo de individuos avirémicos bajo TAR a corto plazo (mediana de 2.9 años) [57] (**Cuadro 1.1**).

Más recientemente, otro estudio encontró bajos niveles de metilación del ADN en el LTR 5' dentro de células T CD4+ en reposo provenientes de personas que recibieron TAR a corto plazo (mediana de 2.3 años). Sin embargo, también se observó un alto grado de metilación (probablemente acumulación) en el LTR 5' en el reservorio latente después de un TAR a largo plazo (mediana de 12.5 años) y se sugirió que la estimulación transitoria de las células que albergan a los provirus latentes puede contribuir a la metilación del LTR 5' [59] (**Cuadro 1.1**).

Así, se necesitan más estudios para determinar en qué medida la metilación del LTR 5' desempeña un papel en la latencia del VIH, con el fin de proveer más información útil para el diseño de estrategias de cura.

Estado del arte en la investigación del nivel de metilación del LTR del VIH.

AUTOR [REFERENCIA]	AÑO	MODELO	GRUPO (S) ESTUDIO	TIPO CELULAR	TIEMPO BAJO TAR (MEDIANA, AÑOS)	TIPO DE TAR	TIEMPO DE INFECCIÓN (AÑOS)	METILACIÓN DEL LTR 5'	METILACIÓN GLOBAL	PRINCIPALES ASOCIACIONES O LA AUSENCIA DE ESTA
Kauder S., [54]	2009	Línea celular	-	-Línea celular J-Lat -Células T CD4+ primarias con infección latente	-	-	-	Hipermetilación	-	-Se encontró hipermetilado el LTR 5' en las dos islas CpG cercanas al TSS. -Las proteínas MDB2 y HDAC2 fueron reclutadas al LTR vía CpG metiladas (en la segunda isla).
Blazkova J., [53]	2009	Células de pacientes VIH+	1.-Avirémicos (<50 copias) 2.- Virémicos	Células T CD4+ de memoria en reposo	11.5	-2 NRTI+ NNRTI -2 NRTI+PI -TN -NRTI+NNRTI+PI	11-16	1-Avirémicos: Hipermetilación (19-100% de mCpG)* 2- Virémicos: Hipometilación (0.1% de mCpG)	-	-Se encontró hipermetilado el LTR 5' en los individuos avirémicos. -Se encontró hipometilado el LTR 5' en los individuos virémicos. -Se encontró una correlación negativa entre los niveles de metilación del LTR 5' y el porcentaje de provirus reactivables <i>ex vivo</i> . <i>*El autor no dió mucha importancia al encontrar individuos avirémicos con baja metilación.</i>
Blazkova J., [57]	2012	Células de pacientes VIH+	Individuos bajo TAR con viremia indetectable (<50 copias)	Células T CD4+ en reposo de sangre periférica	2.9	*Ver referencia	-	Hipometilación (Mediana de 2.4% mCpG)	-	-Se encontró hipometilado el LTR 5'. -Se encontró hipermetilado el gen Env (envoltura). <i>Esto indica que el ADN viral integrado fue metilado.</i>
Palacios J., [55]	2012	Células de pacientes VIH+	1-LTNP/EC 2-Progresores avirémicos (<50 copias)	PBMCs infectadas	1.-No 2.- 2	1.-No 2.-N/D	1.-20 2.-21	Hipometilación en ambos grupos	Similares en ambos grupos	-Se encontró hipometilado el LTR 5' en el grupo de LTNP/EC (5.4% mCpG). -Se encontró hipometilado el LTR 5' en el grupo de progresores avirémicos (0-1 % mCpG). -Se encontró una correlación positiva entre el porcentaje de metilación en el LTR y el tiempo de infección (en el grupo LTNP/EC). <i>-No se encontró correlación entre carga viral vs porcentaje de metilación ni entre niveles de células T CD4+ vs porcentaje de metilación.</i>
Weber S., [58]	2014	Células de pacientes VIH+	1-EC 2-LTNP 3-Progresores (SIDA y crónicos)	PBMCs infectadas	No	Naive a TAR	-	Hipometilación en las regiones 3' (LTR, nef, rev, tat, gp41) y 5' (LTR y gag)	Revisaron 55/94 CpGs	-Se encontró muy bajo porcentaje de metilación en los tres grupos de estudio. <i>-Se trata del primer trabajo que intentó realizar un seguimiento, sin embargo solo lograron amplificar el 38% de las muestras (al ser regiones largas) y por tanto no lograron secuenciar una región específica del mismo paciente en más de una ocasión.</i>
Trejbalova K., [59]	2016	Células de pacientes VIH+	Individuos bajo TAR (carga viral <320 copias)	Células T CD4+ en reposo	1.-2.3 2.-12.5	*Ver referencia	*Ver referencia, material suplementario	1.-Hipometilación (2-8% mCpG) 2.-De hipometilación a hipermetilación (0-45% mCpG)	-	-Se encontró hipometilado el LTR 5' en los pacientes del grupo 1: de 1 a 2 sitios mCpG por molécula. -El grupo 2, mostró mayor grado de metilación en comparación al grupo 1; sin embargo 6 pacientes mostraron de 0 a 3% de mCpG. -Encontraron una correlación positiva entre el tiempo de TAR y el grado de metilación. <i>-La anterior correlación puede deberse tanto al tiempo de TAR y/o tiempo de infección.</i>

Cuadro 1.1: **TSS**-Sitio de inicio de la transcripción; **EC**-Controlador de Elite; **LTNP**-No progresor; **TN**-Naive a TAR; **MDB2**-Dominio de la proteína 2 de unión a sitios CpG metilados; **HDAC**-Desacetilasa de histonas; **mCpG**- Sitios CpG metilados; **PBMC**-células mononucleares de sangre periférica "-" Ausente y N/D, no determinado.

Capítulo 2

Objetivos.

En este trabajo, deseamos implementar una metodología basada en secuenciación de siguiente generación, capaz de analizar detalladamente la composición de las variantes del LTR 5' de VIH y su estado de metilación con alta sensibilidad en las células T CD4+ provenientes de muestras de sangre de personas infectadas con VIH con una historia corta de TAR supresor en comparación con personas bajo TAR supresor a largo plazo y finalmente buscamos una posible asociación entre los patrones de metilación y las características clínicas basales y de seguimiento de los participantes.

Capítulo 3

Hipótesis.

Los pacientes que presenten un LTR 5' de VIH hipometilado, tendrán un alto grado/mayor porcentaje de la carga proviral, de carga viral residual de VIH y de células de memoria.

Capítulo 4

Metodología.

”Todo ser vivo es también un fósil.
Dentro de él, todo el camino hasta la
estructura microscópica de sus
proteínas conserva las huellas y hasta
los estigmas de su ascendencia”.

Jacques Monod.

4.1. Aprobación ética.

Este estudio fue revisado y aprobado por el Comité de Ética en Investigación del Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas (INER) (registro: DHHS IORG0003948; CONBIOETICA-09-CEI-003-20160427) en la Ciudad de México (código del proyecto: B26-09), y se realizó de acuerdo con los principios de la Declaración de Helsinki. Todos los participantes dieron su consentimiento informado por escrito antes de la donación de muestras de sangre.

4.2. Pacientes.

Este estudio incluyó dos diseños paralelos, uno longitudinal y uno transversal. Se obtuvieron muestras de sangre longitudinales de 12 personas infectadas por el VIH con una historia corta de TAR supresor (21-30 meses con TAR y carga viral <40 copias/mL). Las muestras fueron obtenidas en intervalos de tres meses durante 2 años.

Adicionalmente, se obtuvieron muestras de sangre transversales de 10 personas infectadas por el VIH con una larga historia de TAR supresor (6-15 años). La historia del tratamiento antirretroviral de todos los participantes en este estudio puede revisarse en

el (**Cuadro 4.1**).

Todas las muestras de sangre fueron obtenidas y procesadas en el Centro de Investigación en Enfermedades Infecciosas (CIENI) del INER en la Ciudad de México. El plasma y el ADN de las células T CD4+ totales de cada muestra de sangre fueron almacenados. La carga viral en plasma para el seguimiento clínico fue determinada utilizando el sistema m2000 (Abbott, Abbott Park, IL, EE.UU.). Los conteos de células T CD4+ y CD8+ fueron realizados mediante citometría de flujo utilizando el kit Tru-Count en un equipo FACSCanto II (BD Bioscience, San José, CA, EE.UU.). Ambos procesos se realizan rutinariamente como parte del seguimiento clínico de los pacientes atendidos en el CIENI, INER.

4.3. Medición de la carga viral residual en plasma.

El ARN viral fue extraído a partir de 9 mL de plasma. El plasma fue centrifugado a 15 000 rpm durante 3 h, a 4°C. El sobrenadante fue removido y el pellet fue resuspendido en 140 μ L. El ARN viral fue purificado a partir del pellet utilizando el kit QIAamp viral RNA mini kit (Qiagen, Valencia, CA, EE.UU.; se resuspendió en 41 μ L de buffer de elución) y se almacenó a -80°C hasta su uso. La carga viral residual en plasma (<40 copias de ARN/mL) fue determinada mediante PCR en tiempo real (qPCR) utilizando TaqMan Fast Virus One-Step Master Mix (ThermoFisher, Waltham, MA, EE.UU), y un estándar casero de ARN derivado de cultivos de la cepa viral de referencia NL4-3.

El ARN derivado de NL4-3 fue extraído del sobrenadante del cultivo (QIAamp viral RNA mini kit, QIAGEN) y fue cuantificado (sistema m2000, Abbott). La curva estándar fue generada al realizar diluciones 1:2 que iban de 256 a 1 copia de ARN. La qPCR fue realizada por triplicado para los puntos de curva estándar y por duplicado para los puntos de las muestras utilizando los primers Fwd 5'-GGTCTCTCTGGTTAGACCAGAT-3' (posiciones en HXB2: 455-476), Rv 5'-CTGCTAGAGATTTTCCACACTG-3' (635-614) y la sonda TaqMan 5'-6FAM-AGTAGTGTGTGCCCGTCTGTT-TAMRA-3' (552-572), dirigidas al LTR viral, en reacciones de 50 μ L que contenían 20.5 μ L de ARN estándar o de muestra, utilizando un equipo 7500 Real Time PCR System (ThermoFisher). Las condiciones de qPCR fueron las siguientes: 50°C durante 5 min para la transcripción inversa, 95°C durante 20 s, seguido de 40 ciclos de 95°C durante 15 s y 60°C durante 60 s.

4.4. Medición de carga proviral a partir de células T CD4+ totales.

Las células mononucleares de sangre periférica (PBMC, por sus siglas en inglés) fueron obtenidas a partir de 16 mL de sangre mediante centrifugación en gradiente

de Ficoll. Las células T CD4+ totales fueron purificadas mediante selección negativa a partir de 50 millones de PBMC, utilizando el kit EasySep Human CD4+ T Cell (StemCell Technologies, Vancouver, Canadá), y el ADN fue purificado utilizando el kit iPrep Purelink gDNA Blood (ThermoFisher) en un sistema de purificación de ácidos nucleicos iPrep (ThermoFisher). La carga proviral fue medida a partir de células T CD4+ totales utilizando TaqMan Gene Expression Master Mix (ThermoFisher).

Se utilizó un estándar casero de ADN derivado de la línea celular ACH-2 (obtenido a través del Programa de Reactivos sobre SIDA de los Institutos Nacionales de Salud de Estados Unidos; catálogo 349) [75], que contiene dos copias provirales de VIH por célula [76, 77]. Para construir la curva estándar, el ADN fue extraído a partir de 1×10^6 células ACH-2 (iPrep Purelink gDNA Blood kit; ThermoFisher) y fue cuantificado (NanoDrop 1000; Thermo Scientific). La curva estándar fue obtenida con 12 diluciones seriales 1:3 a partir de un stock de ADN que contiene 400 ng, por triplicado. La qPCR fue realizada a partir de 400 ng totales de ADN de células T CD4+, utilizando los mismos primers y la sonda utilizados para la cuantificación de la carga viral residual en plasma (**Sección 4.3**). Se utilizó la cuantificación de la RNasa P como referencia interna (Taqman RNase P Control Reagents kit; ThermoFisher). Las condiciones de PCR fueron las siguientes: 95°C durante 10 min, seguido de 45 ciclos de 95°C durante 15 y 60°C durante 1 min. Todos los valores fueron corregidos por la pureza de las células T CD4+, evaluado mediante citometría de flujo multiparamétrica como se explica en (**Sección 4.10**). , y se ajustó a 1×10^6 de genomas.

4.5. Método de análisis de la metilación del LTR viral.

4.5.1. Conversión con bisulfito

La metilación se evaluó mediante la conversión con bisulfito de sodio utilizando 900 ng de ADN de células T CD4+ totales, utilizando el kit EpiTect Fast DNA Bisulfite (Qiagen), mediante el uso del instrumento QIAcube (Qiagen). Las muestras fueron eluidas en 16 μL de buffer de elución. El volumen completo de ADN tratado con bisulfito fue usado inmediatamente para realizar la primera ronda de PCR LTR-Gag por duplicado, como se describe a continuación. Se utilizó el kit Human Methylated & Non-methylated DNA Set (Zymo Research, Irvine, CA, EE. UU) como control para evaluar la eficacia de la conversión por bisulfito, de acuerdo con las especificaciones del fabricante, utilizando la enzima ADN polimerasa Taq Platinum (ThermoFisher). Los adaptadores de Illumina necesarios para la construcción de librerías sujetas a secuenciación fueron agregados a los primers recomendados por el fabricante (**Cuadro 4.2**).

Historia del tratamiento antirretroviral de los participantes en el estudio.

ID del Paciente	TAR inicial		Cambios de TAR	
	Fecha	Régimen de TAR	Fecha	Régimen de TAR
Grupo de TAR a corto plazo (seguimiento longitudinal)				
TP23	27-02-12	TDF+FTC+EFV	03/08/12	ABC+3TC+EFV
TP24	04/04/12	TDF+FTC+EFV	11/09/12	TDF+FTC+ATV/r
TP25	13/04/12	TDF+FTC+EFV	05/12/12	ABC+3TC+EFV
TP26	04/04/12	TDF+FTC+EFV	-	
TP27	26/04/12	TDF+FTC+EFV	29/10/13 21/08/14	ABC+3TC+ATV/r TDF+FTC+EFV
TP28	16/04/12	TDF+FTC+EFV	30/10/12 06/08/13	TDF+FTC+ATV/r TDF+FTC+NVP
TP29	26/04/12	TDF+FTC+EFV	-	
TP30	03/05/12	TDF+FTC+EFV	-	
TP31	14/05/12	TDF+FTC+EFV	-	
TP32	13/03/12	TDF+FTC+EFV	-	
TP33	21/05/12	TDF+FTC+EFV	-	
TP34	13/12/12	TDF+FTC+AZT+RTV	-	
Grupo de TAR a largo plazo (transversal)				
TP35	21/03/04	AZT+3TC+EFV	15/08/13	TDF+FTC+EFV
TP36	06/07/01	AZT+3TC+IDV	15/10/2008	TDF+FTC+EFV
TP37	1997	AZT+ddl+IDV/r	15/10/04 15/11/04 15/09/05 15/03/09 15/09/16	AZT+3TC+EFV AZT+3TC+NFV AZT+3TC+EFV+LPV/r TDF+FTC+LPV/r TDF+FTC+DRV/r
TP38	15/01/03	AZT+3TC+EFV	2010	TDF+FTC+LPV/r
TP40	15/01/06	AZT+3TC+EFV	2009	TDF+FTC+EFV
TP41	15/05/02	ddl+d4T+IDV	15/05/02 2009	ABC+3TC+EFV TDF+FTC+EFV
TP43	15/08/09	TDF+FTC+EFV	-	
TP44	15/11/10	TDF+FTC+ATV/r	15/05/16	TDF+FTC+DRV/r
TP45	26/07/10	TDF+FTC+EFV	-	
TP46	19/02/02	AZT+3TC+NFV	15/09/04 15/10/09	ddl+3TC+EFV TDF+FTC+EFV

Cuadro 4.1: TAR-tratamiento antirretroviral; **Inhibidores de la Reverso Transcriptasa Nucleósidos (NRTI)**: TDF-Tenofovir; FTC-Emtricitabina; AZT-Zidovudina (Azidotimidina); 3TC-Lamivudina; ddl-Didanosina; d4T-Estavudina y ABC-Abacavir. **Inhibidores de la Reverso Transcriptasa No Nucleósidos (NNRTI)**: EFV-Efavirenz y NVP-Nevirapina. **Inhibidores de Proteasa (PI)**: RTV-Ritonavir; IDV-Indinavir; NFV-Nelfinavir; ATV/r-Atazanavir/Ritonavir; DRV/r-Darunavir/Ritonavir y LPV/r-Lopinavir/Ritonavir. Los regímenes de TAR actuales están en **negritas**. Los participantes sin cambios en el régimen de TAR son mostrados con ”-”

Primers utilizados para amplificar y secuenciar la región del LTR 5´ del VIH y los controles de Metilación de DAPK1.

PRIMERS PARA SECUENCIAR EL VIH		
Nombre del primer	Secuencia del primer	Posiciones respecto a HXB2
LTR1-20Fw	TGGAAGGGCTAATTCCTCC	1-20
LTR1-20TxFw	TGGAAGGGTTAATTATTTT	1-20
LTR835-856Rv	GCCTTAACCGAATTTTTTCCC	835-856
LTR836-856TxRv	ACCTTAACCAAAATTTTTTCC	836-856
Nex-LTR213-236Fw	<i>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG</i> GTGAGCCTGCATGGAATGGATGAC	213-236
Nex-LTR213-236TxFw	<i>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG</i> GTGAGTTTGTATGGAATGGATGAT	213-236
Nex-LTR471-495Rv	<i>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG</i> GTCTAACCAGAGAGACCCAGTACAG	471-495
Nex-LTR471-495TxRv	<i>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG</i> ATCTAACCAAAAAACCCAATACAA	471-495
PRIMERS PARA SECUENCIAR LOS CONTROLES DE METILACION DAPK1		
Nombre del primer	Secuencia del primer	Posiciones respecto al gen DAPK1
Nex-DAPK1_-867_-836Fw	<i>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG</i> ATTGGGAAGTTAAGGYGGAGGGAAATTTGGT	-867 to -836
Nex-DAPK1_-628_-594Rv	<i>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG</i> CCCCAAACRAACAATCCCCAAAACACATTCTA	-628 to -594

Cuadro 4.2: Las secuencias de los adaptadores de Illumina Nextera XT se muestran en itálicas en rojo. Los cambios en las secuencias de los primers para tener en cuenta en la conversión con bisulfito se muestran en azul. Fw-primers Forward; Rv-primers Reverse; Tx-Primers utilizados en las muestras convertidas con bisulfito de sodio.; LTR-región terminal repetida larga. El gen de la proteína cinasa asociada a muerte-1 humana (DAPK1) fue usado para medir la eficiencia de la conversión por bisulfito de sodio. "Y , R" son posiciones degeneradas que consideran secuencias metiladas y no metiladas.

4.6. Amplificación del LTR 5’.

La región viral correspondiente al LTR 5’ fue amplificada utilizando una PCR anidada, mediante el uso de la enzima ADN polimerasa Taq Platinum (ThermoFisher), obteniendo amplicones U3-R tanto de las muestras tratadas con bisulfito de sodio, como de las muestras no tratadas. Las secuencias de los adaptadores de Illumina necesarios para la construcción de librerías para secuenciación fueron incluidas como parte de los primers para la segunda ronda de PCR (**Cuadro 4.2**).

Para las muestras no tratadas con bisulfito de sodio, se utilizaron los primers LTR1-20Fw y LTR835-856Rv (**Cuadro 4.2**) para la primera ronda de PCR, lo que permitió la amplificación del fragmento LTR-gag, mediante las siguientes condiciones: 94°C por 3 min, 35 ciclos de 94°C durante 30s, 50°C durante 30 s, y 72°C durante 1 min, y un paso final de 72°C durante 5 min. A continuación, se muestra la secuencia del amplicón resultante,

>K03455.1 Human immunodeficiency virus type 1 (reference genome)(1-856)

```
TGGAAGGGCTAATTC ACTCCCAACGAAGACAAGATATCCTTGATCTGTGGATCTACCA  
CACAAAGGCTACTTCCCTGATTAGCAGA ACTACACACCAGGGCCAGGGATCAGATATCCAC  
TGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGAGAAGTTAGAAGAAGCCA  
ACAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGAATGGATGACCCGG  
AGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCTAGCATTTCATCACATGGCCCGAG  
AGCTGCATCCGGAGTACTTCAAGAACTGCTGACATCGAGCTTGCTACAAGGGACTTTCGG  
CTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGAT  
CCTGCATATAAGCAGCTGCTTTTTG CCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGA  
GCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCCT  
TGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTA ACTAGAGATCCCTC  
AGACCCTTTTAGTCA GTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGAAAG  
CGAAAAGGGAAACCAGAGGAGCTCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGG  
CAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTGACTAGCGGAGGCTAGA  
AGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGATGGGAA  
AAAATTCGGTTAAGGC
```

Para la segunda ronda de PCR, se utilizaron los primers Nex-LTR213-236Fw y Nex-LTR471-495Rv, permitiendo la amplificación de la región U3-R del LTR 5’, con las siguientes condiciones: 94°C durante 3 min, 35 Ciclos de 94°C durante 30 s, 56°C durante 30 s, y 72°C durante 1 min, y un paso final de 72°C durante 5 min (**Figura 4.1**). A continuación, se muestra la secuencia del amplicón resultante,

>K03455.1 Human immunodeficiency virus type 1 (reference genome)(213-495)

```
GTGAGCCTGCATGGAATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGC  
CGCCTAGCATTTCATCACATGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGA  
CATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGG  
GGGACTGGGGAGTGGCGAGCCCTCAGATCCTGCATATAAGCAGCTGCTTTTTG CCTGTAC  
TGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCT
```

Para las muestras tratadas con bisulfito de sodio, se utilizaron los primers LTR1-20TxFw y LTR835-856TxRv para realizar la primera ronda de PCR con las siguientes condiciones: 94°C durante 3 min, 35 ciclos de 94°C durante 30 s, 48°C durante 30 s y 72°C durante 1 min, y un paso final de 72°C durante 5 min. Para la segunda ronda

de PCR, se utilizaron los primers Nex-LTR213-236TxFw y Nex-LTR471-495TxRv, mediante las siguientes condiciones: 94°C durante 3 min, 35 ciclos de 94°C durante 30 s, 47°C durante 30 s y 72°C durante 1 min, y un paso final de 72°C durante 5 min.

4.7. Secuenciación de siguiente generación del LTR

Las librerías de secuenciación de siguiente generación (SSG) fueron construidas a partir de los amplicones purificados de la región U3-R del LTR (tanto de ADN tratado con bisulfito, así como con ADN no tratado), utilizando la estrategia de amplicón mediante el uso de los kits Nextera XT DNA sample preparation and index (Illumina, San Diego, CA, EE. UU) (**Figura 4.1**). Se agregaron diferentes combinaciones únicas de índices a cada muestra para incluir múltiples muestras diferentes por corrida. Las librerías fueron normalizadas a 60 nM y purificadas en gel de agarosa (bandas de 338-440 pb, kit QIAquick gel extraction, Qiagen).

Después de normalizar nuevamente, las librerías se agruparon, se cuantificaron (Qubit 3.0 Fluorometer; ThermoFisher) y se corrieron en el equipo MiSeq (Illumina), utilizando kits v2 300-cycle pair-end (Illumina), de acuerdo con las instrucciones del fabricante.

4.8. Alineamiento de lecturas (reads) y filtrado.

Las lecturas de secuenciación (reads) fueron ensambladas y filtradas utilizando el programa USEARCH versión 9.2.64 [78]. El siguiente algoritmo fue utilizado: 1) los archivos fastq R1 y R2 fueron fusionados, 2) los reads fueron filtrados por la presencia de los sitios de unión al primer (se permitió una diferencia máxima de 2 nucleótidos) y por tamaño (209-309 nucleótidos para las muestras y 224-324 nucleótidos para los controles de metilación DAPK1), 3) se eliminaron las secuencias del primer previo al filtrado de calidad (24 y 25 nucleótidos en los extremos 5' y 3' respectivamente para las muestras y 32 y 35 para los controles de metilación DAPK1), 4) los reads fueron filtrados por calidad (máximo error esperado=1), 5) se determinó la abundancia de cada secuencia, 6) los singlets (lecturas únicas) fueron descartadas, y 7) los reads con una abundancia <2% fueron filtrados.

Los reads obtenidos con el programa USEARCH mostraron una tasa de conversión del 97.6% (intervalo: 95.1% -100%) de Citosinas fuera de los sitios CpG para las muestras y del 100% para los controles, utilizando el software Quma (Quantification Tool for Methylation Analysis) [79].

4.9. Análisis de la metilación del LTR.

Las lecturas de secuenciación obtenidas para cada punto de tiempo en cada individuo fueron normalizadas a 100 y alineadas a la referencia HXB2 utilizando Clustal Omega [80]. Se construyó una matriz de peso por posición (PSSM) utilizando R y Seqinr [81]. Los 9 sitios CpG canónicos de la región del LTR del VIH fueron localizados en la PSSM y se estimó la proporción de secuencias metiladas (CpG), no metiladas (TpG) y mutadas para cada una. La PSSM también proporcionó información acerca de los sitios de metilación adicionales CpG, CHG y CHH. Estos datos fueron representados mediante múltiples gráficos circulares utilizando el programa Raw Graphs [82].

Utilizamos el software Quma [79], (optimizado para el análisis de reads provenientes de SSG), para cuantificar y evaluar las diferencias en la metilación global, y estimamos una métrica llamada "Índice de Metilación CpG", promediando la proporción de secuencias metiladas para cada uno de los 9 sitios CpG canónicos, presentes en el LTR de HXB2 en cada muestra (**Figura 4.1**). El Índice de Metilación CpG varió de 0 (completamente sin metilar) a 100 (completamente metilado). Las diferencias en la metilación entre las muestras de diferentes puntos de tiempo se evaluaron mediante la prueba U de Mann-Whitney, utilizando Quma. Con el objetivo de ser más exhaustivo en el análisis de metilación, también estudiamos los patrones de metilación no canónica a lo largo del período de seguimiento, estableciendo una medida de metilación en los sitios CpG diferentes a los canónicos, (Índice de Metilación CpG adicional), y la metilación en sitios distintos a CpG (Índice de Metilación no CpG), promediando la proporción de secuencias metiladas para cada posición relevante en cada categoría.

4.10. Inmunofenotipificación.

Se seleccionó un panel de 12 anticuerpos para identificar por citometría de flujo, la proporción de subpoblaciones de linfocitos T CD4+ en cada punto de tiempo para todos los participantes incluidos en el estudio. Los anticuerpos utilizados fueron los siguientes: Biologend: HLA-DR BV785 clona L243 (307641), CD45RO BV650 clona UCHL1 (304231), CD3 BV570 clona UCHT1 (300435), CD95 PECy5 clona DX2 (305610), CCR7 APC Cy7 clona G043H7 (353211), CCR5 BV421 clona J418F1 (359117), PD1 PE clona EH12.2H7 (329906) y CD127 FITC clona A019D5 (351311). BD: CD28 BV711 clona CD28.2 (563131), CD38 AF700 clona HIT2 (560676), CD45 APC clona RUO (340943). Invitrogen: CD4 PE Cy5.5 clona S3.5 (MHCD0418). Life technologies: kit de tinción de células muertas LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (L34957).

Todos los anticuerpos fueron previamente titulados utilizando linfocitos T CD4+ de donadores sanos. Para la tinción, fueron resuspendidas 1×10^6 de células T CD4+ en 2 mL de buffer de tinción (PBS 1x, 2% de Suero Fetal Bovino y 0.5% de EDTA). Las células fueron centrifugadas a 2000 rpm durante 5 minutos y se decantó el sobrena-

dante. La tinción se llevó a cabo en el volumen remanente (aproximadamente 50 μL), agregando primero el anticuerpo anti-CCR7. Después de 10 minutos de incubación, se añadió el resto de los anticuerpos y el tubo fue incubado durante 1 h a temperatura ambiente y en oscuridad. Las células fueron lavadas con 3 mL de buffer de tinción, se centrifugaron a 2000 rpm durante 5 minutos, se decantó el sobrenadante y las muestras fueron fijadas con 0.3 mL de paraformaldehído al 1 %. Las células fueron adquiridas en un citómetro LSR Fortessa (BD).

Las subpoblaciones de linfocitos T CD4+ fueron definidas de la siguiente manera: Células T CD4+naïve(CD3+/CD4+/CD45RO-/CCR7+/CD28+/CD95-), Células T CD4+ de memoria (CD3+/CD4+/CD28+ -/CD95+), Stem Cell Memory (T_{SCM} : CD3+/CD4+/CD45RO-/CCR7+/CD28+/CD95+), Memoria Central (T_{CM} : CD3+/CD4+/CD45RO+/CCR7+/CD28+), Memoria transicional (T_{TM} : CD3 +/CD4+/CD45RO+/CCR7-/CD28+), Memoria Efectora (T_{EM} : CD3+/CD4+/CD45RO+/CCR7-/CD28-), Terminalmente Efectora (T_{TE} : CD3+/CD4+/CD45RO-/CCR7-/CD28-) y T new (T_{NEW} : CD3+/CD4+/CD45RO-/CCR7-/CD28+) utilizando el programa Cytobank [83]. Para las subpoblaciones de células con un número limitado de eventos capturados, utilizamos una distribución de Poisson para establecer el umbral mínimo para el análisis en 100 eventos positivos y 10 % de coeficiente de variación. La estrategia de gating se muestra en la (**Figura 4.2**).

4.11. Análisis estadístico.

Las comparaciones de los parámetros clínicos y de metilación de los datos longitudinales fueron realizados para cada punto de tiempo contra el mes 30 como referencia (tanto considerando a todos los pacientes combinados como separados en dos grupos, el grupo de “metilación incrementada” y el grupo de “metilación constante”; ver (**Sección 5**), y contra el grupo de TAR a largo plazo (los meses 30,42,48 y 54 se muestran en figuras y tablas). Se utilizaron medianas con rangos intercuartiles (RIC) para el análisis descriptivo. Las variables de las muestras relacionadas (diferentes puntos de tiempo provenientes del mismo participante) fueron comparadas mediante la prueba de Wilcoxon y las variables de muestras independientes fueron comparadas utilizando la prueba de Mann-Whitney. Una $p < 0.05$ fue considerada significativa. Los análisis estadísticos se realizaron utilizando SPSS versión 23.0 y STATA versión 14.0.

Se empleó el Análisis Factorial Exploratorio utilizando Componentes Principales (AFECP) como un método de reducción de variables para describir la estructura de los datos y resumir la varianza en las variables clínicas basales y las variables clínicas de seguimiento en menos dimensiones. Se construyeron índices, seguido de rotación ortogonal (varimax) (**Cuadros A.2, A.3, A.4**). Las variables incluidas fueron: edad, nadir de células T CD4+, pVL pre-TAR, tiempo entre el diagnóstico de VIH y el tratamiento, delta de células T CD4+ (cuenta y % a los 48 meses), delta de células T

CD8+ (cuenta y % a los 48 meses), delta del cociente CD4/CD8 a los 48M, tiempo de TAR, tiempo para lograr una pVL indetectable, pVL residual, carga proviral, Índice de Metilación CpG, el número de cambios en el régimen de TAR, “esqueleto” (“backbone”) en el TAR (tenofovir+emtricitabina vs. abacavir+lamivudina), tercer fármaco en el TAR (inhibidores de la RT no nucleósidos vs. inhibidores de proteasa), presencia de blips de carga viral y la proporción de subpoblaciones de células T CD4+: naive, memoria en general, T_{SCM} , T_{CM} , T_{EM} , T_{TM} , T_{NEW} y T_{TE} . (**Cuadros A.2, A.3, A.4**) . Seis factores mostraron un valor eigenvalor mayor que 1 (varianza común acumulada 76.78 %), de los cuales tres fueron seleccionados (varianza común acumulada 50.86 %) para presentación y análisis adicional. Se obtuvieron y graficaron las puntuaciones de los factores. Adicionalmente se obtuvo la varianza única de cada variable. El AFECP se realizó utilizando STATA versión 14.0.

Figura 4.1: ...*de página anterior*. La estructura de la región LTR 5' (basada en la secuencia de referencia HXB2) se detalla en esta figura. Se incluyeron los sitios de unión a factores de transcripción, así como los sitios CpG canónicos. La región del LTR fue amplificada por medio de una PCR anidada a partir del ADN de células T CD4+ totales tratadas con bisulfito de sodio. La primera ronda de PCR generó un producto de 856 pb que incluyó un fragmento de gag, con el fin de amplificar selectivamente el LTR 5' y no el LTR 3'. La PCR anidada fue realizada utilizando primers que contienen a los adaptadores de Illumina para la secuenciación de siguiente generación. Se generó un amplicón de 349 pb que incluyó a las 9 posiciones CpG canónicas presentes en la referencia HXB2 en la región U3-R del LTR y este fragmento generado fue secuenciado a profundidad. Asumiendo que cada lectura (read) corresponde a un solo amplicón, cuantificamos, normalizamos y alineamos todas las variantes obtenidas para cada muestra (el número de lecturas y la proporción en la muestra se indican al lado de cada variante). Cada posición CpG canónica fue representada mediante un círculo. Las posiciones resistentes a la transformación con bisulfito de sodio se consideraron como metiladas (negro) y las posiciones susceptibles a la transformación con bisulfito de sodio como no metiladas (blanco). Las posiciones con mutaciones que causaron la pérdida de sitios CpG fueron también identificadas (gris). Posteriormente, generamos una representación resumida de todas las variantes observadas en cada muestra, utilizando un gráfico de pastel por cada posición susceptible a la metilación (los colores representan la proporción de variantes metiladas, no metiladas y mutadas en ese sitio específico dentro de la muestra). Los sitios fueron numerados según el orden de aparición en la referencia HXB2. También incluimos otros sitios susceptibles a la metilación tales como los CpG adicionales, CHG y CHH, no observados en la secuencia de HXB2 (marcados con *). Finalmente, definimos una métrica llamada índice de metilación CpG por muestra, promediando la proporción de variantes metiladas en cada uno de los nueve sitios CpG canónicos analizados dentro del LTR 5'.

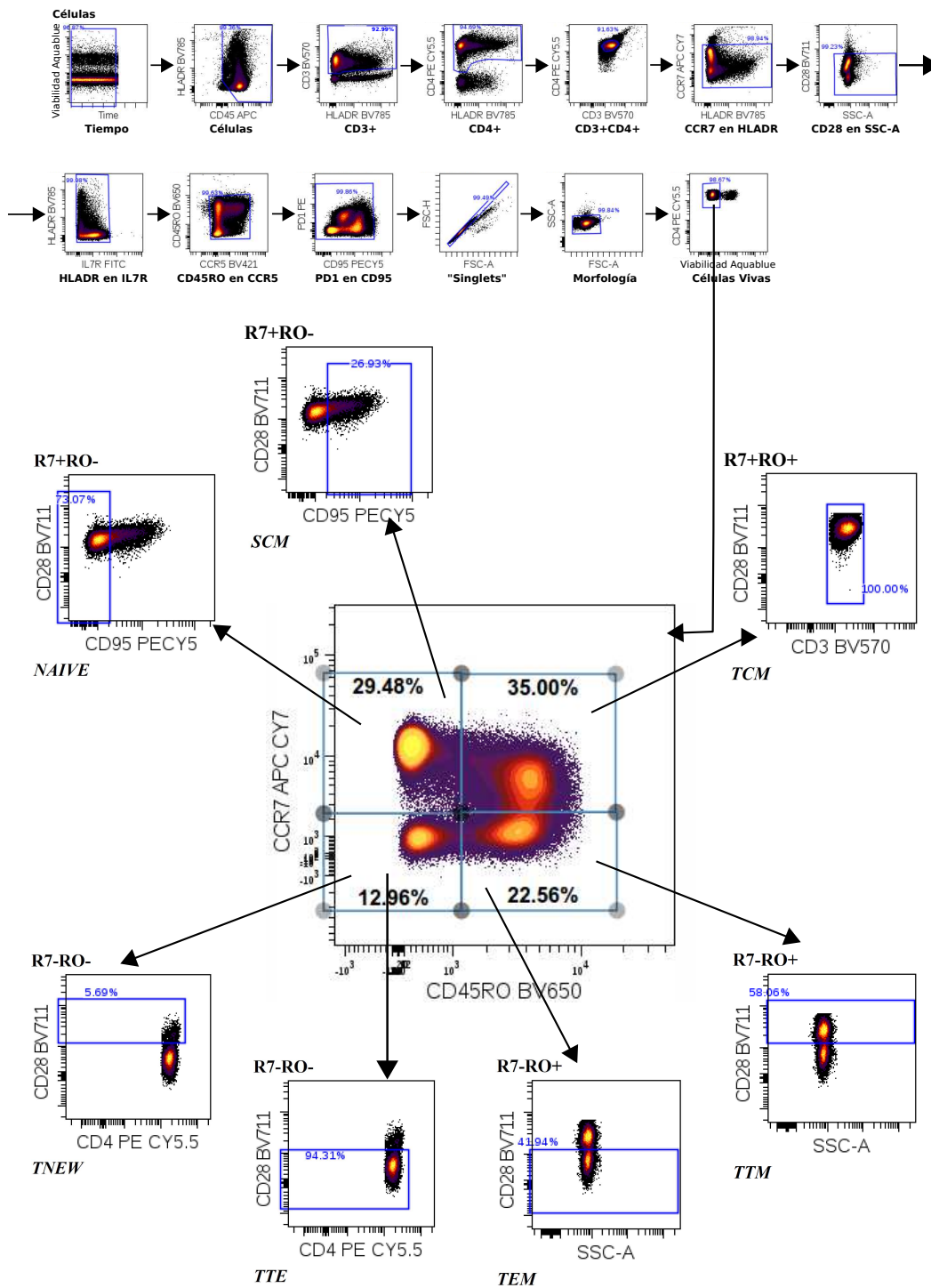


Figura 4.2: Estrategia de compuertas "gating" para definir a las subpoblaciones de células T CD4+. *Continúa siguiente página...*

Figura 4.2: ...de página anterior. La estrategia de análisis consistió en la realización de compuertas (gates) de tiempo vs viabilidad-Aqua blue (para analizar a los eventos que fueron capturados homogéneamente), CD45 vs HLA DR (para eliminar a los restos celulares), CD3 vs CD4 (para analizar únicamente a los linfocitos T CD4+), FSC-A vs FSC-H (para analizar solo a los singletes que son células individuales), FSC-A vs SSC-A (para analizar a las células que tienen morfología de linfocitos), Viabilidad-Aqua blue vs CD4 (para evaluar la viabilidad celular) y finalmente CD45 RO vs CCR7 (para acotar a las diferentes subpoblaciones de interés dentro de los cuatro cuadrantes). Las subpoblaciones fueron posteriormente identificadas de acuerdo a la expresión combinada de los siguientes marcadores: Células T CD4+ naïve (CD3+/ CD4+/ CD45RO- / CCR7+/CD28+/CD95-), Células T CD4+ de memoria (CD3+/CD4+/ CD28+ - /CD95+), Stem Cell Memory (T_{SCM} : CD3+/ CD4+/ CD45RO- / CCR7+/ CD28+/ CD95+), Memoria Central (T_{CM} : CD3+/ CD4+/ CD45RO+ / CCR7+/ CD28+), Memoria transicional (T_{TM} : CD3 +/ CD4+/ CD45RO+ / CCR7- / CD28+), Memoria Efectora (T_{EM} : CD3+/CD4+/CD45RO+ / CCR7- / CD28-), Terminalmente Efectora (T_{TE} : CD3+/CD4+/CD45RO- / CCR7- / CD28-) y T new (T_{NEW} : CD3+/CD4+/CD45RO- / CCR7- / CD28+).

Capítulo 5

Resultados.

”Sólo el azar está en el origen de toda innovación, de toda creación en la biosfera. El puro azar, la sola posibilidad, la libertad absoluta, absolutamente libre pero ciega, está en la raíz misma del prodigioso edificio de la evolución. Esta es hoy en día la única hipótesis concebible, la única que cuadra con los hechos observados y comprobados”.

*Jacques Monod, citado en el libro
"Risk and Meaning" de Nicolas
Bouleau.*

”El silencio es el ruido más fuerte,
quizás el más fuerte de los ruidos”

Miles Davis.

5.1. Perfil longitudinal de la metilación del LTR 5' presente en las células T CD4+ totales provenientes de personas bajo TAR a corto plazo.

Evaluamos los perfiles longitudinales de metilación basados en 9 sitios CpG canónicos dentro del LTR 5', presentes en la secuencia de referencia HXB2, en células T CD4+

totales de sangre provenientes de 12 pacientes bajo TAR a corto plazo (24-30 meses), utilizando secuenciación de siguiente generación, como se explica en las (**Secciones 4.5, 4.6, 4.7, 4.8,4.9**). (**Figura 4.1**). Para cada participante, se realizó un seguimiento de dos años en intervalos de tres meses (**Figura 5.1**). Para tener una medida de la metilación que fuese cuantitativa, estimamos un "Índice de Metilación CpG", promediando la contribución ponderada de la metilación de cada sitio CpG a la metilación global del LTR 5' de cada muestra. Además, para cada punto de tiempo, determinamos la carga proviral y la carga viral residual en plasma, y las distribuciones de las subpoblaciones de células T CD4+, definidas como se muestra en la (**Figura 4.2**). Los participantes fueron en su mayoría hombres (92%), con una mediana de edad de 39 años (RIC 33-47) y con una mediana en el nadir de CD4+ de 71 células/mm³ (RIC 7-251) (**Cuadro 5.1**). La mayoría de las personas bajo TAR a corto plazo (11/12) empezaron con un régimen de tenofovir+emtricitabina+efavirenz. 7/12 de los participantes se quedaron con su régimen original de TAR durante el período de seguimiento, 3/12 de los participantes cambiaron su backbone "esqueleto" a abacavir+lamivudina y 3/12 cambiaron el tercer fármaco de efavirenz a atazanavir/ritonavir o nevirapina debido a los efectos adversos del fármaco (no se observó resistencia al fármaco en ninguno de los participantes) (**Cuadro 4.1**). La mediana del número de lecturas de secuenciación (reads) obtenidas para evaluar a las variantes virales fue 16 270 (RIC 4 107-46 760). El número de variantes observadas en cada punto de tiempo varió de 1 a 13 y no se observó asociación entre el número de variantes y el tiempo de TAR en ninguno de los participantes ($p > 0.05$ en todos los casos). En general, la metilación del LTR 5' fue común y se observó en al menos un punto de tiempo en todos los participantes (**Figura 5.1**). Además, la metilación del LTR 5' mostró un comportamiento dinámico a lo largo del tiempo. En general, se observaron dos patrones globales de metilación CpG: 1) la metilación estuvo presente durante todo el período de seguimiento ("grupo de metilación constante"; e.g el participante TP31) o 2) la metilación aumentó significativamente en algún momento (principalmente alrededor del mes 42) a lo largo del período de estudio ("grupo de metilación incrementada"; e.g el participante TP25) (**Figura 5.2**). Esta separación en dos grupos considerando a la variable metilación fue significativa a los meses 30, 33, 39 y 54 del seguimiento (en los meses 30 y 54, valores $p < 0.01$ y a los meses 33 y 39, valores $p < 0.05$) (**Figura 5.4**). Las medianas del índice de metilación CpG, los RIC y los valores p, pueden encontrarse en el (**Cuadro A.1**).

Al examinar a las variantes individuales, se observaron diferencias en el grado de metilación entre los diferentes puntos de tiempo y éstas coincidieron frecuentemente con la alternancia en el predominio de variantes altamente metiladas vs. variantes no metiladas (**Figura 5.2**) y (**Figura 5.3**).

En general, las variantes que carecían de sitios metilados mostraron la pérdida de uno o más sitios CpG debido a mutación, en comparación con HXB2 (por e.g, (**Figura 5.2**), TP25 36M). Las variantes que mostraron 8 sitios CpG canónicos metilados (el sitio CpG2 generalmente se encontró mutado o no metilado) fueron comunes y observadas

en todos los participantes en al menos un punto de tiempo.

En general, el número de sitios CpG en el fragmento del LTR 5' estudiado varió de 1 a 10 entre los participantes del estudio y a diferentes puntos de tiempo. Observamos la presencia de mutaciones que conducen finalmente a la pérdida y a la ganancia de sitios CpG en variantes predominantes. Además, todos los participantes ganaron sitios CpG, CHG o CHH, potencialmente susceptibles a la metilación, sitios originalmente ausentes en la referencia HXB2, y esta adquisición estuvo presente en al menos un punto de tiempo (**Figura 5.1**).

Tomados en conjunto, nuestros resultados muestran que la metilación del LTR 5' del VIH es variable a lo largo del tiempo en células T CD4+ totales de la sangre provenientes de individuos bajo TAR a corto plazo y que la metilación puede ser observada desde las primeras etapas de la iniciación del TAR. Además, al menos dos patrones temporales de metilación diferentes fueron observados entre los participantes.

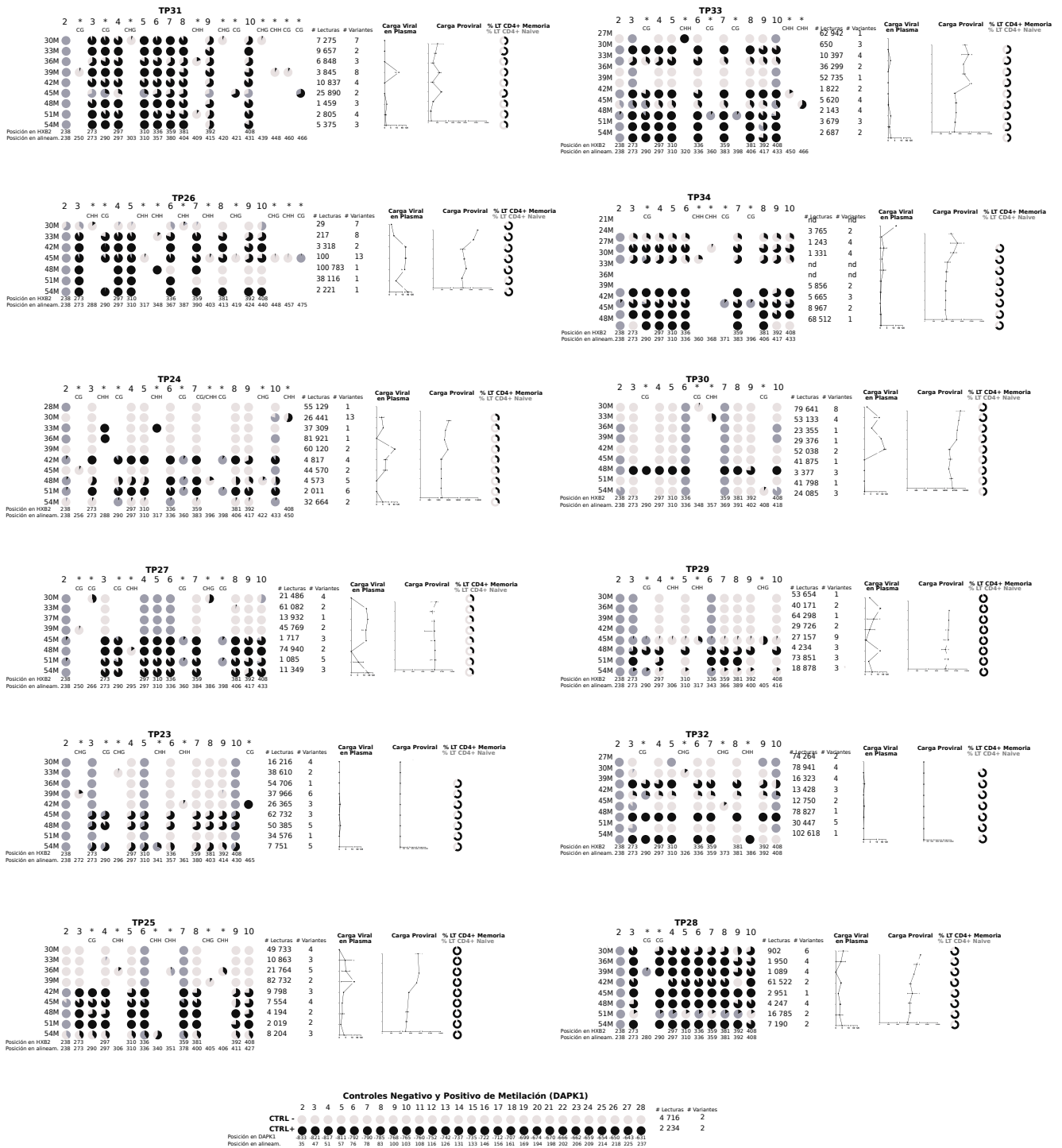


Figura 5.1: Perfiles longitudinales de la metilación del LTR 5' del VIH en células T CD4+ totales provenientes de doce participantes bajo TAR a corto plazo. *Continúa siguiente página...*

Figura 5.1: *...de página anterior.* Cada grupo representa a un solo participante bajo TAR a corto plazo (24 a 30 meses), seguido durante dos años con intervalos de 3 meses. Cada línea representa el resumen de todas las variantes observadas en ese punto de tiempo. Los sitios CpG son mostrados como gráficos de pastel que indican la proporción de las variantes que se encontraron metiladas (negro), no metiladas (gris claro) o mutadas (gris oscuro) dentro de la muestra. Las posiciones de los sitios CpG canónicos fueron numeradas de acuerdo a la referencia HXB2. También se incluyeron sitios susceptibles a la metilación como CpG adicionales, CHG y CHH, no observados en la secuencia de HXB2 (marcados con *). El número de lecturas y el número total de variantes que fueron obtenidas por punto de tiempo, después del filtrado y del alineamiento también son mostradas. Finalmente, las mediciones de la carga viral residual en plasma, la carga proviral y la proporción de células T CD4+ de memoria (CD3+/CD4+/CD28+/-/CD95+; negro) y naive (CD3+/CD4+/CD45RO-/ CCR7+/CD28+/CD95-; gris claro) en cada punto de tiempo son mostradas.

Características basales de los grupos de metilación constante y metilación incrementada bajo TAR a corto plazo comparados con el grupo de TAR a largo plazo.

	Grupo bajo TAR a corto plazo combinado n=12	Valor-p ^a	Grupo bajo TAR a corto plazo metilación constante n=5	Valor-p ^b	Grupo bajo TAR a corto plazo metilación incrementada n=7	Valor-p ^c	Valor-p ^d	Grupo bajo TAR a largo plazo n=10
Género (Masculino/Femenino)	11:1		5:0		6:1			9:1
Edad (años)	39 (33-47)	0.013	41 (38-45)	0.910	34 (31-48)	0.004	<0.001	45 (39-56)
Nadir de células T CD4+ (células/mm ³)	71 (7-251)	0.294	128 (12-264)	0.843	13 (4-212)	0.106	0.075	96 (47-234)
pVL Pre-TAR (copias ARN/mL)	194077 (67293-285046)	0.077	144952 (90736-243202)	0.174	248219 (59479-285552)	0.048	0.205	47800 (7581-154809)
Delta de cuentas de células T CD4+ a los 48 meses (células/mm ³)	388 (174-502)	<0.001	330 (130-412)	0.843	455 (214-580)	0.029	<0.001	291 (193-341)
Delta % CD4 a los 48 meses	13 (11-20)	0.163	12 (10-15)	0.190	13 (12-32)	0.023	0.041	11 (6-14.5)
Delta de cuentas de células T CD8+ a los 48 meses (células/mm ³)	172 (-104-315)	0.053	102 (-109-246)	0.751	202 (-89-325)	0.105	0.023	-85 (-405-331)
Delta % CD8 a los 48 meses	-15 (-24_-5)	0.268	-13 (-17-3)	0.015	-24 (-25_-10)	0.162	<0.001	-32 (-46_-9)
Delta CD4/CD8 a los 48 meses	0.40 (0.24-0.80)	0.058	0.33 (0.23-0.35)	0.668	0.53 (0.28-0.84)	0.036	<0.001	0.35 (0.25-0.55)
Días para conseguir una pVL indetectable	183 (101-368)	0.287	186 (99-384)	0.551	180 (107-344)	0.771	0.205	148 (95-379)
Tercer fármaco en el régimen de TAR actual (NNRTIs / PIs)	83% / 17%	0.624	80% / 20%	1	86% / 14%	0.603	1	70% / 30%
Número de cambios de regimen de TAR	0 (0-1)	0.027	0 (0-0)	0.003	1 (0-1)	0.102	0.008	1 (0.8-2)

Cuadro 5.1: Para cada variable se muestra la mediana (rango intercuartil). ^a Grupo bajo TAR a corto plazo vs Grupo bajo TAR a largo plazo; ^b Grupo de metilación constante vs. Grupo bajo TAR a largo plazo; ^c Grupo de metilación incrementada vs Grupo bajo TAR a largo plazo; ^d Grupo de metilación constante vs Grupo de metilación incrementada. TAR=terapia antirretroviral; pVL=Carga viral residual en plasma; NNRTI=Inhibidores de la reverso transcriptasa no análogos de nucleósidos; PI=Inhibidores de proteasa. Los valores p significativos (p<0.05) se muestran en **negritas**.

5.2. Patrones temporales en la metilación del LTR 5'

Considerando a todos los participantes con TAR a corto plazo, observamos un aumento significativo en la metilación del LTR 5' al mes 42 con respecto al mes 30 (mediana del Índice de Metilación CpG: 74.7 % vs. 0 %, $p=0.025$) (**Cuadro 5.2**), (**Figura 5.4A**) posteriormente observamos un incremento aún más significativo en la metilación del LTR 5' al mes 48 con respecto al mes 30 (mediana del Índice de Metilación CpG: 81.8 % vs. 0 %, $p=0.003$) (**Cuadro 5.2**), (**Figura 5.4A**) y finalmente observamos una disminución en la metilación del LTR 5' al mes 54 con respecto al mes 48 (mediana del Índice de Metilación CpG: 55.6 % vs. 81.8 %, aunque el valor p no es significativo $p=0.083$). Continuamos observando el aumento significativo en la metilación del LTR 5' al mes 54 con respecto al mes 30 (mediana del Índice de Metilación CpG: 55.6 % vs. 0 %, $p=0.006$) (**Cuadro 5.2**), (**Figura 5.4A**). Para revisar los detalles de las pruebas estadísticas empleadas en este apartado, ver (**Sección 4.11**).

Este incremento en el grado de metilación fue evidente en un subconjunto de los participantes (grupo de metilación incrementada) que a continuación se detalla. Observamos un aumento significativo en la metilación del LTR 5' (grupo de metilación incrementada) al mes 48 con respecto al mes 30 (mediana del Índice de Metilación CpG: 82.9 % vs. 0 %, $p=0.022$) y este incremento continuo siendo significativo al mes 54 con respecto al mes 30 (mediana del Índice de Metilación CpG: 36.4 % vs. 0 %, $p=0.036$) (**Cuadro 5.3**), (**Figura 5.4B**), mientras que el resto de los participantes mantuvo una metilación bastante alta y constante a lo largo del período de seguimiento (grupo de metilación constante) (**Figuras 5.1, 5.2, 5.4B**).

Específicamente, el grupo de metilación incrementada fue significativamente diferente del grupo de metilación constante respecto al nivel inicial de la metilación CpG dentro del LTR 5' (a los 30 meses de TAR), es decir el primer grupo fue significativamente menor en comparación con este último (mediana del índice de metilación CpG: 0 vs 65, $p=0.009$).

Considerando a todos los participantes con TAR a corto plazo, observamos un incremento significativo en la metilación del LTR 5' (específicamente en los sitios CpG adicionales) al mes 48 con respecto al mes 30 (mediana del índice de metilación CpG adicional: 35 % vs. 1 %, $p=0.004$), manteniéndose esta tendencia al comparar el mes 54 con respecto al mes 30 (mediana del índice de metilación CpG adicional: 31 % vs. 1 %, $p=0.006$) (**Cuadro 5.2**).

También notamos que los grupos metilación incrementada y metilación constante, son diferentes al mes 30 del seguimiento respecto al índice de metilación CpG adicional (mediana del índice de metilación CpG adicional: 0 vs 21.7 respectivamente $p=0.030$) (**Cuadro 5.3**).

El incremento en el índice de metilación CpG adicional a lo largo del tiempo fue evidente en un subconjunto de los participantes (grupo de metilación incrementada), comparando al mes 48 con respecto al mes 30 (mediana del Índice de Metilación

CpG adicional: 45 % vs. 0 %, $p=0.015$) y este incremento continuo siendo significativo al comparar al mes 54 con respecto al mes 30 (mediana del Índice de Metilación CpG adicional: 17.6 % vs. 0 %, $p=0.015$) (**Cuadro 5.3**).

Posteriormente nosotros observamos una disminución general en la carga proviral con respecto al tiempo de TAR. p. eg. al mes 42 en comparación con el mes 30 (mediana de la carga proviral: 3 638 vs. 14 146 copias de ADN/ 10^6 células T CD4+, $p=0.005$) (**Cuadro 5.2**), (**Figura 5.4C**), esta disminución general en la carga proviral también fue observada en el mes 48 en comparación con el mes 30 (mediana de la carga proviral: 4 973 vs. 14 146 copias de ADN/ 10^6 células T CD4+, $p=0.00097$), y esta tendencia continua al mes 54 en comparación con el mes 30 (mediana de la carga proviral: 3 915 vs. 14 146 copias de ADN/ 10^6 células T CD4+, $p=0.002$) (**Cuadro 5.2**), (**Figura 5.4C**). En su conjunto estos resultados indican que la carga proviral disminuyó con el tiempo de TAR (excepto en la comparación del mes 39 vs mes 30) (**Figura 5.4C**).

Posteriormente para ser más específicos en el análisis de la variación en la carga proviral, dividimos a los pacientes bajo TAR a corto plazo en dos grupos (metilación constante y metilación incrementada) por cada punto de tiempo durante el seguimiento y lo comparamos con la carga proviral al mes 30 (**Cuadro 5.3**), (**Figura 5.4D**). Encontramos una disminución general en la carga proviral (únicamente en el grupo de metilación incrementada) al mes 48 en comparación con el mes 30 (mediana de la carga proviral: 5 174 vs. 11 250 copias de ADN/ 10^6 células T CD4+, $p=0.031$) y manteniéndose esta disminución en la carga proviral también al mes 54 en comparación con el mes 30 (mediana de la carga proviral: 3 915 vs. 11 250 copias de ADN/ 10^6 células T CD4+, $p=0.031$) (**Cuadro 5.3**), (**Figura 5.4D**).

Respecto a la carga viral residual en plasma (pVL), esta se mantuvo en general estable en la mayoría de los participantes durante el seguimiento, aunque con ocasionales blips: TP30 39M (43 copias/mL), TP30 42M (99 copias/mL), TP24 39M (71 copias/mL), TP27 54M (42 copias/mL) y TP26 54M (60 copias/mL) (**Figuras 5.1, 5.4E**).

Analizando la variable de pVL y considerando la división en los grupos de metilación constante y metilación incrementada por cada punto de tiempo, encontramos un incremento en la pVL (grupo de metilación incrementada) al mes 39 en comparación con el mes 30 (mediana de la pVL residual: 10.5 vs. 1 copia ARN/mL $p=0.031$), esta diferencia podría ser explicada debido al ligero incremento de la carga proviral (grupo de metilación incrementada) al mes 39 en comparación con el mes 30 (mediana de la carga proviral: 13 084 vs. 11 250 copias de ADN/ 10^6 células T CD4+, $p=0.031$). En general no se observaron tendencias significativas de variación en el tiempo para la carga viral residual en plasma. Sin embargo, el grupo de metilación constante mostró en general menos dispersión que el grupo de metilación incrementada (excepto en el mes 54 cuando el participante TP26 tuvo un blip) (**Figura 5.4E**).

Considerando los porcentajes de las subpoblaciones de células T CD4+ (Naive vs Memoria), se observaron tres patrones generales entre los participantes: 1) individuos

con mayor proporción de células T CD4+ naive (17 %): TP24 y TP27, 2) individuos con mayor proporción de células T CD4+ de memoria (58 %) : TP26, TP23, TP25, TP34, TP29, TP32 y TP28, y 3) individuos con igual proporción de células T CD4 + naive y de memoria (25 %): TP31, TP33 y TP30; sin embargo, estos grupos no mostraron asociaciones específicas con los patrones de metilación (**Figura 5.1**).

Respecto a las variaciones temporales dentro de las subpoblaciones de células T CD4+ en el grupo de TAR a corto plazo, no se encontraron diferencias significativas en el porcentaje tanto de células T CD4+ naive y de memoria al comparar 30M vs 42M ($p=0.114$ y $p=0.114$ respectivamente) y continuó la misma tendencia al comparar 30M vs 48M ($p=0.465$ y $p=0.465$ respectivamente). Sin embargo, esta tendencia cambio al comparar 30M vs 54M mostrando diferencias significativas ($p=0.002$ y $p=0.002$ respectivamente). Incrementándose las células naive y disminuyendo las células de memoria. Este cambio es importante, debido a que trae consigo un rearrreglo en las subpoblaciones de células T CD4+ de memoria. Específicamente, las células T_{CM} y T_{TM} (disminuyeron $p=0.046$ y $p=0.002$ respectivamente), en cambio las células T_{SCM} , T_{NEW} y T_{TE} (incrementaron $p=0.004$, $p=0.004$ y $p=0.004$ respectivamente).

Es de destacar que las células T_{NEW} se incrementaron en el tiempo al realizar las comparaciones 30M vs 42M ($p=0.005$), 30M vs 48M ($p=0.005$) y 30M vs 54M($p=0.004$). (**Cuadro 5.2**)

Posteriormente para revisar las variaciones en el porcentaje de las subpoblaciones de células T CD4+ de memoria, ahora ya considerando la división entre los grupos de metilación constante y metilación incrementada por punto de tiempo, encontramos lo siguiente: Al mes 48, se observó un mayor porcentaje de células T_{SCM} en el grupo de metilación constante vs grupo de metilación incrementada ($p=0.017$) (**Cuadro 5.3**).

Mas adelante al realizar una comparación en el grupo de metilación incrementada a los 54M vs 30M respecto al porcentaje de células naive y memoria encontramos lo siguiente: el porcentaje de células T CD4+ Naive (mediana del porcentaje de células T CD4+ Naive: 40 % vs. 28 %, $p=0.031$) y el porcentaje de células T CD4+ Memoria (mediana del porcentaje de células T CD4+ Memoria: 60 % vs. 72 %, $p=0.031$). Estos datos muestran que el grupo de metilación incrementada, gana significativamente células naive en el tiempo, y por tanto logro una disminución en el porcentaje de células de memoria, y la variación en el porcentaje de las células T CD4+ de memoria generó una reestructuración en las subpoblaciones(**Cuadro 5.3**). Para conocer específicamente que subpoblaciones de las células de memoria tuvieron cambios significativos en el grupo de metilación incrementada comparando los tiempos de seguimiento 54M vs 30M, encontramos que el porcentaje de células T_{NEW} y T_{TE} se incrementaron, mientras que el porcentaje de células T_{TM} disminuyeron (en los tres casos una $p<0.05$). También es de destacar que realizamos la misma comparación temporal con el grupo de metilación constante. Aquí no se observaron cambios significativos en el porcentaje de células de naive y de memoria, ni tampoco en las diferentes subpoblaciones de células de memoria ($p>0.05$)(**Cuadro 5.3**).

Adicionalmente, se incluyó a un grupo de participantes que tenían una larga historia de TAR, con el objetivo de explorar las asociaciones entre la metilación con la carga viral residual y la carga proviral. En este grupo de corte transversal se incluyó a personas con una mediana de tiempo en TAR de 12 años (RIC 7-15), en su mayoría hombres (90%), con una mediana de edad de 45 años (RIC 39-56) y una mediana del nadir de células T CD4+ de 96 células/mm³ (RIC 47-234) (**Cuadro 5.1**). Los participantes bajo TAR a largo plazo mostraron significativamente más cambios en el régimen de TAR que las personas bajo TAR a corto plazo (8/10 de los participantes cambiaron de régimen de TAR al menos una vez), principalmente por motivos de optimización o cambio a nuevos medicamentos. El régimen de TAR actual de los participantes en este grupo fue tenofovir+emtricitabina+efavirenz para 7/10 y tenofovir+emtricitabina+ un inhibidor de proteasa para 3/10 (**Cuadro 4.1**). Las personas bajo TAR supresor a largo plazo mostraron en general un alto grado de metilación CpG con una mediana del índice de metilación CpG de 78% (RIC 71-87%) (**Cuadro 5.2**). Las variantes con sitios CpG mutados fueron raras en este grupo y la presencia de nuevos sitios susceptibles a la metilación fueron también poco frecuentes, en comparación con los individuos que fueron seguidos longitudinalmente (**Figura 5.5**).

Comparando a los participantes del grupo de seguimiento longitudinal a los 30 meses bajo TAR con el grupo transversal de TAR a largo plazo, observamos una diferencia significativa en la metilación global (mediana del índice de metilación CpG: 0% vs 78%, $p=0.005$), pero esta diferencia fue perdida al mes 42 de TAR (mediana del índice de metilación CpG: 75% vs. 78%, $p=0.273$) (**Cuadro 5.2**). Específicamente, la diferencia en la metilación temprana (30 meses vs TAR a largo plazo) fue observada en las personas del grupo de metilación incrementada ($p=0.001$), pero no en el grupo de metilación constante ($p=0.244$) (**Cuadro 5.3**). No se observaron diferencias en la pVL residual o en la carga proviral, comparando a los individuos con TAR a corto plazo (tanto a 30 como a 42, 48 y 54 meses de seguimiento) y los de TAR a largo plazo (**Cuadros 5.2, 5.3**).

Para ser más exhaustivos en el análisis de metilación, estudiamos a los patrones de metilación no canónicos, estableciendo una medida de metilación en sitios CpG diferentes a los 9 canónicos, presentes en la referencia HXB2 (índice de metilación CpG adicional) y también estudiamos los patrones de metilación en sitios distintos de CpG (índice de metilación no-CpG). De manera similar a nuestras observaciones para la metilación CpG canónica, observamos una diferencia significativa en la metilación CpG adicional en personas bajo 30M de TAR dentro del grupo de metilación incrementada ($p=0.003$), pero no en el grupo de metilación constante ($p=0.282$) cuando los comparamos con personas bajo TAR a largo plazo. A los 48M de TAR dentro del grupo de metilación incrementada también encontramos diferencias significativas ($p=0.017$) vs grupo bajo TAR a largo plazo, pero no encontramos diferencias al comparar el grupo de metilación constante a los 48M vs grupo bajo TAR a largo plazo ($p=0.121$). Sin embargo, la comparación a los 54M de TAR dentro del grupo de metilación incrementada

vs grupo bajo TAR a largo plazo ya no obtenemos diferencias significativas ($p=0.803$), pero si obtenemos diferencias significativas al comparar el grupo de metilación constante a los 54M ($p=0.012$) vs grupo bajo TAR a largo plazo (**Cuadro 5.3**).

No se observaron diferencias en la metilación no-CpG al comparar los grupos de TAR a corto plazo (tanto en 30, 42, 48 y 54 meses de seguimiento) y el grupo de TAR a largo plazo (**Cuadros 5.2, 5.3**).

Aunque no se observaron diferencias significativas en el porcentaje de células T CD4+ naive y de memoria entre individuos bajo TAR a corto y a largo plazo (**Cuadro 5.2**), pero al considerar a las diferentes subpoblaciones de células T CD4+ de memoria, observamos una prevalencia más baja de células T de memoria stem cell (T_{SCM}) en el grupo de TAR a corto plazo 30M ($p=0.049$), 42M ($p=0.035$) y 48M ($p=0.042$) comparado con el grupo de TAR a largo plazo (**Cuadro 5.2**). Esta menor prevalencia en T_{SCM} también se mantiene al comparar el grupo de metilación incrementada 48M vs grupo de TAR a largo plazo $p=0.009$ (**Cuadro 5.3**).

Respecto a las células T de memoria nuevas (T_{NEW}), también las encontramos en baja prevalencia en los individuos con 30 meses de TAR al compararlas con el grupo de TAR a largo plazo ($p=0.003$) (**Cuadro 5.2**). Más específicamente el grupo de metilación incrementada a los 30 y 48 meses mostró menor prevalencia de células T_{NEW} al compararlo con el grupo de TAR a largo plazo ($p=0.002$ y $p=0.018$ respectivamente), pero no obtuvimos diferencias si la comparación se realiza con el grupo de metilación constante ($p=0.075$ y 0.902 , respectivamente) (**Cuadro 5.3**), lo que sugiere diferencias tempranas en la composición de las subpoblaciones de células T CD4+ entre personas pertenecientes a ambos grupos.

5.3. Características basales y de seguimiento en personas con distintos patrones de metilación del LTR 5'.

A continuación, buscamos diferencias dentro de las características basales de las personas que pertenecían a los grupos de metilación constante y de metilación incrementada (**Cuadro 5.1**), es de destacar que no encontramos diferencias en el tercer fármaco en el régimen de TAR actual (NNRTIs/PIs) entre ambos grupos. Las personas dentro del grupo de metilación incrementada fueron significativamente más jóvenes (mediana de edad: 34 vs. 41 años, $p<0.001$), tuvieron una mayor ganancia en el recuento de CD4+ a los 48 meses después del inicio del TAR (delta de cuentas de células T CD4+: 455 vs 330, $p<0.001$; delta % CD4: 13 vs. 12, $p=0.041$), mayor disminución en las cuentas de CD8+ a los 48 meses (delta de cuentas de células T CD8+: 202 vs. 102, $p=0.023$; delta % CD8: -24 vs. -13, $p<0.001$), un mayor incremento en la relación CD4/CD8 a los 48 meses (delta CD4/CD8: 0.53 vs. 0.33, $p<0.001$), y un mayor número

de cambios de régimen de TAR (mediana del número de cambios de régimen de TAR: 1 vs. 0, $p=0.008$) (**Cuadro 5.1**). Estos resultados sugieren que los diferentes comportamientos en la metilación de LTR 5' podrían estar asociados con las características clínicas basales de las personas que comienzan el TAR (principalmente edad y probablemente una mayor fuente de células T naive capaces de reconstituir al sistema inmune después del TAR). Finalmente, aplicamos el Análisis Factorial Exploratorio utilizando Componentes Principales (AFECP) con rotación varimax con el objetivo de explicar la estructura de los datos. Se incluyeron a las variables clínicas basales y de seguimiento (**Cuadros A.3, A.4**), (**Figura 5.6**), ver (**Sección 4.11**). Considerando a la varianza común, los tres primeros factores explicaron el 50.86 % de la varianza común total. **El Factor 1** explicó el 21.39 % de la varianza común total y abarcó 10 variables (edad, ganancia de células T CD4+ (numérico y porcentual) a los 48M, disminución en las células T CD8+ (numérico y porcentual) a los 48M, cambio en el cociente CD4/CD8 a los 48M, porcentaje de células T CD4+ (Memoria, Naive, T_{CM} y T_{TM}). **El Factor 2**, explicó el 16.11 % de la varianza común total y abarcó 5 variables (tiempo transcurrido para lograr una carga viral en plasma indetectable, tiempo bajo TAR: 30 vs 54 meses en el seguimiento, tercer fármaco en el régimen de TAR: un inhibidor de la reverso transcriptasa del tipo no nucleósido (NNRTI) o un inhibidor de proteasa (PI). **El Factor 3**, explicó el 13.36 de la varianza común total y abarcó 4 variables (nadir del recuento de células T CD4+, porcentaje en las subpoblaciones de CD4+ (T_{EM} y T_{TE}) y el número de cambios en el régimen de TAR. Es de destacar que el índice de metilación CpG, la carga viral pre-TAR, el tiempo entre el diagnóstico de VIH y el tratamiento y el porcentaje de (T_{NEW}) fueron incluidos hasta el **Factor 4** (que explicó el 9.37 % de la varianza común total); el backbone “esqueleto” del régimen de TAR y el % de T_{SCM} se encontraron hasta el **Factor 5** (que explicó el 9.19 % de la varianza común total); y la carga viral residual, la carga proviral y la presencia de blips fueron encontrados hasta el **Factor 6** (que explicó el 7.36 % de la varianza común total (**Cuadro A.4**).

Observamos que 4/29 variables que fueron ingresadas al análisis factorial, mostraron un valor de uniqueness (unicidad o varianza única) >0.5 . Las variables fueron: % T_{SCM} (0.6588), % T_{NEW} (0.5401), carga proviral (0.5037) e índice de metilación CpG (0.6655). La unicidad significa la proporción de la varianza de la variable, en el cual el factor en donde fue asignada la variable, no fue capaz de capturarla y por lo tanto esas variables no están bien explicadas por su factor.

Considerando a los 3 primeros Factores (varianza común) y la división entre los grupos de metilación constante y metilación incrementada, los grupos mostraron baja evidencia de separación significativa a lo largo del tiempo (prueba de Wilcoxon para el Factor 2: $p=0.04$; para los Factores 1 y 3: $p>0.05$) (**Figura 5.6**). En el caso del Factor 1, los grupos se encuentran en general bien separados, pero es notorio un solapamiento entre los grupos al mes 42 del seguimiento y en el caso del Factor 3, en general los grupos también se encuentran bien separados, aunque es notorio un solapamiento entre los grupos al mes 54. Es importante señalar que considerando únicamente a la variable Índice

de metilación CpG, los grupos de metilación constante y metilación incrementada se separan muy bien en los primeros tiempos del seguimiento (30, 33 y 39 meses) y al final del seguimiento (54 meses), pero existe una zona comprendida entre 42-51 meses donde los grupos se comportan de manera similar, no encontrándose diferencias significativas (**Figura 5.4A y B**)(**Cuadro A.1**), (**Figura 5.6**).

En conjunto, estas observaciones muestran que las diferencias temporales en los patrones de Metilación en el LTR 5' se asocian con la combinación de características clínicas basales y de seguimiento. Dentro de las características basales destacan: Edad (univariado y Factor 1), ganancia de células T CD4+ a los 48M de TAR (univariado y Factor 1), disminución de células T CD8+ a los 48M de TAR (univariado y Factor 1), el número de cambios en el régimen de TAR (univariado y Factor 3), el tiempo transcurrido para lograr una carga viral indetectable (Factor 2), la clase de tercer fármaco en el régimen de TAR (Factor 2) y el nadir del recuento de células T CD4+ (Factor 3).

Dentro de las características del seguimiento destacan: disminución significativa de la carga proviral durante el seguimiento (univariado y únicamente en el grupo de metilación incrementada), tiempo bajo TAR (Factor 2), la composición de las subpoblaciones de células T CD4 (univariado, destacable en el grupo de metilación incrementada al final del seguimiento vs mes 30), porcentaje de células T CD4+ de memoria, naive, T_{CM} y T_{TM} (Factor 1) y el porcentaje de células T CD4+ T_{EM} y T_{TE} (Factor 3).

Finalmente, las variables $\%T_{SCM}$, $\%T_{NEW}$, carga proviral e índice de metilación CpG tuvieron una varianza única alta (análisis multivariado).

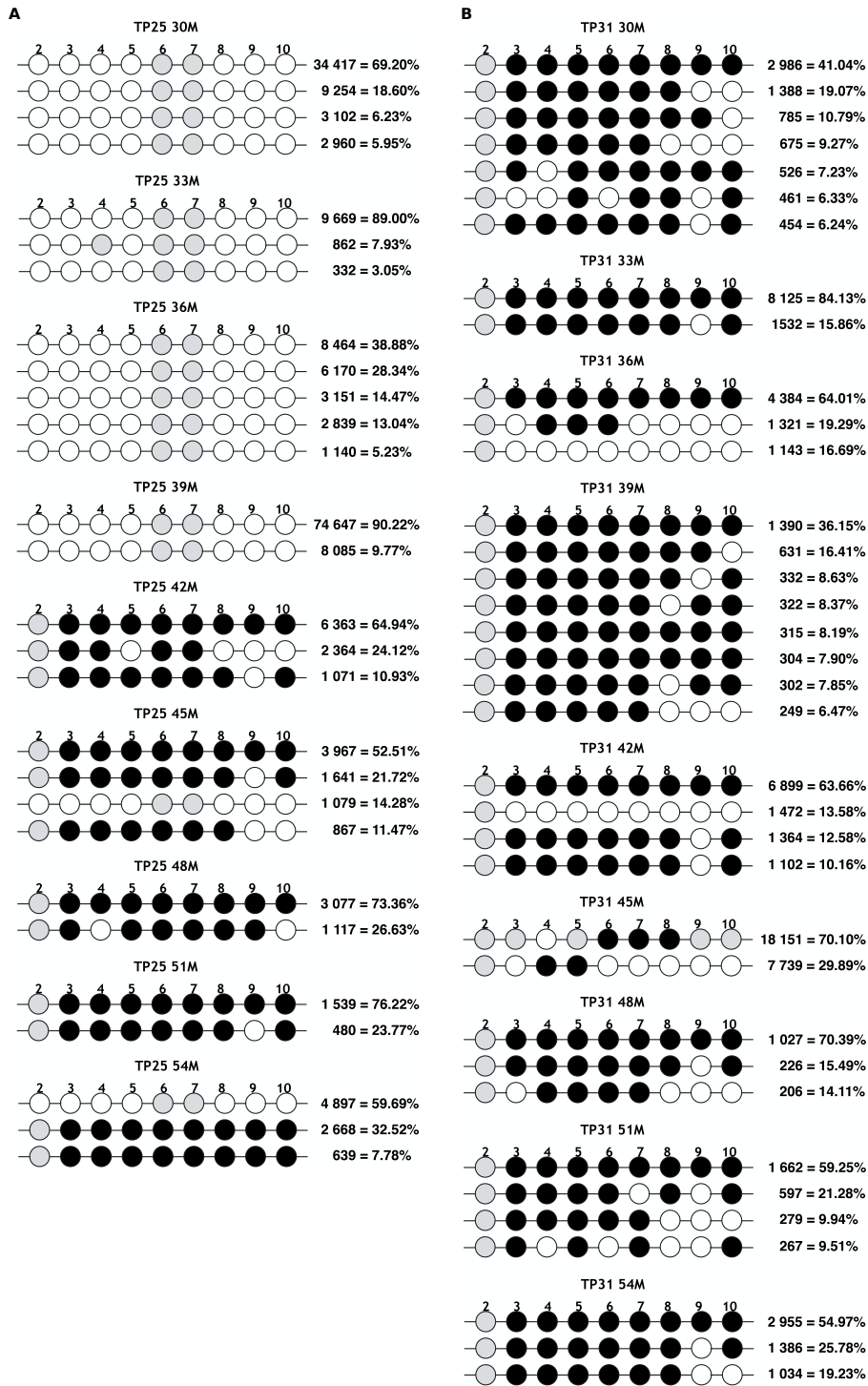
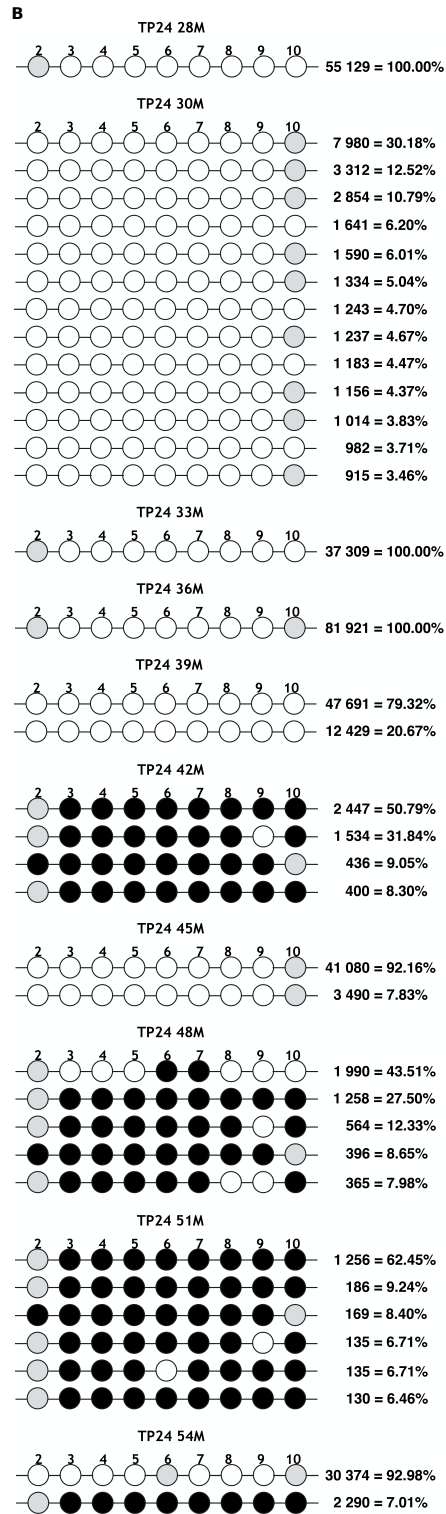
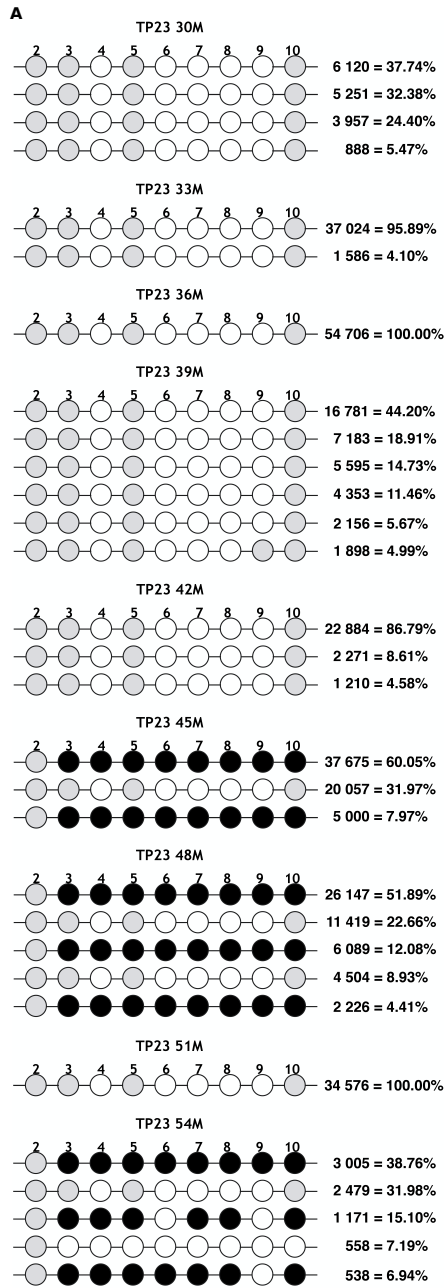
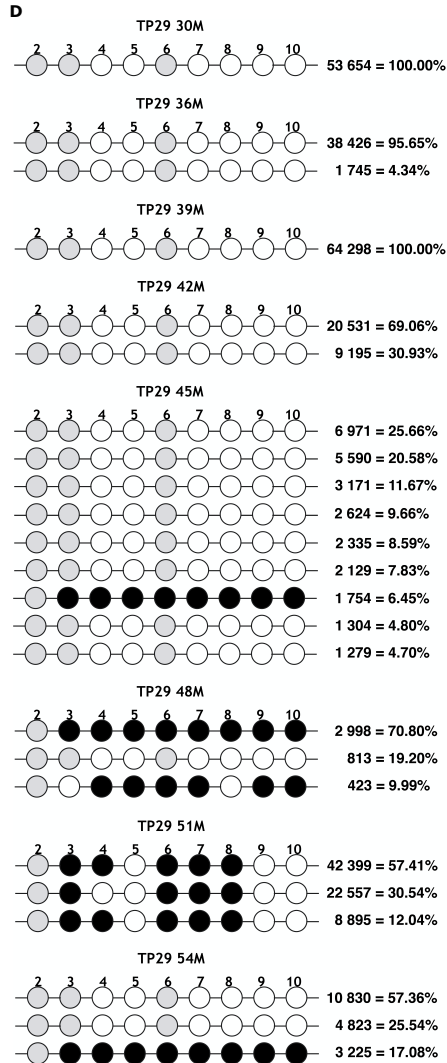
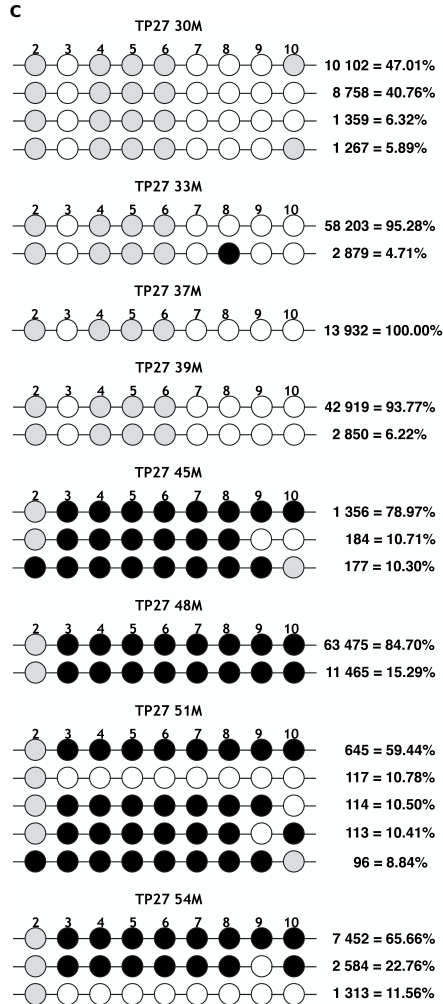
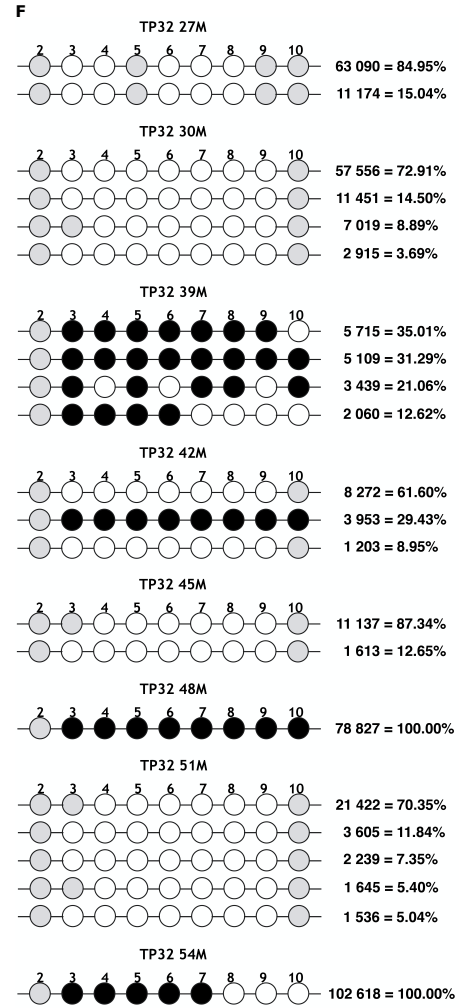
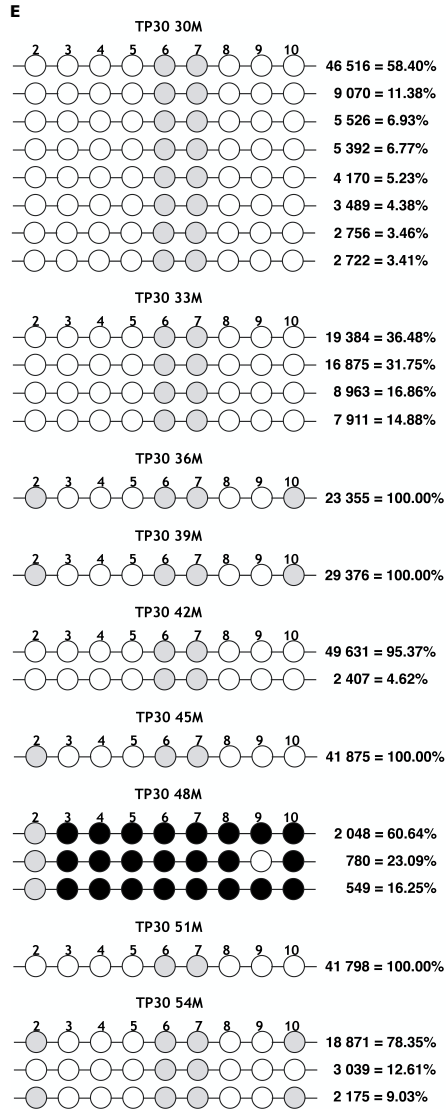


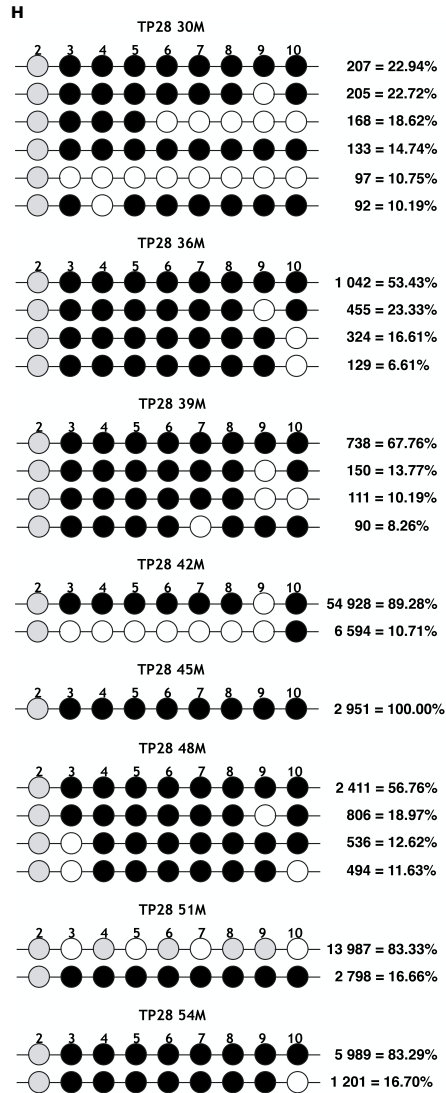
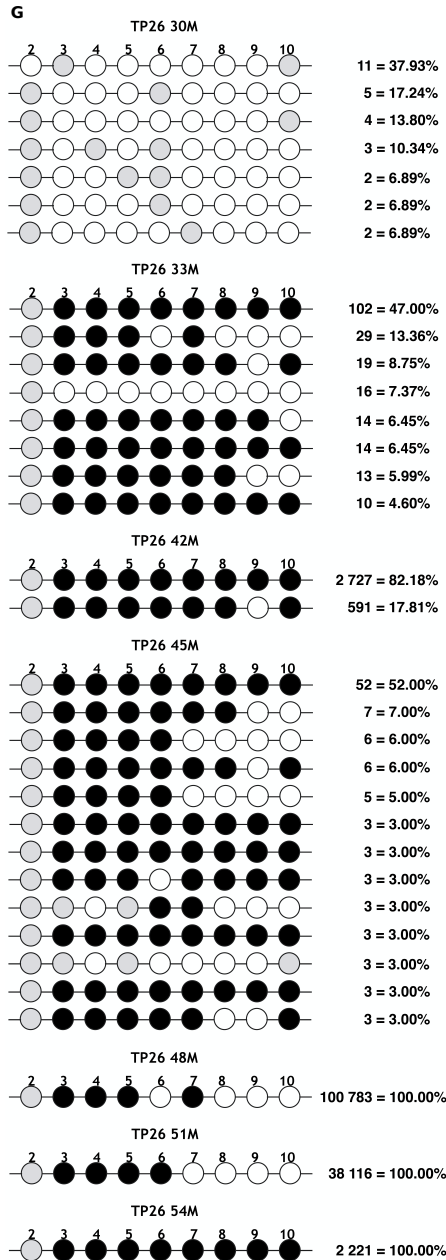
Figura 5.2: Tendencias de la metilación del LTR 5' del VIH en el tiempo considerando a dos participantes modelo. *Continúa siguiente página...*

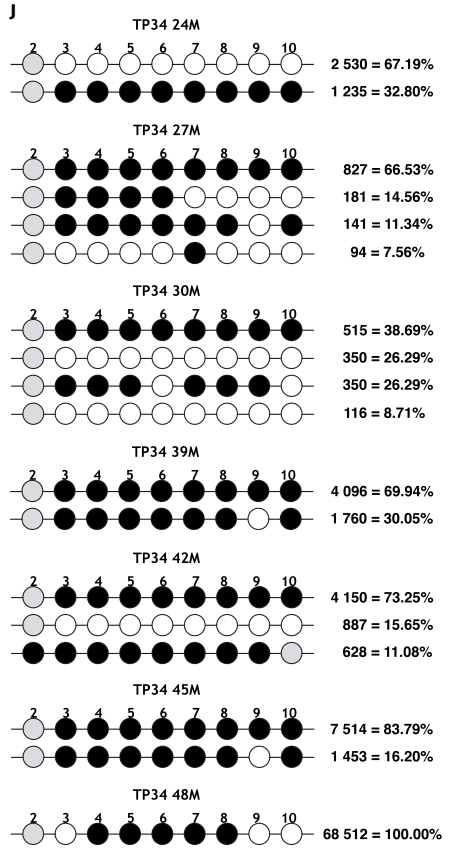
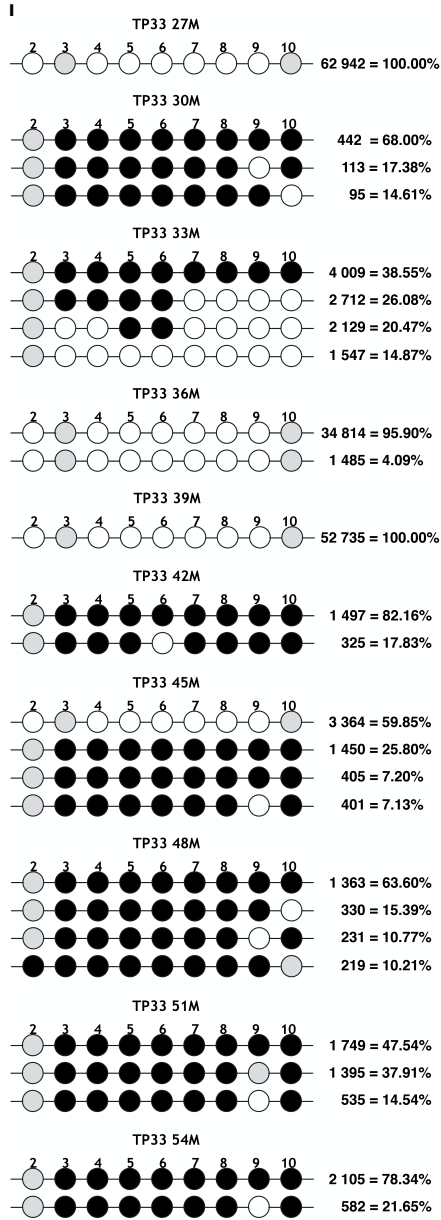
Figura 5.2: ...de página anterior. Se muestra la composición de variantes a detalle en cada punto de tiempo (30, 33, 36, 39, 42, 45, 48, 51 y 54 meses) para los participantes **A**) TP25, representativo del grupo de metilación incrementada, y **B**) TP31, representativo de grupo de metilación constante. Cada línea horizontal representa una variante. El número de lecturas que representan a esa variante y su proporción en la muestra se muestran a la derecha. Cada círculo representa a un sitio CpG canónico, numerado de acuerdo con la referencia HXB2: negro —metilado, blanco —no metilado, gris—mutado.











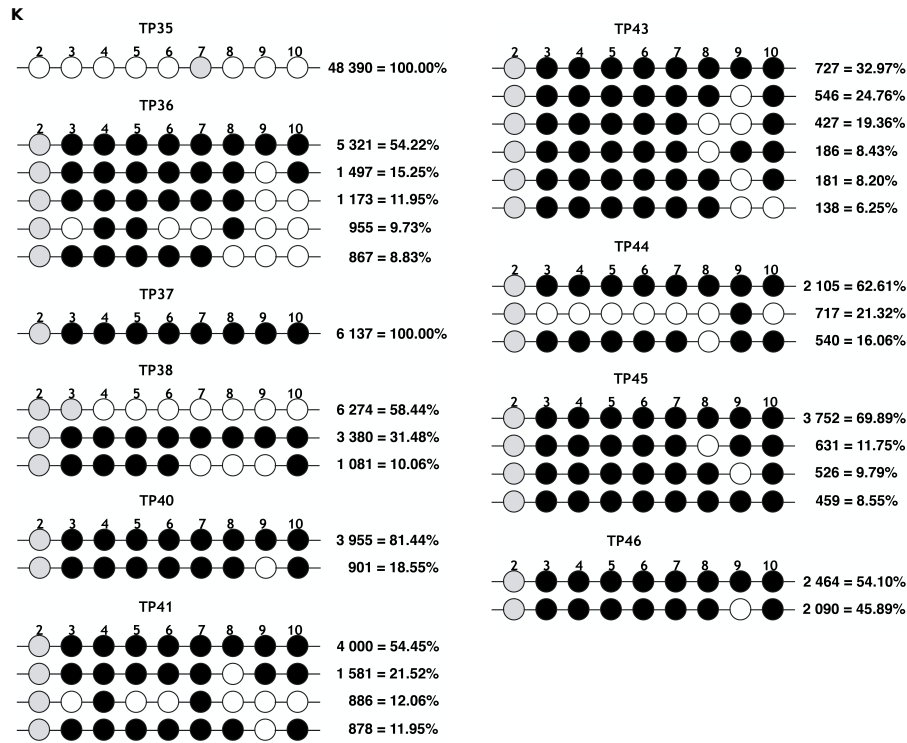


Figura 5.3: Composición detallada de las variantes del LTR 5' del VIH y análisis de metilación del LTR 5' en el resto de los participantes. Se muestra la composición detallada de las variantes por punto de tiempo para diez participantes bajo TAR a corto plazo no incluidos en la **Fig. 5.2** (TP23, TP24, TP27, TP29, TP30, TP32, TP26, TP28, TP33, TP34; **A-J**), así como puntos de tiempo únicos disponible para 10 participantes con TAR a largo plazo (TP35, TP36, TP37, TP38, TP40, TP41, TP43, TP44, TP45, TP46; **K**). Cada línea horizontal representa una variante. El número de lecturas que representan a esa variante y su proporción dentro de en la muestra se muestran a la derecha. Cada círculo representa a un sitio CpG canónico; la numeración de los sitios se basa en la referencia HXB2: negro-metilado, blanco-no metilado, gris-mutado.

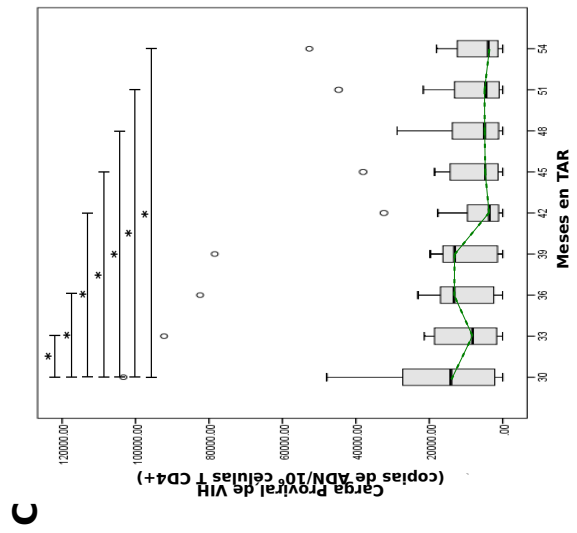
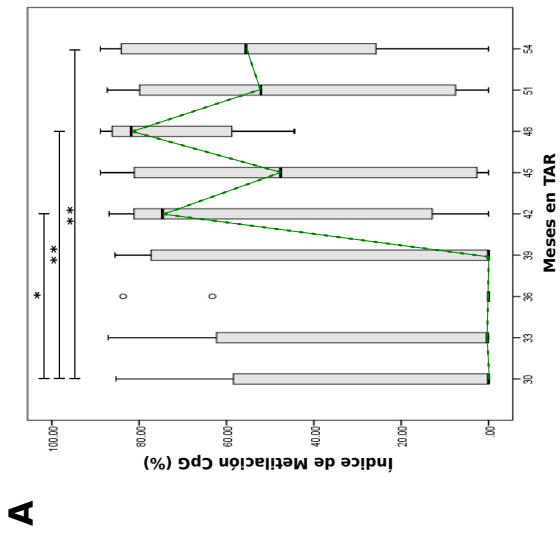
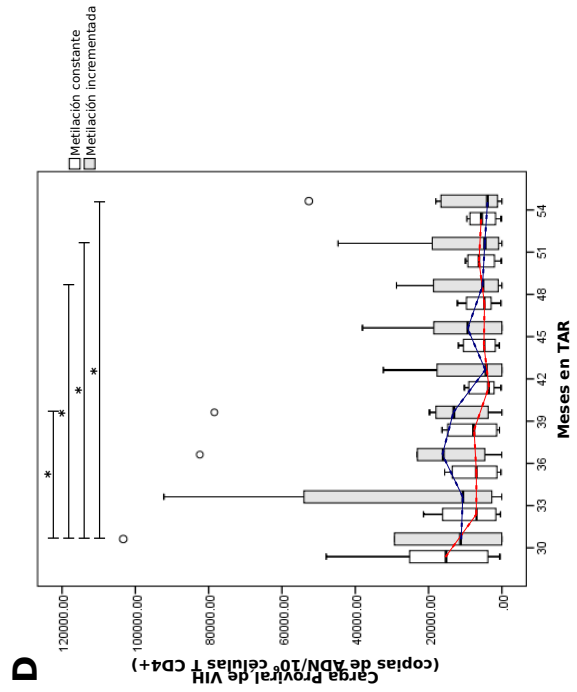
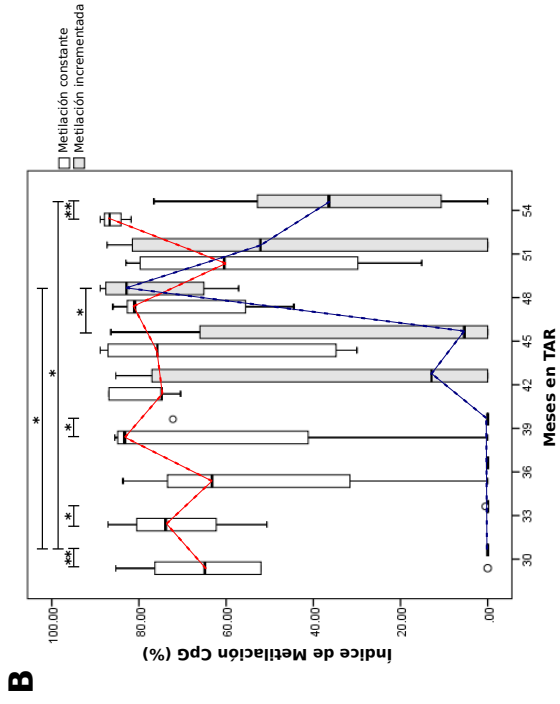
Cambios tempranos en los patrones de metilación y su asociación con las variables clínicas.

	Grupo bajo TAR a corto plazo (30 meses)	Valor-p ^a	Grupo bajo TAR a corto plazo (42 meses)	Valor-p ^b	Valor-p ^d	Grupo bajo TAR a corto plazo (48 meses)	Valor-p ^c	Valor-p ^e	Grupo bajo TAR a largo plazo
Índice de metilación CpG	0 (0-58.4)	0.005	74.7 (0-85.3)	0.273	0.025	81.8 (57.9-86.3)	0.973	0.003	77.9 (70.8-86.7)
Índice de metilación CpG adicional ^h	0.8 (0-21.0)	0.155	27.0 (4.3-45.1)	0.163	0.074	34.8 (20-66.5)	0.014	0.004	20.0 (13.8-20.0)
Índice de metilación no-CpG ⁱ	1.3 (0-9.5)	0.388	0 (0-0)	0.332	0.093	0 (0-5)	0.661	0.208	0 (0-2.2)
pVL residual (copias ARN/mL)	1.4 (0-3.6)	0.737	3.4 (0-7.1)	0.305	0.176	1.4 (0.05-3.3)	0.417	0.449	0.4 (0-2.9)
Carga proviral (copias ADN/10 ⁶ células T CD4+)	14146 (517-29277)	0.526	3638 (228-10216)	0.260	0.005	4972.8 (688.9-14482.2)	0.674	0.00097	9780 (4 057-13716)
Células T CD4+ naïve (%)	25.2 (20.6-52.6)	0.944	29.1 (18.3-49.1)	0.888	0.114	34.3 (17.4-58.1)	0.974	0.465	39.4 (11.2-45.4)
Células T CD4+ memoria (%)	74.8 (47.4-79.4)	0.944	70.9 (50.9-81.7)	0.888	0.114	65.7 (41.9-82.6)	0.974	0.465	60.6 (54.6-88.8)
T _{SCM} (%)	5.6 (3.0-6.9)	0.049	4.2 (2.9-7.1)	0.035	0.508	4.9 (3.4-7.9)	0.042	0.831	9.4 (5.7-15.9)
T _{CM} (%)	23.2 (15.3-31.2)	0.673	21.7 (12.7-33.0)	0.439	0.169	23.8 (16.2-31.7)	0.674	0.764	23.8 (19.5-40.9)
T _{EM} (%)	12.5 (2.1-20.4)	0.398	10.1 (3.2-12.5)	0.205	0.333	8 (1.7-14.5)	0.674	0.413	6.3 (4.8-9.3)
T _{TM} (%)	15.0 (11.3-25.4)	0.091	13.8 (11.4-31.5)	0.049	0.799	13.2 (8.6-29.7)	0.093	0.320	8.3 (4.5-16.6)
T _{NEW} (%)	1.1 (0.6-1.5)	0.003	2.5 (1.6-3.6)	0.526	0.005	1.4 (1-2.7)	0.099	0.005	2.0 (1.5-3.9)
T _{TE} (%)	2.3 (0.3-11.1)	0.181	5.0 (1.2-9.8)	0.944	0.285	1.4 (0.4-7.5)	0.180	0.5771	4.8 (2.7-7.8)

	Grupo bajo TAR a corto plazo (54 meses)	Valor-p ^f	Valor-p ^g	Grupo bajo TAR a largo plazo
Índice de metilación CpG	55.6 (15.1-86.4)	0.459	0.006	77.9 (70.8-86.7)
Índice de metilación CpG adicional ^h	30.5 (17-49.5)	0.162	0.006	20.0 (13.8-20.0)
Índice de metilación no-CpG ⁱ	0 (0-8)	0.834	0.932	0 (0-2.2)
pVL residual (copias ARN/mL)	3.8 (0.19-5.7)	0.091	0.058	0.4 (0-2.9)
Carga proviral (copias ADN/10 ⁶ células T CD4+)	3915.3 (170.1-15174.1)	0.511	0.002	9780 (4 057-13716)
Células T CD4+ naïve (%)	39.9 (24.7-62.5)	0.654	0.002	39.4 (11.2-45.4)
Células T CD4+ memoria (%)	60.1 (37.5-75.3)	0.654	0.002	60.6 (54.6-88.8)
T _{SCM} (%)	8.7 (6.5-9)	0.863	0.004	9.4 (5.7-15.9)
T _{CM} (%)	19.1 (9.4-28.6)	0.281	0.046	23.8 (19.5-40.9)
T _{EM} (%)	8.3 (4-13.3)	0.314	0.492	6.3 (4.8-9.3)
T _{TM} (%)	6.3 (4.5-14)	0.863	0.002	8.3 (4.5-16.6)
T _{NEW} (%)	3.0 (1.6-3.4)	0.511	0.004	2.0 (1.5-3.9)
T _{TE} (%)	5.6 (3.3-13)	0.511	0.004	4.8 (2.7-7.8)

Cuadro 5.2: ^a Grupo bajo TAR a corto plazo (30 meses) vs Grupo bajo TAR a largo plazo;

Cuadro 5.2: ...de pagina anterior. ^b Grupo bajo TAR a corto plazo (42 meses) vs Grupo bajo TAR a largo plazo; ^c Grupo bajo TAR a corto plazo (48 meses) vs Grupo bajo TAR a largo plazo; ^d Grupo bajo TAR a corto plazo (30 meses) vs Grupo bajo TAR a corto plazo (42 meses); ^e Grupo bajo TAR a corto plazo (30 meses) vs Grupo bajo TAR a corto plazo (48 meses); ^f Grupo bajo TAR a corto plazo (54 meses) vs Grupo bajo TAR a largo plazo; ^g Grupo bajo TAR a corto plazo (30 meses) vs Grupo bajo TAR a corto plazo (54 meses); ^h Considera a la Metilación de posiciones CpG que no están presentes en la referencia HXB2 e ⁱ Considera únicamente a la metilación en posiciones diferentes a CpG (CHG y CHH). En todos los casos son mostradas, las medianas (rango intercuartil). Los valores p significativos ($p < 0.05$) se muestran en **negritas**. TAR—terapia antirretroviral; pVL—Carga Viral Residual en Plasma; T_{SCM} —células T de Memoria Stem cell; T_{CM} —células T de Memoria Central; T_{EM} —células T de Memoria Efectora; T_{TM} —células T de Memoria Transicional; T_{NEW} —células T "nuevas"; T_{TE} —células T terminalmente efectoras.



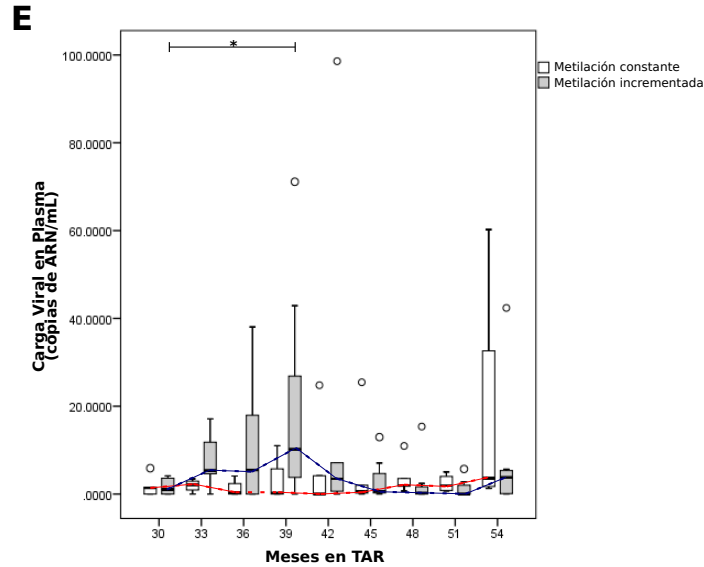


Figura 5.4: Cambios en el índice de metilación CpG, la carga proviral y la carga viral residual en plasma a lo largo del seguimiento del estudio. **A)** El índice de metilación CpG se calculó como la proporción promedio de las variantes metiladas, considerando a los nueve sitios CpG canónicos que fueron analizados y que están presentes en HXB2, en la muestra de cada participante y en cada punto de tiempo; todos los participantes están incluidos. **B)** Los participantes fueron divididos en dos grupos de acuerdo a los patrones de metilación observados a lo largo del tiempo de estudio: *metilación constante* y *metilación incrementada*. Se muestran las variaciones en el índice de metilación CpG para cada grupo. **C)** Se muestra la variación de la carga proviral a lo largo del tiempo; todos los participantes están incluidos. **D)** Se muestra la variación de la carga proviral a lo largo del tiempo para los grupos de *metilación constante* y *metilación incrementada*. **E)** Se muestra la carga viral residual en plasma a lo largo del tiempo para los grupos de *metilación constante* y *metilación incrementada*. En todos los casos, las cajas representan 25-75 % cuartiles, los bigotes valores máximo y mínimo de la serie o hasta 1.5 veces el RIC (Cuando los datos se extienden más allá de esto, significa que hay valores atípicos y entonces se calcula los límites superior (Ls) e inferior (Li)), y las líneas horizontales dentro de las cajas representan a las medianas. Las diferencias entre dos grupos fue señalada mediante unas líneas. * para $p < 0.05$, ** para $p < 0.01$ y *** $p < 0.001$. La línea punteada verde (considera a todos los participantes), la línea punteada roja (considera únicamente al grupo de metilación constante) y la línea punteada azul (considera únicamente al grupo de metilación incrementada).

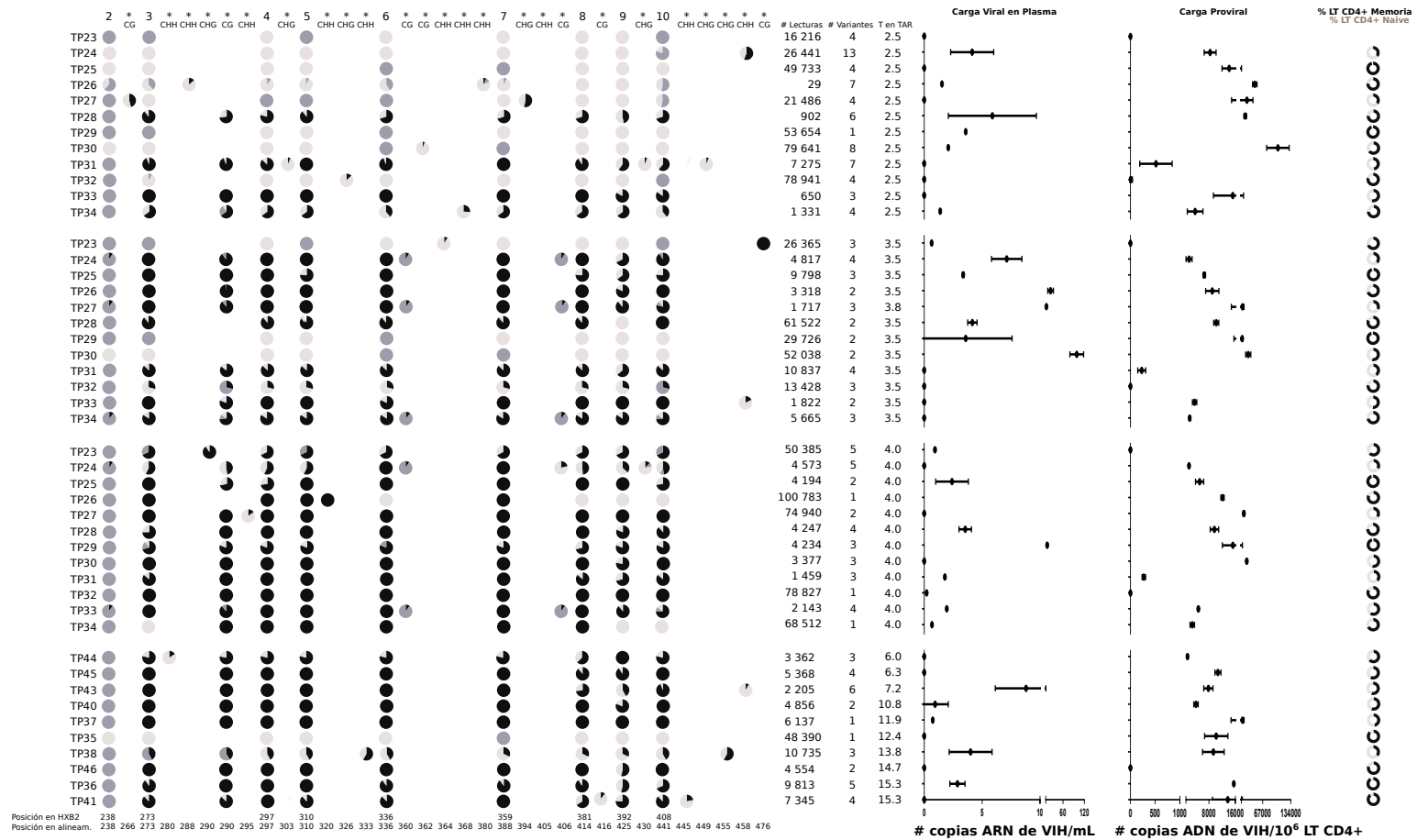


Figura 5.5: Cambios en los patrones de metilación del LTR 5' a lo largo del tiempo.

Figura 5.5: ...de página anterior. Se muestra una comparación de los patrones de metilación para todos los participantes bajo TAR a corto plazo a los 30 meses (2.5 años de seguimiento–**superior**), todos los participantes con TAR a corto plazo a los 42 meses (3.5 años de seguimiento–**segundo grupo**), todos los participantes con TAR a corto plazo a los 48 meses (4 años de seguimiento–**tercer grupo**) y todos los participantes bajo TAR a largo plazo (único punto de tiempo disponible, de 6 a 15 años–**inferior**). Cada línea representa un promedio de todas las variantes observadas por punto de tiempo. Los sitios CpG son mostrados como gráficos de pastel que indican la proporción de variantes metiladas (negro), no metiladas (gris claro) o mutadas (gris oscuro) dentro de la muestra. Las posiciones CpG canónicas están numeradas de acuerdo a HXB2. También se incluyeron a los sitios susceptibles a la metilación tales como CpG adicionales, CHG y CHH, que no se observan en la secuencia de HXB2 (marcados con *). Se muestran el número de lecturas y el número total de variantes obtenidas por punto de tiempo, después del filtrado y del alineamiento. Finalmente se incluyeron las mediciones de la carga viral residual en plasma, la carga proviral y la proporción de células T CD4+ de memoria (CD3+/CD4+/CD28+/-/CD95+; negro) y naive (CD3+/CD4+/CD45RO-/CCR7+/CD28+/CD95-; gris claro) por punto de tiempo.

Variación diferencial de las variables clínicas a lo largo del tiempo entre los grupos de metilación constante y metilación incrementada (continua en la siguiente página).

	Grupo bajo TAR a corto plazo (30 meses)					Grupo bajo TAR a corto plazo (42 meses)						
	Grupo de metilación constante	Valor-p ^a	Grupo de metilación incrementada	Valor-p ^b	Valor-p ^c	Grupo de metilación constante	Valor-p ^a	Grupo de metilación incrementada	Valor-p ^b	Valor-p ^d	Valor-p ^e	Valor-p ^f
Índice de metilación CpG	64.9 (52.0-76.3)	0.244	0 (0-0)	0.001	0.009	74.8 (74.7-86.9)	0.902	12.9 (0-77.0)	0.071	0.116	0.187	0.181
Índice de metilación CpG adicional ^m	21.7 (9.4-34.9)	0.282	0 (0-1.7)	0.003	0.030	27 (8.6-40.6)	0.282	27.0 (0-50.0)	0.219	0.853	1	0.105
Índice de metilación no-CpG ⁿ	2.6 (0-8.3)	0.352	0 (0-11.2)	0.585	1	0 (0-3)	0.560	0 (0-0)	0.356	0.892	0.583	0.422
pVL residual (copias ARN/mL)	1.4 (0-1.5)	0.699	1.0 (0-3.6)	0.862	1	0 (0-4.2)	0.842	3.5 (0.6-7.1)	0.164	0.454	1	0.201
Carga proviral (copias ADN/10 ⁵ células T CD4+)	15241 (3800-25166)	0.462	11250 (22-29277)	0.745	0.792	3638 (2155-9012)	0.221	4264 (10-17670)	0.515	0.930	0.062	0.062
Células T CD4+ naive (%)	25.2 (23.9-45.7)	0.713	28.0 (4.2-64.1)	0.828	0.792	29.1 (27.7-49.1)	0.713	26.3 (4.8-47.0)	0.588	0.662	0.437	0.437
Células T CD4+ memoria (%)	74.8 (54.3-76.2)	0.713	72.1 (35.9-95.8)	0.828	0.792	70.9 (50.9-72.3)	0.713	73.7 (53.0-95.2)	0.588	0.662	0.437	0.437
T _{SCM} (%)	6.4 (4.7-6.6)	0.221	4.5 (1.0-6.9)	0.051	0.662	4.2 (2.9-7.4)	0.111	4.7 (3.4-6.2)	0.065	0.792	0.437	1
T _{CM} (%)	23.2 (23.2-28.8)	0.713	22.3 (15.3-35.1)	0.745	0.930	21.7 (12.7-33.0)	0.540	21.9 (15.5-32.5)	0.515	1	1	0.062
T _{EM} (%)	12.5 (5.4-17.2)	0.270	9.2 (2.1-26.5)	0.745	1	9.0 (7.6-10.1)	0.462	12.1 (3.2-13.8)	0.193	0.246	0.187	1
T _{TM} (%)	13.3 (12.6-15.1)	0.221	15.2 (11.3-25.4)	0.129	0.792	11.6 (11.4-19.5)	0.142	17.4 (11.7-31.5)	0.083	0.536	0.625	1
T _{NEW} (%)	1.4 (0.6-1.56)	0.075	1.1 (0.6-1.2)	0.002	0.536	2.5 (1.6-4.1)	0.624	2.7 (2.3-3.3)	0.588	0.930	0.062	0.062
T _{TE} (%)	4.0 (0.3-11.8)	0.713	1.5 (0.3-4.5)	0.083	0.662	3.9 (1.2-9.6)	0.806	6.3 (2.2-9.8)	0.914	0.792	1	0.187

	Grupo bajo TAR a corto plazo (48 meses)					Grupo bajo TAR a corto plazo (54 meses)						
	Grupo de metilación constante	Valor-p ^a	Grupo de metilación incrementada	Valor-p ^b	Valor-p ^c	Valor-p ^h	Valor-p ⁱ	Grupo de metilación constante	Valor-p ^a	Grupo de metilación incrementada	Valor-p ^b	Valor-p ^j
Índice de metilación CpG	81 (50-84.3)	0.859	82.9 (60.4-88.9)	0.806	0.254	0.062	0.022	86.7 (82.9-88.4)	0.119	36.4 (6.2-55.6)	0.031	0.006
Índice de metilación CpG adicional ^m	33.3 (10-43.2)	0.121	45 (20-81)	0.017	0.290	0.100	0.015	41.4 (23.3-49.9)	0.012	17.6 (3-40)	0.803	0.230
Índice de metilación no-CpG ⁿ	0 (0-6.3)	0.609	0 (0-6.6)	0.864	1	0.789	0.181	0 (0-0)	0.184	0 (0-11.8)	0.623	0.186
pVL residual (copias ARN/mL)	1.9 (1.2-7.2)	0.151	0.2 (0-2.4)	1	0.190	0.625	0.787	3.6 (1.5-46.4)	0.070	3.8 (0-5.7)	0.288	0.635
Carga proviral (copias ADN/10 ⁵ células T CD4+)	4771.7 (1600.7-10896.1)	0.439	5173.8 (10.6-21999.5)	1	0.876	0.062	0.031	5580.9 (942-9127.7)	0.303	3915.3 (13.3-18021)	0.886	0.927
Células T CD4+ naive (%)	31.6 (21.2-51.3)	1	37 (4.4-64.8)	0.962	1	0.812	0.437	43.2 (27-59.9)	0.539	39.9 (5.6-65.3)	0.886	0.927
Células T CD4+ memoria (%)	68.4 (48.7-78.8)	1	62.9 (35.2-95.6)	0.962	1	0.812	0.437	56.8 (40-73)	0.539	60.1 (34.7-94.4)	0.886	0.927
T _{SCM} (%)	7 (6-9.9)	0.594	3.6 (1.5-4.4)	0.009	0.017	0.312	0.562	11.4 (8.5-14.6)	0.635	6.6 (6.4-8.9)	0.536	0.109
T _{CM} (%)	24.9 (16.8-31.3)	1	21.1 (16-37.4)	0.536	1	0.312	0.687	22.5 (11.9-27.9)	0.453	18 (9.1-35.3)	0.363	0.927
T _{EM} (%)	8.7 (5-13.3)	0.309	2.9 (1.3-15.9)	0.886	0.755	0.437	0.687	8.3 (2.7-10.6)	0.733	8.3 (4-17.6)	0.269	0.648
T _{TM} (%)	12.1 (6.6-23.1)	0.371	13.7 (11.6-38.2)	0.087	0.431	0.187	0.843	6.8 (4.7-9.8)	0.635	6.3 (4.4-22)	0.962	0.787
T _{NEW} (%)	2.4 (1.2-3.2)	0.902	1.2 (1-1.8)	0.018	0.343	0.062	0.093	3.2 (2.4-3.4)	0.303	2.3 (1.5-3.6)	0.886	0.648
T _{TE} (%)	3.1 (0.6-15.6)	0.859	0.6 (0.2-6.9)	0.070	0.202	0.437	0.687	6.1 (1.6-13.6)	0.945	5.6 (3.3-13)	0.417	0.927

	Continúa (54 meses)		Grupo bajo TAR a largo plazo
	Valor-p ^a	Valor-p ^b	
Índice de metilación CpG	0.125	0.036	77.9 (70.8-86.7)
Índice de metilación CpG adicional ^m	0.181	0.015	20.0 (13.8-20.0)
Índice de metilación no-CpG ⁿ	0.371	0.787	0 (0-2.2)
pVL residual (copias ARN/mL)	0.250	0.177	0.4 (0-2.9)
Carga proviral (copias ADN/10 ⁶ células T CD4+)	0.125	0.031	9780 (4057-13716)
Células T CD4+ naive (%)	0.125	0.031	39.4 (11.2-45.4)
Células T CD4+ memoria (%)	0.125	0.031	60.6 (54.6-88.8)
T _{SCM} (%)	0.125	0.062	9.4 (5.7-15.9)
T _{CM} (%)	0.461	0.156	23.8 (19.5-40.9)
T _{EM} (%)	0.875	0.687	6.3 (4.8-9.3)
T _{TM} (%)	0.125	0.031	8.3 (4.5-16.6)
T _{NEW} (%)	0.250	0.031	2.0 (1.5-3.9)
T _{TE} (%)	0.625	0.031	4.8 (2.7-7.8)

Cuadro 5.3: ^a Grupo de metilación constante vs. Grupo bajo TAR a largo plazo; ^b Grupo de metilación incrementada vs. Grupo bajo TAR a largo plazo; ^c Grupo de metilación constante (30 meses) vs. Grupo de metilación incrementada (30 meses); ^d Grupo de metilación constante (42 meses) vs. Grupo de metilación incrementada (42 meses); ^e Grupo de metilación constante (30 meses) vs. Grupo de metilación constante (42 meses); ^f Grupo de metilación incrementada (30 meses) vs. Grupo de metilación incrementada (42 meses); ^g Grupo de metilación constante (48 meses) vs. Grupo de metilación incrementada (48 meses); ^h Grupo de metilación constante (30 meses) vs. Grupo de metilación constante (48 meses); ⁱ Grupo de metilación incrementada (30 meses) vs. Grupo de metilación incrementada (48 meses); ^j Grupo de metilación constante (54 meses) vs. Grupo de metilación incrementada (54 meses); ^k Grupo de metilación constante (30 meses) vs. Grupo de metilación constante (54 meses); ^l Grupo de metilación incrementada (30 meses) vs. Grupo de metilación incrementada (54 meses); ^m Considera la metilación de posiciones CpG, no presentes en la referencia HXB2. ⁿ Considera únicamente a la metilación de posiciones diferentes a CpG (CHG y CHH). En todos los casos se muestran medianas (rango intercuartil). Los valores p significativos ($p < 0.05$) se muestran en **negritas**. TAR—terapia antirretroviral; pVL—Carga viral residual en plasma; T_{SCM}—células T de Memoria Stem cell; T_{CM}—células T de Memoria Central; T_{EM}—células T de Memoria Efectora; T_{TM}—células T de Memoria Transicional; T_{NEW}—células T "nuevas"; T_{TE}—células T terminalmente efectoras.

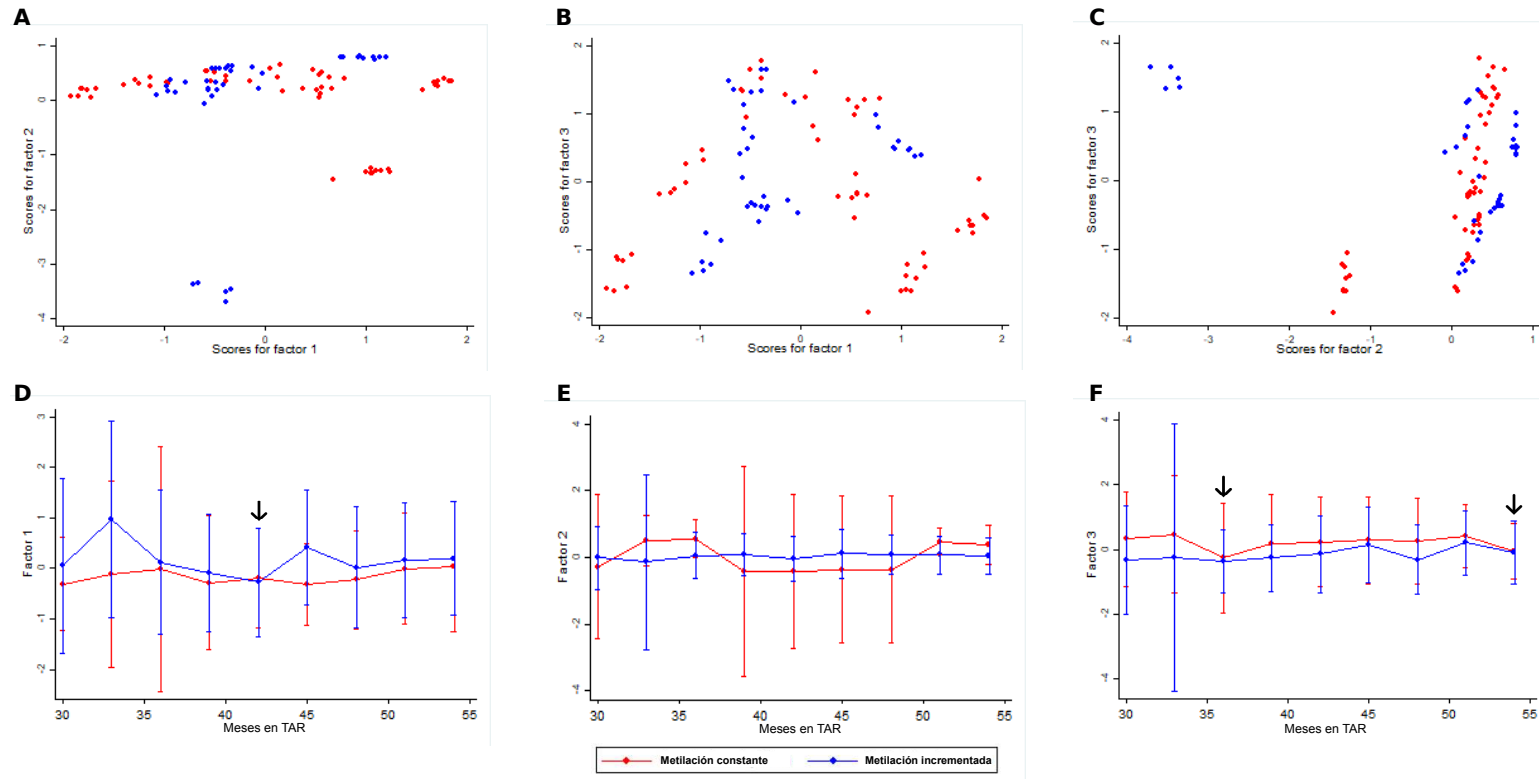


Figura 5.6: Análisis Factorial Exploratorio utilizando Componentes Principales (AFECP). Se realizó un análisis factorial utilizando componentes principales con rotación varimax con el objetivo de describir la estructura de los datos. A), B), C) Las puntuaciones de los factores para los tres factores principales que en conjunto explicaron el 50.86% de la varianza común total son mostrados para cada punto temporal de seguimiento, considerando a los 12 participantes con TAR a corto plazo. D), E), F) Se muestran las tendencias temporales para cada uno de los tres factores. El color denota a los participantes clasificados en los grupos de metilación constante y metilación incrementada. La ↓ señala el tiempo de seguimiento, en donde los dos grupos se traslapan.

Capítulo 6

Discusión.

”No hables al menos que puedas mejorar el silencio”.

Jorge Luis Borges

Muchos de los esfuerzos para lograr una cura del VIH se han centrado en el desarrollo de estrategias para revertir la latencia viral que, en teoría, podrían permitir la expresión de genes virales que conduzcan al reconocimiento y a la erradicación del virus. La contribución de la metilación en la región promotora llamada LTR 5' del VIH a la latencia sigue siendo controversial, ya que los estudios previos mostraron resultados contrastantes [53, 57, 59]. Nuestro estudio mostró por primera vez que la metilación del LTR 5' del VIH es dinámica a lo largo del tiempo, con diferentes patrones en la metilación asociados con las características clínicas basales de las personas que iniciaron el TAR entre las que destacan: Edad, ganancia de células T CD4+ a los 48M de TAR (cuentas y porcentaje), disminución de células T CD8+ a los 48M de TAR (cuentas y porcentaje), delta de CD4/CD8 a los 48 meses, el número de cambios en el régimen de TAR, el tiempo transcurrido para lograr una carga viral indetectable, la clase de tercer fármaco en el régimen de TAR y el nadir del recuento de células T CD4+. Estas observaciones podrían explicar las diferencias entre los estudios previos respecto al papel de la metilación del LTR 5' que mostraron resultados contrastantes [53, 57, 59], ya que algunos trabajos reportaron LTRs 5' hipermetilados y en otros reportaron LTRs 5' hipometilados o no metilados en pacientes VIH+ avirémicos.

Las fortalezas y los aspectos novedosos de este estudio incluyen la implementación de una metodología basada en secuenciación de siguiente generación, capaz de analizar detalladamente la composición de las variantes del LTR 5' y su estado de metilación con alta sensibilidad, en contraste con estudios previos que utilizaron secuenciación clonal tradicional de Sanger, con capacidad limitada de análisis de variantes. Así mismo, en este estudio presentamos por primera vez una evaluación longitudinal de los cambios

en la metilación del LTR 5' en células T CD4+ de una cohorte de individuos bajo TAR a corto plazo y además comparamos estos resultados con un grupo transversal de personas bajo TAR a largo plazo. Nuestro estudio mostró que la metilación del LTR 5' del VIH en células T CD4+ totales de sangre periférica es dinámica a lo largo del tiempo, y que existen al menos dos patrones de metilación diferentes (que se separaron significativamente a los meses 30, 33, 39 y 54), y que los patrones de metilación mostraron asociación tanto con las características clínicas basales como de seguimiento de los individuos (**Sección 5.3**).

Los análisis estadísticos en este trabajo compararon el tiempo de tratamiento (30M vs 42M, 48M y 54M), la decisión en esta selección de tiempos obedece a que fue evidente en los datos de secuenciación, que el grado de metilación CpG comienza a incrementarse al mes 42, llegando al mes 48 al valor más alto (**Figura 5.4A**). Al realizar la comparación de la carga proviral a los 42M vs 30M y 48M vs 30M (**Figura 5.4C**), coincide con una disminución significativa en la carga proviral.

Se ha sugerido previamente que el VIH latente reside principalmente en células T CD4+ en reposo [84, 85] y que la metilación del ADN puede ser un evento tardío que refuerza el silenciamiento de provirus ya latentes [86, 87]. De hecho, la hipermetilación de los sitios CpG dentro del LTR 5' del VIH ha sido asociado con el silenciamiento del promotor del VIH en células Jurkat con infección latente y en células T CD4+ en reposo de pacientes avirémicos bajo TAR [53, 54]. Sin embargo, otros estudios que analizaron el reservorio viral latente de pacientes avirémicos bajo TAR encontraron muy poca o nada de metilación en el LTR 5' de VIH en células T CD4+ de memoria en reposo o PBMC, cuestionando la importancia de la metilación como mecanismo de mantenimiento del reservorio latente [55, 57]. Hay varios factores que podrían explicar las observaciones contrastantes sobre el papel de la metilación de LTR 5' en la latencia del VIH, incluyendo el tiempo de infección con VIH, el régimen de TAR utilizado, el tiempo bajo TAR, el tipo celular estudiado, la composición de las subpoblaciones de células T CD4+, la tasa de reemplazo de células T CD4+ con infección latente, la proporción de provirus defectuosos, los sitios de integración del VIH [56, 57, 59] y posiblemente una dieta rica en donadores de grupo metilo (por e.g Biotina, folato, metionina, colina y betaina) [88].

¿Cómo se encuentra la expresión de las enzimas involucradas en mantener los patrones de metilación del ADN en los pacientes VIH+ bajo TAR? El trabajo de Rosca A., y colab. mostró que los pacientes VIH+ bajo TAR y con una carga viral de VIH indetectable, tuvieron menor nivel de expresión del gen DNMT3A, pero en cambio mostraron un mayor nivel de expresión en el gen de DNMT1, en comparación con los pacientes con una carga viral de VIH detectable. Respecto al tiempo de TAR, encontraron que un tiempo de TAR mayor (mediana de TAR 10.2 años), incrementó los niveles de expresión de DNMT3A vs el grupo con una TAR <10.2 años, mientras que el nivel en la expresión de la DNMT1 no se modificó con el tiempo de TAR [89]. Si consideramos que la DNMT3A, es la enzima responsable de la metilación no canónica, en nuestro trabajo

encontramos que la comparación entre los grupos bajo TAR a corto plazo vs grupo bajo TAR a largo plazo no mostró diferencias significativas entre los grupos respecto al índice de metilación no CpG (**Cuadros 5.2, 5.3**).

Nuestro estudio identificó asociaciones entre algunas variables y los cambios longitudinales en los patrones de metilación. Entre las variables asociadas (características basales) se encontraron: la edad, la ganancia de células T CD4+ a los 48M de TAR (numérico y porcentaje), la disminución de células T CD8+ a los 48M de TAR (numérico y porcentaje), el número de cambios en el régimen de TAR, el tiempo transcurrido para lograr una carga viral indetectable, la clase de tercer fármaco en el régimen de TAR y el nadir del recuento de células T CD4+.

Dentro de las características del seguimiento destacan: disminución significativa de la carga proviral durante el seguimiento (únicamente en el grupo de metilación incrementada), tiempo bajo TAR, la composición de las subpoblaciones de células T CD4 (destacable en el grupo de metilación incrementada al final del seguimiento vs mes 30), porcentaje de células T CD4+ de memoria, naive, T_{CM} y T_{TM} y el porcentaje de células T CD4+ T_{EM} y T_{TE} .

Es importante mencionar que el Análisis Factorial Exploratorio utilizando Componentes Principales (AFECP), es capaz de distinguir y cuantificar tanto a la varianza en común como a la varianza única. En la obtención de las variables asociadas a cada factor, se considera solo a la varianza en común. El análisis, encontró que 4/29 variables $\%T_{SCM}$, $\%T_{NEW}$, carga proviral e índice de metilación CpG son las variables más importantes al tener una varianza única (>0.5).

Además, como se sugirió recientemente [90], otros factores metodológicos también podrían explicar las diferencias y limitar las conclusiones en los análisis de metilación, incluyendo el uso de líneas celulares vs. **células T CD4+ primarias *in vitro***, muestras congeladas vs. **muestras frescas**, el tipo de tratamiento con bisulfito (sin buffer protector de ADN vs **con buffer protector de ADN**), el sesgo en el muestreo, la estrategia de amplificación por PCR (PCR punto final vs **PCR anidada**) y primers utilizados, la técnica de secuenciación (Sanger vs. **secuenciación de siguiente generación**) (en **negritas**, se incluyen las características empleadas en este trabajo).

Utilizando el Análisis Factorial Exploratorio utilizando Componentes Principales (AFECP), nuestro estudio identificó que la edad, la ganancia de células T CD4+ (48 meses después del inicio del TAR), la disminución en las células T CD8+ (48M después del inicio del TAR), el cambio en el cociente CD4/CD8 y el porcentaje de células naïve, memoria, T_{CM} y T_{TM} como las variables más importantes (combinación), representando el 21.4% de la varianza común total en los participantes del estudio. Adicionalmente, el tiempo bajo TAR, el tiempo para lograr una pVL indetectable y el tercer fármaco en el TAR actual fueron también identificadas como variables importantes que explicaron la varianza común total en el segundo eigenvector. Dado que la metilación del ADN podría ser un evento tardío que refuerza el silenciamiento de los virus ya latentes, más que impulsar la latencia desde el principio, el tiempo de infección y el tipo de célula podrían

ser factores muy importantes en los análisis de metilación. De hecho, la metilación CpG puede ser un mecanismo de latencia más común en las células T CD4+ de memoria que son más propensas a la proliferación que en las células T CD4+ en reposo [91]. La selección tardía de los provirus latentes en la subpoblación de células de memoria podría resultar en la acumulación de los provirus con promotores metilados en comparación con los provirus con promotores no metilados encontrados en la subpoblación de células en reposo [91]. Recientemente, Trejbalova et al. mostró una correlación positiva entre el tiempo bajo TAR y el grado de metilación CpG en el LTR 5' del VIH en células T CD4+ en reposo [59]. Sin embargo, es difícil discriminar si la correlación se debió al tiempo bajo TAR únicamente, al tiempo de infección por VIH previo al inicio de TAR o a ambos. Nuestro estudio aporta información adicional al campo, mostrando que la metilación del LTR 5' aparece de forma temprana después del inicio del TAR en al menos dos patrones diferentes que pueden asociarse con las características clínicas basales de cada persona. Interesantemente, mientras un grupo de participantes mantuvo niveles de metilación bastante altos y constantes a lo largo del período de seguimiento (grupo de metilación constante); se produjo un aumento significativo en la metilación CpG en un subgrupo de participantes (58 %) alrededor de los 42 meses (3.5 años) después del inicio del TAR (grupo de metilación incrementada). Al mes 42, comenzó a cambiar el patrón de la metilación CpG en el grupo de metilación incrementada, alcanzando un valor máximo al mes 48 del seguimiento (**Cuadro 5.3**), (**Figura 5.4A**).

Observamos una disminución significativa en la carga proviral a lo largo del seguimiento y más específicamente en el grupo de metilación incrementada (**Cuadros 5.2, 5.3**), (**Figura 5.4D**), una observación que fue reportada recientemente en una cohorte de gran tamaño $n= 1057$ [92].

El subgrupo de metilación incrementada tuvo diferentes características basales en comparación con el subgrupo de metilación constante como se muestra en el Análisis Factorial Exploratorio utilizando Componentes Principales, y estuvo compuesto principalmente por individuos más jóvenes que se presentaron tardíamente a la atención clínica (cuentas de nadir de células T CD4+: 13 células/mm³) y mostraron una mayor ganancia de células CD4+. Estas observaciones podrían ayudar a explicar las observaciones conflictivas de los estudios previos y podrían ser importantes para las estrategias de cura.

Nuestro estudio es único en el sentido de que se utilizó un sistema de conversión por bisulfito altamente efectivo con baja degradación de ADN, una estrategia de PCR anidada que selectivamente amplificó y analizó el LTR 5' convertido con bisulfito y una estrategia de secuenciación de siguiente generación altamente sensible que detectó variantes en baja frecuencia más eficiente que las estrategias de secuenciación clonal basadas en el método Sanger. Como los análisis de metilación del LTR 5' se realizaron en células T CD4+ totales, consideramos la composición de las subpoblaciones de células T CD4+ en cada muestra como una posible variable confusora. Finalmente, el diseño de nuestro estudio fue longitudinal, lo que permitió estudiar la dinámica de la metilación

del LTR 5' a lo largo del tiempo en el mismo individuo.

Con nuestro diseño experimental, fuimos capaces de identificar entre 1 y 13 variantes diferentes (mediana: 3) por punto de tiempo con una mediana de cobertura de 16 270 reads y, además, pudimos estimar la prevalencia de cada variante dentro de la muestra, lo que nos permitió ponderar la contribución de cada variante al panorama proviral general y el peso de la metilación. Esto supera ampliamente el número promedio de las clonas analizadas por muestra en otros estudios que utilizaron estrategias de secuenciación basadas en Sanger. Nuestros controles mostraron un proceso de conversión por bisulfito altamente eficiente (100% de conversión), que combinado con nuestro diseño de primers (que consideró a las citosinas convertidas por bisulfito) (**Cuadro 4.2**) y una estrategia de filtrado que permitió estudiar la metilación específicamente en los productos de PCR que fueron convertidos por bisulfito.

Es importante mencionar algunas debilidades de nuestro estudio:

1) Como secuenciamos exclusivamente el LTR 5' y no el genoma completo de VIH, no pudimos evaluar el papel de la metilación del LTR 5' únicamente en los provirus competentes para la replicación, y tampoco pudimos evaluar el papel de metilación en otras partes del provirus o en el sitio de integración.

2) Se requiere definir *in vitro* si los sitios de metilación estudiados tendrían relevancia para la expresión del virus. En el caso de la metilación canónica CpG, esta ha sido el único tipo de metilación que ha sido estudiada, en el contexto de la infección por VIH y los sitios de unión a los factores de transcripción más estudiados y más relevantes son NFkB, y AP-1. En algunos modelos ya existen las asociaciones entre la secuencia mutada de un sitio de unión a un factor de transcripción, con la no unión del factor de transcripción y conlleva a una menor expresión de GFP [93]. Hasta hace muy poco tiempo, ya es posible reproducir un patrón de metilación en una secuencia de ADN y es posible analizar su expresión. Se podría emplear un modelo que incluya la secuencia de los promotores de estos pacientes (incluyendo el patrón de metilación), debido a que la unión del factor de transcripción puede ser impedida estéricamente ya sea por la ocupación de una Citosina metilada/remodelador/molécula de andamiaje, también es importante el motivo en la secuencia (6-16 nucleótidos), que puede también modificar la afinidad/avidez del factor de transcripción por ese sitio. En el caso de la metilación no canónica (CHG y CHH), al no haber sido estudiada en el contexto de VIH, no conocemos si el sitio metilado se comportará de la misma manera que la metilación CpG.

3) Quizás sería importante el incluir el estado de metilación de un gen/genes de referencia(s) para definir si los cambios en la metilación son específicos de la secuencia proviral o no. Este punto se puede abordar de dos maneras. En la primera, se puede analizar el metiloma global, utilizando el ADN de las células T CD4+ totales. El obtener el Índice de metilación CpG global en el genoma y el relacionarlo con el Índice de metilación CpG de VIH, nos podría indicar si hay una misma tendencia. En la segunda opción, sería el elegir un gen de referencia (que esté presente en todos los estados

de diferenciación de las células T CD4+), se mediría el estado de metilación de este gen y nuevamente se buscaría relacionar con el Índice de metilación CpG de VIH, esto nos podría indicar si hay metilación específica del virus o no. Propongo algunos candidatos para gen de referencia: CD132 (señalización de citocinas cadena gamma común) y CXCR4 (receptor de quimiocina). Ambas moléculas, son expresados en todos los estados de diferenciación de los linfocitos TCD4+ [94]

4) Nuestros análisis fueron realizados en células T CD4+ totales debido a las limitaciones en el número de células y en la mayor posibilidad de amplificación de la región del LTR en el ADN tratado con bisulfito y, por lo tanto, no fue posible evaluar la metilación en subpoblaciones de células específicas o en células T CD4+ en reposo.

5) Debido a que las cargas provirales fueron variables entre los participantes (o aún dentro del mismo individuo a lo largo del tiempo), incluso cuando se utilizó la misma cantidad de ADN para la conversión por bisulfito y como templado para la amplificación por PCR, el número de provirus analizados en las diferentes muestras fue también variable.

6) Finalmente, como fue sugerido previamente [90], la carga proviral y la pVL residual podrían no ser las variables más apropiadas para evaluar el papel de la metilación del LTR 5' en la latencia en el contexto de nuestro estudio, ya que se analizaron tanto a los provirus defectuosos y a los competentes para la replicación (los cuales no pueden ser diferenciados con nuestro método). Y respecto a la carga viral, para obtener una asociación directa y quizás fuerte con el índice de metilación en el provirus se requerirían: a) aislar y trabajar con las células productoras de virus competentes para la replicación y b) retirar el TAR. En el primer caso, aun no se ha logrado descubrir algún marcador de superficie que nos permita distinguir con alta confianza a las células productoras de virus competentes para la replicación y en el segundo caso no es ético quitar el TAR, a los pacientes debido a la adquisición de mutaciones de resistencia y al repunte de la carga viral y proviral.

Sin embargo, son buenas variables proxy, dada nuestra alta profundidad de muestreo, que aumentó significativamente la sensibilidad para incluir a los provirus intactos. De hecho, considerando nuestra mediana de la carga proviral (8 129 copias de ADN/10⁶ células CD4+), al comenzar con 900 ng de ADN para realizar la conversión por bisulfito, estaríamos trabajando con 136 364 genomas humanos, de los cuales 1 109 podrían contener ADN del VIH. Si consideramos que el 88 % de los provirus son defectuosos [47], podríamos tener un total de 133 provirus intactos en la muestra. Como la mitad del volumen de ADN convertido por bisulfito fue utilizado para la PCR, estaríamos analizando 67 provirus intactos (esto sin considerar la degradación inherente del ADN debido al tratamiento con bisulfito).

7) Además, este estudio no contempló grandes extracciones de sangre o la recolección de muestras obtenidas por leucoféresis, necesarias para realizar ensayos de reactivación viral en cultivo ("viral outgrowth assays", VOA) [45, 95], dada la baja proporción de células sanguíneas que teóricamente se espera que contengan el VIH latente *in vivo* (1

de cada 10^6 de células T CD4+ en reposo) [45, 96] en comparación con los modelos de latencia *in vitro*. Por esta razón no fue posible complementar nuestras observaciones con ensayos de reactivación viral.

Finalmente, es necesaria una mayor comprensión respecto a la alternancia observada entre las variantes no metiladas y las metiladas con el objetivo de entender el papel de la metilación del LTR 5' en la latencia del VIH. Esto podría tratarse de un fenómeno complejo con varios mecanismos posibles en juego. Con respecto a la ausencia o al bajo nivel de metilación del LTR 5' en células con infección latente, los mecanismos importantes podrían incluir: 1) prevención de la metilación *de novo* [58, 97], 2) alta tasa de desmetilación [97], o 3) la desaminación de 5mCpG o CpG: la detección de un sitio TpG podría ser el resultado de la desaminación de 5mCpG o CpG sin metilar (ambos producirían UpG, que a su vez se observarían como TpG después de la amplificación por PCR). De hecho, se ha sugerido que la metilación CpG en el ADN podría ser el principal mecanismo responsable de la reducción numérica de los sitios CpG en el genoma del VIH (el trabajo de Alinejad-Rokny y colab. consideró la representación de todos los motivos CpG hasta el nivel de tetranucleótidos en >2000 secuencias genómicas completas de VIH) [98]. Este tipo de mutación podría originar provirus defectuosos para la replicación o eliminar el silenciamiento transcripcional [99].

Con respecto al enriquecimiento de variantes metiladas, algunos mecanismos posibles incluyen: 1) acumulación de provirus con promotores metilados en células de memoria a través de la proliferación homeostática [100, 101], 2) cambio en la proporción de las subpoblaciones celulares que contienen el ADN proviral de VIH [58], 3) metilación *de novo* [59, 89], y 4) la activación celular de bajo nivel.

En cuanto a la pérdida de los sitios CpGs podría explicarse mediante dos mecanismos: 1) enzimático debido a la acción de AID o APOBEC [102, 103, 104] y 2) un mecanismo pasivo debido a la desaminación de sitios mCpGs [98]. Ejemplificando al primer mecanismo, Doerfler et al., analizó el patrón de metilación del ADN proviral tanto del LTR 5' como del LTR 3' del VIH en PBMCs provenientes de pacientes con diferente espectro en la progresión de la enfermedad [56]. Los pacientes con una enfermedad avanzada mostraron provirus de VIH esencialmente no metilados en sus LTRs, pero además detectaron 3 sitios CpG que habían sido selectivamente mutados a TpAs [56]. La observación anterior podría explicarse debido a la acción de la enzima AID (desaminasa de citosina inducida por activación) que es capaz de editar el ADN induciendo un rompimiento en la doble cadena, lo que lleva a la reorganización de un locus. Se genera una edición en el ADN de $C \rightarrow U$ creando apareamiento defectuoso U:G que después de la replicación conduce a la fijación de las transiciones de $C \rightarrow T$ o de $G \rightarrow A$, dependiendo de la cadena que fue editada. Los sitios U:G pueden someterse a una escisión de base o al proceso de reparación (mismatch repair), donde esta disfunción puede conducir tanto a transiciones como a transversiones en estos sitios [102, 103]. Para ejemplificar al segundo mecanismo responsable de la depleción de los sitios CpG en el genoma de VIH, Alinejad-Rokny et al., analizaron la representación de todos los

motivos CpG hasta el nivel de tetranucleótidos en alrededor de 2 368 secuencias genómicas completas de VIH-1 (obtenidas de la base de datos del Laboratorio Nacional de los Álamos) y también en el genoma humano. Para lograr este objetivo, cuantificaron el porcentaje de sitios CpG metilados a nivel de tri y tetranucleótidos, usando datos del metiloma de PBMCs (genoma completo) de la base de datos de NGSmethDB [98]. Para investigar si la metilación del ADN ha jugado un papel importante en la depleción de los sitios CpG en el genoma de VIH-1, Alinejad-Rokny et al., analizaron a los patrones de metilación del ADN provenientes de PBMCs humanas. Los resultados mostraron que la metilación CpG es un mecanismo dependiente de secuencia. En el genoma humano los sitios CpGWW, WWCPG (donde W=T/A) son metilados en mayor proporción en comparación con los flanqueados por C/G y son menos representados en el genoma de VIH-1. Únicamente en los genomas de VIH y del virus de inmunodeficiencia del simio, los motivos CpG flanqueados por T/A tuvieron una más baja representación comparada con los flanqueados por C/G. Al buscar una correlación entre la depleción de los sitios CpG y el porcentaje de metilación, encontraron una correlación inversa. Lo anterior sugiere que la metilación podría ser el principal mecanismo responsable de la depleción de sitios CpG en el genoma de VIH [98].

Nuestras observaciones coinciden con los resultados observados en los dos trabajos anteriores ya que también observamos la pérdida de sitios CpG debido a mutaciones (**Figuras 5.1, 5.2, 5.3, 5.5**).

Los trabajos que han utilizado a monoagentes anti-latencia, nos han permitido conocer, de forma parcial, el efecto de una marca epigenética sobre el reservorio; sin embargo, la tendencia de los trabajos más recientes nos obliga a conocer el papel "directo o indirecto" de este agente en más de una marca epigenética, por ejemplo el ácido valproico (un HDACi), no solo realiza su función de inhibir la función de una enzima desacetilasa de histonas, sino que también es capaz de incrementar la metilación global en el ADN proveniente de una línea celular de neuroblastoma, resultando en efectos netos contrarios [105]. En un caso similar, vorinostat, un HDACi, contribuye también a la metilación *de novo* del LTR 5' del VIH vía reclutamiento de DNMT1 [59]. Por tanto, las estrategias terapéuticas que buscan activar al reservorio latente del VIH, podrían paradójicamente incrementar el grado de metilación en el LTR 5' del VIH y aumentar así la estabilidad del reservorio [59]. A pesar de que existen aún muchas preguntas respecto al papel de la metilación del LTR 5' del VIH, actualmente ya se cuenta con inhibidores de la DNMT1 y DNMT3. La 5' azacitidina es un inhibidor de DNMT1, que evita el reclutamiento de MBD2 y HDAC2 al LTR 5' del VIH [106]. Por otro lado Nanaomicina A, es un antibiótico que actúa como inhibidor selectivo para la DNMT3B, siendo capaz de inducir una desmetilación genómica global [107]. Sin embargo, el empleo de las estrategias basadas en inhibidores de metil-transferasas de mantenimiento dista mucho todavía de llegar al uso clínico debido a su potencial de alterar el genoma a nivel global y el riesgo de activar a las secuencias del tipo SINEs, LINEs y LTRs.

Algunos trabajos han demostrado que los provirus latentes que son refractarios a

la reactivación en respuesta a activadores celulares, poseen altos niveles de H3k27me3 [52]. El trabajo de Matsuda et al., reveló que la marca H3K27me3 se acumula muy tempranamente en la infección por VIH, particularmente en el LTR viral. Este grupo comprobó que utilizando al inhibidor de PRC2, llamado DZNep, combinado con vorinostat era posible disminuir a la marca H3K27me3, logrando una alta reactivación del reservorio [61].

En la búsqueda para inhibir al proceso epigenético ATP dependiente, el trabajo de Rafati et al., identificó a las moléculas BRG1, BAF250a e INI1 como blancos a depletar en el reservorio latente de pacientes con VIH+ [62].

El conocimiento actual en el tema, obliga a la investigación de agentes activadores del provirus de VIH que sean más específicos, para purgar idealmente al reservorio latente. Por tanto la investigación ahora se ha enfocado más en los miARNs. Por ejemplo el miRH3 es un miARN viral localizado en la región del gen de la RT del VIH, que funciona como enhancer, al unirse a la caja TATA dentro del LTR viral. Otro ejemplo es el miR29a, un miARN celular que se ha asociado con el mantenimiento de la latencia viral. Por tanto, un combinación de pequeños ARNs sintéticos que bloqueen a miRH3, impidiendo su unión a la caja TATA del VIH, además de una combinación de inhibidores de miARNs celulares que participen en la latencia, podrían ayudar a activar la transcripción del VIH-1 en linfocitos T CD4+ en reposo, con infección latente [108]. Idealmente los nuevos ensayos sobre cura de VIH deberían probar diferentes combinaciones de compuestos anti-latencia que tengan como blanco al menos a tres procesos epigenéticos distintos y que sean administrados concomitantemente con el TAR en pacientes VIH+. Este nuevo esquema de tratamiento, podría administrarse en un corto periodo de tiempo, haciendo eventualmente innecesario el TAR de por vida.

Finalmente, una estrategia actual, más específica, y muy prometedora en la búsqueda de la cura esterilizante de VIH es el empleo de una variante del sistema Crispr-Cas9. Es conocido que la edición mediante Crispr-Cas9 resulta típicamente en una rotura de ADN de doble cadena y que el sistema de reparación celular conduce a la vía de reparación (unión) por recombinación no homóloga (NHEJ), que a menudo introduce pequeñas inserciones o deleciones que pueden, por ejemplo, cambiar el marco de lectura de una región codificante o puede interrumpir un sitio de unión a factor de transcripción [109].

Una variante de este sistema, utilizaría varios guías de ARN (son varios porque deberán cubrir la heterogeneidad viral) cuyo fin será el ubicar específicamente a los provirus a eliminar. Posteriormente, el sistema estará acoplado a un remodelador epigenético, en este caso una buena opción sería TET y TDG, mismos que editarían la Citosinas metiladas y podrían eliminar el silenciamiento transcripcional. Si aún se siguieran presentando mutaciones y el resultado por tanto no fuera satisfactorio, aún podría emplearse algún otro remodelador de la cromatina, e.g. HDACi.

Capítulo 7

Conclusiones.

”Revertir lo irreversible”.

Richard S. Jones.

En este trabajo mostramos una fuerte evidencia de que la metilación del LTR 5' del VIH puede variar significativamente a lo largo del tiempo en los provirus provenientes de células T CD4+ circulantes y que la metilación puede observarse desde etapas tempranas de la administración del TAR. También mostramos que la metilación del LTR 5' puede seguir diferentes patrones temporales (denominados metilación constante y metilación incrementada). La separación de estos dos grupos (univariado) fue significativa a los meses 30, 33, 39 y 54 del seguimiento. En particular, los diferentes patrones de metilación en el LTR 5' pueden asociarse con la edad (univariado y Factor 1), ganancia de células T CD4+ a los 48M de TAR (univariado y Factor 1), disminución de células T CD8+ a los 48M de TAR (univariado y Factor 1), el número de cambios en el régimen de TAR (univariado y Factor 3), el tiempo transcurrido para lograr una carga viral indetectable (Factor 2), la clase de tercer fármaco en el régimen de TAR (Factor 2) y el nadir del recuento de células T CD4+ (Factor 3), la disminución significativa de la carga proviral durante el seguimiento (univariado y únicamente en el grupo de metilación incrementada), tiempo bajo TAR (Factor 2), la composición de las subpoblaciones de células T CD4 (univariado, destacable en el grupo de metilación incrementada al final del seguimiento vs mes 30), porcentaje de células T CD4+ de memoria, naive, T_{CM} y T_{TM} (Factor 1) y porcentaje de células T CD4+ T_{EM} y T_{TE} (Factor 3). Adicionalmente encontramos que las variables $\%T_{SCM}$, $\%T_{NEW}$, carga proviral e índice de metilación CpG representan variables muy importantes al tener una varianza única alta. En su conjunto nuestros resultados pueden explicar las diferencias observadas en los estudios previos realizados por otros autores (donde encontraron LTRs hipermetilados, hipometilados o no metilados en individuos VIH+ bajo TAR supresora).

Considerando las principales debilidades de este estudio (diseño y metodológicas),

proponemos estudios adicionales en subpoblaciones celulares específicas y en tejidos, el uso de cohortes más grandes con tiempos de seguimiento más prolongados, un control más estricto de las variables confusoras (e.g. tiempo de infección), considerar la variación intrínseca de la carga proviral en el tiempo de TAR, con el fin de tratar y analizar un número similar de provirus durante el seguimiento y el uso de nuevas técnicas de secuenciación, que permitan la evaluación de la metilación a lo largo del provirus completo y en el sitio de integración, en su conjunto, con el fin de evaluar el papel de la metilación del LTR viral en el silenciamiento proviral, ya que consideramos que tanto el conocimiento del fenómeno así como su modulación futura podrían ser clínicamente relevantes.

Capítulo 8

Consideraciones éticas.

Este estudio fue aprobado por el Comité de Ética en Investigación del Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas (INER) (Registro: DHHS IORG0003948; CONBIOETICA-09-CEI-003-20160427) en la Ciudad de México (código del proyecto: B26-09), y fue conducido de acuerdo a los principios de la Declaración de Helsinki. Todos los participantes dieron su consentimiento informado por escrito antes de la donación de muestras de sangre.

Capítulo 9

Fondos.

Este trabajo fue apoyado por subvenciones del Gobierno de México (Comisión de Equidad y Género de las Legislaturas LX-LXI y de la Comisión de Igualdad de Género de la Legislatura LXII de la Cámara de Diputados de la República Mexicana) a GRT y del Consejo Nacional de Ciencia y Tecnología (CONACyT SALUD-2017-01-289725) a SAR. Los financiadores no tuvieron ningún papel en el diseño del estudio, recopilación de datos, decisión del análisis a publicar, o en la preparación del manuscrito.

Capítulo 10

Contribuciones de los autores.

CNCR, GSMO, SAR, GRT concibieron y diseñaron el estudio; CNCR realizó los experimentos, FJPG realizó los análisis estadísticos; MMF brindó asistencia con la secuenciación de siguiente generación; LKC contribuyó con la inscripción de participantes y AMO proporcionó los archivos clínicos; APCA, CEO y CNCR contribuyeron con el análisis de datos de secuenciación de siguiente generación; ICR asistió con la construcción y el manejo de la base de datos; CNCR y SAR escribieron el artículo. Todos los autores leyeron y aprobaron el manuscrito final.

Anexos

Anexo A

A-E

A.1. Dos patrones Globales de Metilación CpG: Metilación Constante e Incrementada.

Tiempo de TAR (corto plazo; meses)	Grupos de Estudio	Mediana del Índice de metilación CpG	RIC	Valor-p*
30	Metilación constante	64.9	26-80.8	0.009
	Metilación incrementada	0	0-0	
33	Metilación constante	73.9	50.7-87.1	0.026
	Metilación incrementada	0	0-0.275	
36	Metilación constante	63.2	0-83.7	0.051
	Metilación incrementada	0	0-0	
39	Metilación constante	83.3	20.6-85.2	0.037
	Metilación incrementada	0	0-0	
42	Metilación constante	74.8	72.6-86.9	0.116
	Metilación incrementada	12.9	0-79	
45	Metilación constante	75.8	32.4-88	0.102
	Metilación incrementada	5.3	0-71.6	
48	Metilación constante	81	50-84.3	0.255
	Metilación incrementada	82.9	60.4-88.9	
51	Metilación constante	60.4	22.4-81.4	0.924
	Metilación incrementada	52.1	0-86.2	
54	Metilación constante	86.7	82.9-88.4	0.006
	Metilación incrementada	36.4	6.2-55.6	

*Las variables de muestras independientes fueron comparadas utilizando la prueba de Mann-Whitney.

Cuadro A.1: En el estudio longitudinal se observaron dos patrones de metilación CpG: Grupo de metilación constante y metilación incrementada. Esta separación fue significativa a los meses 30, 33, 39 y 54 del seguimiento.

B.2. AFECP: Factores.

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	7.03295	1.44480	0.2425	0.2425
Factor2	5.58815	2.44781	0.1927	0.4352
Factor3	3.14034	0.49383	0.1083	0.5435
Factor4	2.64651	0.44222	0.0913	0.6348
Factor5	2.20428	0.55128	0.0760	0.7108
Factor6	1.65300	0.39264	0.0570	0.7678
Factor7	1.26036	0.17115	0.0435	0.8112
Factor8	1.08922	0.19195	0.0376	0.8488
Factor9	0.89727	0.03530	0.0309	0.8797
Factor10	0.86197	0.23246	0.0297	0.9094
Factor11	0.62951	0.14019	0.0217	0.9312
Factor12	0.48932	0.07217	0.0169	0.9480
Factor13	0.41715	0.02965	0.0144	0.9624
Factor14	0.38750	0.14354	0.0134	0.9758
Factor15	0.24395	0.05618	0.0084	0.9842
Factor16	0.18778	0.06916	0.0065	0.9907
Factor17	0.11861	0.05324	0.0041	0.9948
Factor18	0.06538	0.01727	0.0023	0.9970
Factor19	0.04811	0.01138	0.0017	0.9987
Factor20	0.03673	0.03487	0.0013	0.9999
Factor21	0.00186	0.00180	0.0001	1.0000
Factor22	0.00006	0.00006	0.0000	1.0000
Factor23	0.00000	0.00000	0.0000	1.0000
Factor24	0.00000	0.00000	0.0000	1.0000
Factor25	0.00000	0.00000	0.0000	1.0000
Factor26	-0.00000	0.00000	-0.0000	1.0000
Factor27	-0.00000	0.00000	-0.0000	1.0000
Factor28	-0.00000	0.00000	-0.0000	1.0000
Factor29	-0.00000	.	-0.0000	1.0000

Cuadro A.2: Se empleó el Análisis Factorial Exploratorio utilizando Componentes Principales (AFECP) como un método de reducción de variables para describir la estructura de los datos y resumir la varianza de las variables clínicas basales y de seguimiento en menos dimensiones. Se eligieron a los primeros 6 Factores

D.4. AFECF: Rotación

Rotated factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Uniqueness
EDAD30Maños	-0.4077			0.7705			0.0704
NadirCD4			0.7971				0.1388
CVpreTAR				-0.6591			0.2934
tentreeldx~s				0.8268			0.1241
Nu~4salos48M	0.7425	0.4403					0.0820
DeltaCD4sa~M	0.8317						0.1153
Nu~8salos48M	-0.4591	0.4233	-0.5647				0.1831
DeltaCD8~48M	-0.4944				0.5712		0.2327
DeltaCD4CD~M	0.8798						0.1159
tparaindet~s		-0.8388					0.0615
tdeTAR30Ma~s		0.8383					0.1050
tdeTAR54Ma~s		0.8383					0.1050
MEMORIAUNG~D	-0.9263						0.0612
NAIVEUNGATED	0.9263						0.0612
SCM					0.4890		0.6588
TCM	-0.8352						0.1030
TEM			0.6904				0.3684
TTM	-0.6986						0.2512
TNEW				0.5260			0.5401
TTE			0.7525				0.3263
CVR						0.7997	0.3393
CP						0.5607	0.5037
INDICEMET9S				0.4691			0.6655
Numerodeca~R			-0.6700				0.4789
TARskel~PTCf					0.7818		0.1351
TARskel~3TCf					-0.7818		0.1351
NNRTIespecif		0.8914					0.0853
PIespecif		-0.8914					0.0853
blipsespecif						0.8229	0.3091

(blanks represent abs(loading)<.4)

Factor rotation matrix

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
Factor1	0.8127	-0.4397	0.1932	0.0135	0.2992	0.1384
Factor2	0.4858	0.6219	-0.5397	-0.2590	-0.0925	0.1017
Factor3	0.0307	0.4025	0.7418	-0.3069	-0.1480	0.4132
Factor4	0.1840	0.4412	0.2290	0.8147	0.0680	-0.2250
Factor5	-0.2467	0.2067	-0.0729	-0.0454	0.9117	0.2405
Factor6	-0.0887	-0.1433	-0.2517	0.4156	-0.2103	0.8314

Cuadro A.4: En cada factor, se muestra a las variables que son asociadas, considerando a la varianza común. En la última columna, se muestra el valor de *uniqueness* (unicidad o varianza única)

Literatura Citada

- [1] JELTSCH A., *Beyond Watson and Crick: DNA Methylation and Molecular Enzymology of DNA Methyltransferases*, *Chembiochem.* 2002; 3:274-93.
- [2] NORELLI S., *Could DNA-reactive B lymphocytes be activated through HIV-1 DNA-stimulation involving BCR/TLR-9 pathway to yield antibodies targeting viral DNA?*, *Med Hypotheses.* 2014; 83:659-63.
- [3] JURKOWSKA RZ, JURKOWSKI TP, JELTSCH A., *Structure and function of mammalian DNA methyltransferases*, *Chembiochem.* 2011; 12:206-22.
- [4] JOEL P, SHAO W, PRATT K., *A nuclear protein with enhanced binding to methylated Sp1 sites in the AIDS virus promoter*, *Nucleic Acids Res.* 1993; 21:5786-93.
- [5] SHAO W., *Characterization of HMBP-2, a DNA-Binding Protein that binds to HIV-1 LTR when only one of the three SP1 sites is Methylated*, *J Biomed Sci.* 1997; 4:39-46.
- [6] BEDNARIK D, DUCKETT C, KIM S, PEREZ V, GRIFFIS K, GUENTHNER P, ET AL., *DNA CpG methylation inhibits binding of NF-Kappa B proteins to the HIV-1 long terminal repeat cognate DNA motifs*, *New Biol.* 1991; 3:969-76.
- [7] HOLLIDAY R, PUGH J., *DNA Modification Mechanisms and Gene Activity during Development*, *Science* 1975; 187:226-32.
- [8] EHRLICH M, GAMA M, HUANG L, MIDGETT R, KUO K, MCCUNE R, ET AL., *Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells*, *Nucleic Acids Res.* 1982; 10:2709-21.
- [9] LISTER R, PELIZZOLA M, DOWEN RH, HAWKINS RD, HON G, TONTI-FILIPPINI J, ET AL., *Human DNA methylomes at base resolution show widespread epigenomic differences*, *Nature.* 2009; 462:315-22.

- [10] LAURENT L, WONG E, LI G, HUYNH T, TSIRIGOS A, ONG CT, ET AL., *Dynamic changes in the human methylome during differentiation*, Genome Res. 2010; 20:320-31.
- [11] BIRD A., *DNA methylation and the frequency of CpG in animal DNA*, Nucleic Acids Res. 1980; 8:1499-504.
- [12] SVED J, BIRD A., *The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model*, Proc Natl Acad Sci U S A. 1990; 87:4692-6.
- [13] PFEIFER G, TANG M, DENISSENKO M., *Mutation hotspots and DNA methylation*, Curr Top Microbiol Immunol. 2000; 249:1-19.
- [14] SHEN J, RIDEOUT W, JONES P., *The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA*, Nucleic Acids Res. 1994; 22:972-6.
- [15] LANDER E., *Initial sequencing and analysis of the human genome*, Nature. 2001; 409:860-921.
- [16] FRYXELL KJ, MOON WJ., *CpG mutation rates in the human genome are highly dependent on local GC content*, Mol Biol Evol. 2005; 22:650-8.
- [17] SAXONOV S, BERG P, BRUTLAG DL., *A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters*, Proc Natl Acad Sci U S A. 2006; 103:1412-7.
- [18] SELA N, KIM E, AST G., *The role of transposable elements in the evolution of non-mammalian vertebrates and invertebrates*, Genome Biol. 2010; 11:R59.
- [19] YODER J, WALSH C, BESTOR T., *Cytosine methylation and the ecology of intragenomic parasites*, Trends Genet. 1997; 13:335-40.
- [20] SMIT A., *Interspersed repeats and other mementos of transposable elements in mammalian genomes*, Current Opinion in Genetics and Development. 1999; 9:657-63.
- [21] RAMSAHOYE B, BINISZKIEWICZ D, LYKO F, CLARK V, BIRD A, JAENISCH R., *Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a*, Proc Natl Acad Sci U S A. 2000; 97:5237-42.
- [22] ARAND J, SPIELER D, KARIUS T, BRANCO MR, MEILINGER D, MEISSNER A, ET AL., *In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases*, PLoS Genet. 2012; 8:e1002750.
- [23] JONES PA., *Functions of DNA methylation: islands, start sites, gene bodies and beyond*, Nat Rev Genet. 2012; 13:484-92.

- [24] DYACHENKO OV, SCHEVCHUK TV, KRETZNER L, BURYANOV YI, SMITH SS., *Human non-CG methylation: are human stem cells plant-like?*, Epigenetics. 2010; 5:569-72.
- [25] GOWHER H, JELTSCH A., *Enzymatic properties of Recombinant Dnmt3a DNA Methyltransferase from Mouse: The Enzyme Modifies DNA in a Non-processive Manner and also Methylates Non CpA sites*, J Mol Biol. 2001; 309:1201-8.
- [26] PATIL V, WARD RL, HESSON LB., *The evidence for functional non-CpG methylation in mammalian cells*, Epigenetics 2014; 9:823-8.
- [27] ICHIYANAGI T, ICHIYANAGI K, MIYAKE M, SASAKI H., *Accumulation and loss of asymmetric non-CpG methylation during male germ-cell development*, Nucleic Acids Res. 2013; 41:738-45.
- [28] CHEN PY, FENG S, JOO JW, JACOBSEN SE, PELLEGRINI M., *A comparative analysis of DNA methylation across human embryonic stem cell lines*, Genome Biol. 2011; 12:R62.
- [29] ZILLER MJ, MULLER F, LIAO J, ZHANG Y, GU H, BOCK C, ET AL., *Genomic distribution and inter-sample variation of non-CpG methylation across human cell types*, PLoS Genet. 2011; 7:e1002389.
- [30] LISTER R, PELIZZOLA M, KIDA YS, HAWKINS RD, NERY JR, HON G, ET AL., *Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells*, Nature. 2011; 471:68-73.
- [31] TOMIZAWA S, KOBAYASHI H, WATANABE T, ANDREWS S, HATA K, KELSEY G, ET AL., *Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes*, Development. 2011; 138:811-20.
- [32] KOBAYASHI H, SAKURAI T, IMAI M, TAKAHASHI N, FUKUDA A, YAYOI O, ET AL., *Contribution of intragenic DNA Methylation in Mouse Gametic DNA Methylomes to establish Oocyte specific heritable marks*, PLoS Genet. 2012; 8:e1002440.
- [33] XIE W, BARR CL, KIM A, YUE F, LEE AY, EUBANKS J, ET AL., *Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome*, Cell. 2012; 148:816-31.
- [34] LISTER R, MUKAMEL E, NERY J, URICH M, PUDDIFOOT C, JOHNSON N, ET AL., *Global Epigenomic Reconfiguration During Mammalian Brain Development*, Science. 2013; 341:1237905-.

- [35] VARLEY KE, GERTZ J, BOWLING KM, PARKER SL, REDDY TE, PAULI-BEHN F, ET AL., *Dynamic DNA methylation across diverse human cell lines and tissues*, Genome Res. 2013; 23:555-67.
- [36] PINNEY SE., *Mammalian Non-CpG Methylation: Stem Cells and Beyond*, Biology (Basel). 2014; 3:739-51.
- [37] SHIRANE K, TOH H, KOBAYASHI H, MIURA F, CHIBA H, ITO T, ET AL., *Mouse Oocyte Methylomes at base resolution reveal Genome Wide Accumulation of Non-CpG Methylation and Role of DNA Methyltransferases*, PLoS Genet. 2013; 9:e1003439.
- [38] GUO W, CHUNG WY, QIAN M, PELLEGRINI M, ZHANG MQ., *Characterizing the strand-specific distribution of non-CpG methylation in human pluripotent cells*, Nucleic Acids Res. 2014; 42:3009-16.
- [39] YU M, HON GC, SZULWACH KE, SONG CX, ZHANG L, KIM A, ET AL., *Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome*, Cell 2012; 149:1368-80.
- [40] MALONE CS, MINER MD, DOERR JR, JACKSON JP, JACOBSEN SE, WALL R, ET AL., *CmC(A/T)GG DNA methylation in mature B cell lymphoma gene silencing*, Proc Natl Acad Sci U S A. 2001; 98:10404-9.
- [41] FLEXNER C, PLUMLEY B, BROWN RIPIN DH., *Treatment optimization: an outline for future success*, Curr Opin HIV AIDS. 2013; 8:523-7.
- [42] FURTADO MR, CALLAWAY DS, PHAIR JP, KUNSTMAN KJ, STANTON JL, MACKEN CA, ET AL., *Persistence of HIV-1 transcription in peripheral-blood mononuclear cells in patients receiving potent antiretroviral therapy*, N Engl J Med. 1999; 340:1614-22.
- [43] CHUN TW, STUYVER L, MIZELL SB, EHLER LA, MICAN JA, BASELER M, ET AL., *Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy*, Proc Natl Acad Sci U S A. 1997; 94:13193-7.
- [44] CHUN TW, ENGEL D, MIZELL SB, HALLAHAN CW, FISCHETTE M, PARK S, ET AL., *Effect of interleukin-2 on the pool of latently infected, resting CD4+ T cells in HIV-1-infected patients receiving highly active anti-retroviral therapy*, Nat Med. 1999; 5:651-5.
- [45] ERIKSSON S, GRAF EH, DAHL V, STRAIN MC, YUKL SA, LYSENKO ES, ET AL., *Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies*, PLoS Pathog. 2013; 9:e1003174.

- [46] CILLO AR, SOBOLEWSKI MD, BOSCH RJ, FYNE E, PIATAK M, JR., COFFIN JM, ET AL., *Quantification of HIV-1 latency reversal in resting CD4+ T cells from patients on suppressive antiretroviral therapy*, Proc Natl Acad Sci U S A. 2014; 111:7078-83.
- [47] HO YC, SHAN L, HOSMANE NN, WANG J, LASKEY SB, ROSENBLUM DI, ET AL., *Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure*, Cell. 2013; 155:540-51.
- [48] RASMUSSEN TA, TOLSTRUP M, BRINKMANN CR, OLESEN R, ERIKSTRUP C, SOLOMON A, ET AL., *Panobinostat, a histone deacetylase inhibitor, for latent-virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: a phase 1/2, single group, clinical trial*, The Lancet HIV. 2014; 1:e13-e21.
- [49] DU CHENE I, BASYUK E, LIN YL, TRIBOULET R, KNEZEVICH A, CHABLE-BESSIA C, ET AL., *Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency*, EMBO J. 2007; 26:424-35.
- [50] KEEDY KS, ARCHIN NM, GATES AT, ESPESETH A, HAZUDA DJ, MARGOLIS DM., *A limited group of class I histone deacetylases acts to repress human immunodeficiency virus type 1 expression*, J Virol. 2009; 83:4749-56.
- [51] IMAI K, TOGAMI H, OKAMOTO T., *Involvement of histone H3 lysine 9 (H3K9) methyltransferase G9a in the maintenance of HIV-1 latency and its reactivation by BIX01294*, J Biol Chem. 2010; 285:16538-45.
- [52] FRIEDMAN J, CHO WK, CHU CK, KEEDY KS, ARCHIN NM, MARGOLIS DM, ET AL., *Epigenetic silencing of HIV-1 by the histone H3 lysine 27 methyltransferase enhancer of Zeste 2*, J Virol. 2011; 85:9078-89.
- [53] BLAZKOVA J, TREJBALOVA K, GONDOIS-REY F, HALFON P, PHILIBERT P, GUIGUEN A, ET AL., *CpG methylation controls reactivation of HIV from latency*, PLoS Pathog. 2009; 5:e1000554.
- [54] KAUDER SE, BOSQUE A, LINDQVIST A, PLANELLES V, VERDIN E., *Epigenetic regulation of HIV-1 latency by cytosine methylation*, PLoS Pathog. 2009; 5:e1000495.
- [55] PALACIOS JA, PEREZ-PINAR T, TORO C, SANZ-MINGUELA B, MORENO V, VALENCIA E, ET AL., *Long-term nonprogressor and elite controller patients who control viremia have a higher percentage of methylation in their HIV-1 proviral promoters than aviremic patients receiving highly active antiretroviral therapy*, J Virol. 2012; 86:13081-4.

- [56] DOERFLER W, WEBER S, KEMAL K, WEISER B, KORN K, ANASTOS K, ET AL., *Epigenetic modifications of HIV proviral LTRs: potential targets for cure*, *Retrovirology*. 2012; 9:O4.
- [57] BLAZKOVA J, MURRAY D, JUSTEMENT JS, FUNK EK, NELSON AK, MOIR S, ET AL., *Paucity of HIV DNA methylation in latently infected, resting CD4+ T cells from infected individuals receiving antiretroviral therapy*, *J Virol*. 2012; 86:5390-2.
- [58] WEBER S, WEISER B, KEMAL KS, BURGER H, RAMIREZ CM, KORN K, ET AL., *Epigenetic analysis of HIV-1 proviral genomes from infected individuals: predominance of unmethylated CpG's*, *Virology*. 2014; 449:181-9.
- [59] TREJBALOVA K, KOVAROVA D, BLAZKOVA J, MACHALA L, JILICH D, WEBER J, ET AL., *Development of 5' LTR DNA methylation of latent HIV-1 provirus in cell line models and in long-term-infected individuals*, *Clin. Epigenetics*. 2016; 8:19.
- [60] YOON CH, JANG DH, KIM KC, PARK SY, KIM HY, KIM SS, ET AL., *Disruption of polycomb repressor complex-mediated gene silencing reactivates HIV-1 provirus in latently infected cells*, *Intervirology*. 2014; 57:116-20.
- [61] MATSUDA Y, KOBAYASHI-ISHIHARA M, FUJIKAWA D, ISHIDA T, WATANABE T, YAMAGISHI M., *Epigenetic heterogeneity in HIV-1 latency establishment*, *Sci Rep*. 2015; 5:7701.
- [62] RAFATI H, PARRA M, HAKRE S, MOSHKIN Y, VERDIN E, MAHMOUDI T., *Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency*, *PLoS Biol*. 2011; 9:e1001206.
- [63] VAN DUYN R, GUENDEL I, NARAYANAN A, GREGG E, SHAFAGATI N, TYAGI M, ET AL., *Varying modulation of HIV-1 LTR activity by Baf complexes*, *J Mol Biol*. 2011; 411:581-96.
- [64] HARWIG A, DAS AT, BERKHOUT B., *Retroviral microRNAs*, *Curr Opin Virol*. 2014; 7:47-54.
- [65] PATEL P, ANSARI MY, BAPAT S, THAKAR M, GANGAKHEDKAR R, JAMEEL S., *The microRNA miR-29a is associated with human immunodeficiency virus latency*, *Retrovirology*. 2014; 11:108.
- [66] BARICHIEVY S, NAIDOO J, MHLANGA MM., *Non-coding RNAs and HIV: viral manipulation of host dark matter to shape the cellular environment*, *Front Genet*. 2015; 6:108.
- [67] RICE AP., *Roles of microRNAs and long-noncoding RNAs in human immunodeficiency virus replication*, *Wiley Interdiscip Rev RNA*. 2015; 6:661-70.

- [68] SUZUKI K, AHLENSTIEL C, MARKS K, KELLEHER AD., *Promoter Targeting RNAs: Unexpected Contributors to the Control of HIV-1 Transcription*, Mol Ther Nucleic Acids. 2015; 4:e222.
- [69] LI E., *Chromatin modification and epigenetic reprogramming in mammalian development*, Nat Rev Genet. 2002; 3:662-73.
- [70] CHRISTMAN JK, WEICH N, SCHOENBRUN B, SCHNEIDERMAN N, ACS G., *Hypomethylation of DNA during differentiation of Friend erythroleukemia cells*, J Cell Biol. 1980; 86:366-70.
- [71] KOIWA T, HAMANO-USAMI A, ISHIDA T, OKAYAMA A, YAMAGUCHI K, KAMIHIRA S, ET AL., *5'-Long Terminal Repeat-Selective CpG Methylation of Latent Human T-Cell Leukemia Virus Type 1 Provirus In Vitro and In Vivo*, Journal of Virology. 2002; 76:9389-97.
- [72] HARBERS K, SCHNIEKE A, STUHLMANN H, JAHNER D, JAENISCH R., *DNA methylation and gene expression: endogenous retroviral genome becomes infectious after molecular cloning*, Proc Natl Acad Sci U S A. 1981; 78:7609-13.
- [73] HEJNAR J, PLACHY J, GERYK J, MACHON O, TREJBALOVA K, GUNTAKA RV, ET AL., *Inhibition of the rous sarcoma virus long terminal repeat-driven transcription by in vitro methylation: different sensitivity in permissive chicken cells versus mammalian cells*, Virology. 1999; 255:171-81.
- [74] LAVIE L, KITOVA M, MALDENER E, MEESE E, MAYER J., *CpG methylation directly regulates transcriptional activity of the human endogenous retrovirus family HERV-K(HML-2)*, J Virol. 2005; 79:876-83.
- [75] CLOUSE KA, POWELL D, WASHINGTON I, POLI G, STREBEL K, FARRAR W, ET AL., *Monokine regulation of human immunodeficiency virus-1 expression in a chronically infected human T cell clone*, J Immunol. 1989; 142:431-8.
- [76] CANNON P, KIM SH, ULICH C, KIM S., *Analysis of Tat function in human immunodeficiency virus type 1-infected low-level-expression cell lines U1 and ACH-2*, J Virol. 1994; 68:1993-7.
- [77] SYMONS J, CHOPRA A, MALATINKOVA E, DE SPIEGELAERE W, LEARY S, COOPER D, ET AL., *HIV integration sites in latently infected cell lines: evidence of ongoing replication*, Retrovirology. 2017; 14:2.
- [78] EDGAR RC, FLYVBJERG H., *Error filtering, pair assembly and error correction for next-generation sequencing reads*, Bioinformatics. 2015; 31:3476-82.

- [79] KUMAKI Y, OKANO M.: *QUMA (QUantification tool for Methylation Analysis)*, <http://quma.cdb.riken.jp> (2019). Accessed 13 Feb 2019.
- [80] EMBL-EBI.: *Clustal Omega*, <https://www.ebi.ac.uk/Tools/msa/clustalo/> (2018). Accessed 7 Apr 2018.
- [81] CHARIF D, LOBRY JR. SEQINR 1.0-2.: *A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and analysis. In: Bastolla U, Porto M, Roman E, Vendruscolo M, editors. Structural approaches to sequence evolution: Molecules, networks, populations*, New York: Springer Verlag; 2007. p. 207-32.
- [82] MAURI M, ELLI T, CAVIGLIA G, UBOLDI G, AZZI M., *RAWGraphs: A Visualisation Platform to Create Open Outputs*, Proceedings of CHIItaly '17, Cagliari, Italy. 2017:1-5.
- [83] KOTECHA N, KRUTZIK PO, IRISH JM., *Web-based analysis and publication of flow cytometry experiments*, Curr Protoc Cytom. 2010; Chapter 10:1-40.
- [84] PERSAUD D, PIERSON T, RUFF C, FINZI D, CHADWICK KR, MARGOLICK JB, ET AL., *A stable latent reservoir for HIV-1 in resting CD4(+) T lymphocytes in infected children*, J Clin Invest. 2000; 105:995-1003.
- [85] TRONO D, VAN LINT C, ROUZIOUX C, VERDIN E, BARRE-SINOUSI F, CHUN TW, ET AL., *HIV persistence and the prospect of long-term drug-free remissions for HIV-infected individuals*, Science. 2010; 329:174-80.
- [86] DUVERGER A, JONES J, MAY J, BIBOLLET-RUCHE F, WAGNER FA, CRON RQ, ET AL., *Determinants of the establishment of human immunodeficiency virus type 1 latency*, J Virol. 2009; 83:3078-93.
- [87] KARN J., *The molecular biology of HIV latency: breaking and restoring the Tat-dependent transcriptional circuit*, Curr Opin HIV AIDS. 2011; 6:4-11.
- [88] XUE J, ZEMPLENI J., *Epigenetic synergies between biotin and folate in the regulation of pro-inflammatory cytokines and repeats*, Scand J Immunol. 2013; 78:419-25.
- [89] ROSCA A, ANTON G, ENE L, IANCU I, TEMEREANCA A, ACHIM CL, ET AL., *Immunoassay and molecular methods to investigate DNA methylation changes in peripheral blood mononuclear cells in HIV infected patients on cART*, J Immunoassay Immunochem. 2017; 38:299-307.
- [90] LAMERE SA, CHAILLON A, HUYNH C, SMITH DM, GIANELLA S., *Challenges in Quantifying Cytosine Methylation in the HIV Provirus*, MBio. 2019; 10.

- [91] VAN LINT C, BOUCHAT S, MARCELLO A., *HIV-1 transcription and latency: an update*, *Retrovirology*. 2013; 10:67.
- [92] BACHMANN N, VON SIEBENTHAL C, VONGRAD V, NEUMANN K, TURK T, BEERENWINKEL N, ET AL., *Determinants of HIV-1 reservoir size and long-term dynamics under suppressive ART* In: *25th CROI Conference on Retroviruses and Opportunistic Infections*, 2018 March 4-7; Boston, Massachusetts. United States of America; 2018. Available from: <http://www.croiconference.org/sessions/determinants-hiv-1-reservoir-size-and-long-term-dynamics-under-suppressive-art>.
- [93] BURNETT JC, MILLER-JENSEN K, SHAH PS, ARKIN AP, SCHAFFER DV., *Control of stochastic gene expression by host factors at the HIV promoter*, *PLoS Pathog*. 2009; 5:e1000260.
- [94] MAHNKE YD, BRODIE TM, SALLUSTO F, ROEDERER M, LUGLI E., *The who's who of T-cell differentiation: human memory T-cell subsets*, *Eur J Immunol*. 2013; 43:2797-809.
- [95] HUANG SH, REN Y, THOMAS AS, CHAN D, MUELLER S, WARD AR, ET AL., *Latent HIV reservoirs exhibit inherent resistance to elimination by CD8+ T cells*, *J Clin Invest*. 2018; 128:876-89.
- [96] FINZI D, HERMANKOVA M, PIERSON T, CARRUTH LM, BUCK C, CHAISSON RE, ET AL., *Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy*, *Science*. 1997; 278:1295-300.
- [97] PION M, JORDAN A, BIANCOTTO A, DEQUIEDT F, GONDOIS-REY F, RONDEAU S, ET AL., *Transcriptional suppression of in vitro-integrated human immunodeficiency virus type 1 does not correlate with proviral DNA methylation*, *J Virol*. 2003; 77:4025-32.
- [98] ALINEJAD-ROKHY H, ANWAR F, WATERS SA, DAVENPORT MP, EBRAHIMI D., *Source of CpG Depletion in the HIV-1 Genome*, *Mol Biol Evol*. 2016; 33:3205-12.
- [99] HERNANDO-HERRAEZ I, GARCIA-PEREZ R, SHARP AJ, MARQUES-BONET T., *DNA Methylation: Insights into Human Evolution*, *PLoS Genet*. 2015; 11:e1005661.
- [100] CHOMONT N, EL-FAR M, ANCUTA P, TRAUTMANN L, PROCOPIO FA, YASSINE-DIAB B, ET AL., *HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation*, *Nat Med*. 2009; 15:893-900.

- [101] VON STOCKENSTROM S, ODEVALL L, LEE E, SINCLAIR E, BACCHETTI P, KILLIAN M, ET AL., *Longitudinal Genetic Characterization Reveals That Cell Proliferation Maintains a Persistent HIV Type 1 DNA Pool During Effective HIV Therapy*, J Infect Dis. 2015; 212:596-607.
- [102] HWANG JK, ALT FW, YEAP LS., *Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination*, Microbiol Spectr. 2015; 3:MDNA3-0037-2014.
- [103] CHEN Z, WANG JH., *Generation and repair of AID-initiated DNA lesions in B lymphocytes*, Front Med. 2014; 8:201-16.
- [104] KNISBACHER BA, GERBER D, LEVANON EY., *DNA Editing by APOBECs: A Genomic Preserver and Transformer*, Trends Genet. 2016; 32:16-28.
- [105] GU S, TIAN Y, CHLENSKI A, SALWEN HR, LU Z, RAJ JU, ET AL., *Valproic acid shows a potent antitumor effect with alteration of DNA methylation in neuroblastoma*, Anticancer Drugs. 2012; 23:1054-66.
- [106] FERNANDEZ G, ZEICHNER SL., *Cell line-dependent variability in HIV activation employing DNMT inhibitors*, Virol J. 2010; 7:266.
- [107] KUCK D, CAULFIELD T, LYKO F, MEDINA-FRANCO JL., *Nanaomycin A selectively inhibits DNMT3B and reactivates silenced tumor suppressor genes in human cancer cells*, Mol Cancer Ther. 2010; 9:3015-23.
- [108] ZHANG Y, FAN M, GENG G, LUI B, HUANG Z, LUO H, ET AL., *A novel HIV-1 encoded microRNA enhances its viral replication by targeting the TATA box region*, Retrovirology. 2014; 11:1-23.
- [109] DOENCH JG., *Am I ready for CRISPR? A user's guide to genetic screens*, Nat Rev Genet. 2018;19:67-80.


E.5. Artículo requisito para la obtención de grado.

RESEARCH

Open Access



Longitudinal variation in human immunodeficiency virus long terminal repeat methylation in individuals on suppressive antiretroviral therapy

César N. Cortés-Rubio¹, Gonzalo Salgado-Montes de Oca¹, Francisco J. Prado-Galbarro², Margarita Matías-Florentino¹, Akio Murakami-Ogasawara¹, Leticia Kuri-Cervantes³, Ana P. Carranco-Arenas¹, Christopher E. Ormsby¹, Ivette K. Cortés-Rubio⁴, Gustavo Reyes-Terán¹ and Santiago Ávila-Ríos^{1*} 

Abstract

Background: Persistence of latent, replication-competent provirus in CD4⁺ T cells of human immunodeficiency virus (HIV)-infected individuals on antiretroviral treatment (ART) is the main obstacle for virus eradication. Methylation of the proviral 5' long terminal repeat (LTR) promoter region has been proposed as a possible mechanism contributing to HIV latency; however, conflicting observations exist regarding its relevance. We assessed 5'-LTR methylation profiles in total CD4⁺ T cells from blood of 12 participants on short-term ART (30 months) followed up for 2 years, and a cross-sectional group of participants with long-term ART (6–15 years), using next generation sequencing. We then looked for associations between specific 5'-LTR methylation patterns and baseline and follow-up clinical characteristics.

Results: 5'-LTR methylation was observed in all participants and behaved dynamically. The number of 5'-LTR variants found per sample ranged from 1 to 13, with median sequencing depth of 16270x (IQR 4107x–46760x). An overall significant 5'-LTR methylation increase was observed at month 42 compared to month 30 (median CpG Methylation Index: 74.7% vs. 0%, $p = 0.025$). This methylation increase was evident in a subset of participants (methylation increase group), while the rest maintained fairly high and constant methylation (constant methylation group). Persons in the methylation increase group were younger, had higher CD4⁺ T cell gain, larger CD8% decrease, and larger CD4/CD8 ratio change after 48 months on ART (all $p < 0.001$). Using principal component analysis, the constant methylation and methylation increase groups showed low evidence of separation along time (factor 2: $p = 0.04$). Variance was largely explained (21%) by age, CD4⁺/CD8⁺ T cell change, and CD4⁺ T cell subpopulation proportions. Persons with long-term ART showed overall high methylation (median CpG Methylation Index: 78%; IQR 71–87%). No differences were observed in residual plasma viral load or proviral load comparing individuals on short-term (both at 30 or 42 months) and long-term ART.

Conclusions: Our study shows evidence that HIV 5'-LTR methylation in total CD4⁺ T cells is dynamic along time and that it can follow different temporal patterns that are associated with a combination of baseline and follow-up clinical characteristics. These observations may account for differences observed between previous contrasting studies.

Keywords: HIV, 5'-LTR, Latency, DNA methylation, CpG, Longitudinal, Proviral load, Residual plasma viral load, Non-canonical methylation

* Correspondence: santiago.avila@cieni.org.mx

¹Centre for Research in Infectious Diseases, National Institute of Respiratory Diseases, Tlalpan 4502, 14080 Mexico City, Mexico

Full list of author information is available at the end of the article



Background

Modern combination antiretroviral treatment (ART) reduces human immunodeficiency virus (HIV) plasma viral load (pVL) to undetectable levels (< 40 HIV RNA copies/mL), achieving a 4–5 log₁₀ reduction during the first year of treatment [1]. However, HIV proviral load is more stable with only 1 log₁₀ reduction in the same treatment period [2]. Persistence of infectious provirus in CD4⁺ T cells (and maybe other cell lineages) in HIV-infected individuals on ART is the main obstacle for virus eradication [3]. The major HIV-1 reservoir is a small pool of latently infected resting memory CD4⁺ T cells [4], which are present in a frequency of 1/10⁶ resting CD4⁺ T cells in persons on suppressive ART [5]. "Kick and kill" HIV cure strategies aim to achieve complete activation of latently-infected cells so that they can be eliminated by diverse immunologic/pharmacologic strategies. To this day, these strategies have shown limited success [6].

Complicating HIV cure efforts, it is estimated that 88% of HIV proviruses are defective, while only 12% constitute intact proviruses [7], which in theory, would have the ability to activate and produce virions. Furthermore, ex vivo reactivation of only a small fraction of latent proviruses (0.12%) has been achieved from resting CD4⁺ T cells of individuals under ART using strategies based on vorinostat, a histone deacetylase inhibitor (HDACi) [6]. In vivo, treatment with panobinostat, another HDACi, did not achieve a cohort-wide reduction in total HIV DNA, integrated HIV DNA, proportion of cells carrying replication competent virus, or number of latently infected cells [8].

A plausible explanation of why certain proviruses fail to reactivate after stimulation may be due to the fact that HIV epigenetic silencing could be the product of a complex and heterogeneous pattern of histone marks [9–12], DNA methylation [13–19], the effect of chromatin remodeling complexes such as Polycomb [12, 20, 21] and BRG1-associated factors (BAF) [22, 23], and non-coding RNAs [24–28], not to mention genetic and microenvironment factors associated with the viral integration site [7]. Indeed, it has been observed that HDACi-based strategies to activate the latent HIV reservoir could paradoxically increase the degree of DNA methylation of the HIV 5' long terminal repeat (LTR) and thus increase the stability of the reservoir [19]. Therefore, more profound knowledge on individual HIV latency mechanisms as well as on their interrelationships is warranted.

DNA methylation patterns are defined early in development and are closely related to both cell differentiation and gene expression; being modified in many disease states, including cancer, autoimmunity, and infectious diseases [29]. There is evidence that DNA from

retroviruses integrated into mammalian genomes can become methylated de novo [30]. Furthermore, CpG methylation of retroviral promoters and enhancers located in the 5'-LTR has been correlated with silencing of several retroviruses such as human T cell leukemia virus type 1 [31], Moloney murine leukemia virus [32], Rous sarcoma virus [33], and human endogenous retroviruses [34]. These observations have inspired several research groups to explore the role of LTR methylation in HIV infection.

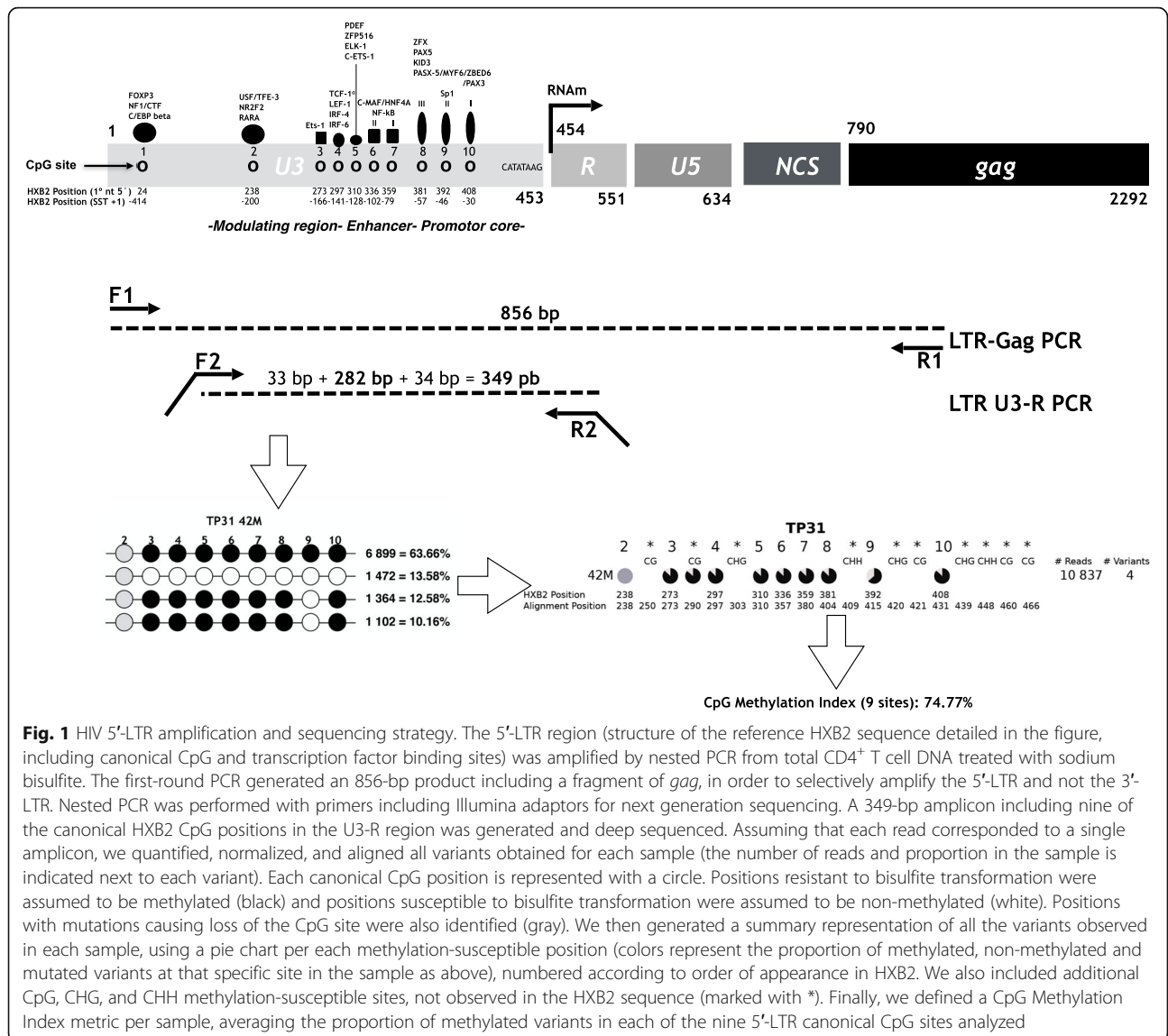
There are contrasting results regarding the role of methylation of the HIV 5'-LTR promoter region on provirus expression in individuals on suppressive ART. Early work suggested that resting memory CD4⁺ T cells of individuals on long-term ART (median 11.5 years) without detectable viremia frequently contained hypermethylated and activation-resistant 5'-LTR, as opposed to hypomethylated 5'-LTR in viremic individuals [13]. Nevertheless, another study challenged these observations, finding that 5'-LTR methylation is rare in resting CD4⁺ T cells of aviremic individuals on short-term ART (median 2.9 years) [17]. More recently, another study found low levels of 5'-LTR DNA methylation in resting CD4⁺ T cells of persons on ART for up to 3 years, but accumulation of 5'-LTR DNA methylation in the latent reservoir after long-term ART and suggested that the transient stimulation of cells harboring latent proviruses may contribute to 5'-LTR methylation [19]. More studies are warranted to determine to what extent 5'-LTR methylation plays a role on HIV latency in order to inform cure strategies.

In this work, we describe longitudinal changes in sequence, variant frequency, and methylation status of the HIV 5'-LTR region, observed in CD4⁺ T cells from blood samples of HIV-infected persons with a short history of suppressive ART compared to persons with long-term suppressive ART. We also studied possible associations between methylation patterns and baseline and follow-up clinical characteristics of the participants.

Results

Longitudinal 5'-LTR methylation profiles from total CD4⁺ T cells of persons with short-term ART

We assessed longitudinal methylation profiles based on nine canonical 5'-LTR CpG sites, present in the reference HXB2 sequence, in total CD4⁺ T cells from blood of 12 participants on short-term ART (24–30 months), using next generation sequencing, as explained in the "Methods", (Fig. 1). For each participant, a 2-year follow up in 3-month intervals was performed (Fig. 2). To have a quantitative methylation measure, we estimated a "CpG Methylation Index," averaging the weighed contribution of each CpG site to the overall 5'-LTR methylation of each sample. Additionally, for each time point, we



determined proviral and residual plasma viral loads, and CD4⁺ T cell subpopulation distributions (defined as shown in Additional file 1: Figure S1). Participants were mostly male (92%), with a median age of 39 years (IQR 33–47) and a median CD4⁺ nadir of 71 cells/mm³ (IQR 7–251) (Table 1). Most persons on short-term ART (11/12) started with a tenofovir + emtricitabine + efavirenz regimen. Further, 7/12 participants remained with their original ART regimen throughout the follow-up period, 3/12 changed the backbone to abacavir + lamivudine, and 3/12 changed the third drug from efavirenz to atazanavir/ritonavir or nevirapine due to drug adverse effects (no drug resistance was observed in any of the participants) (Additional file 2: Table S1). The median number of reads obtained to assess viral variants was 16270 (IQR 4107–46760). The number of variants observed in each time

point ranged from 1 to 13 and no association between the number of variants and the time on ART was observed in any of the participants ($p > 0.05$ in all cases). In general, methylation of the 5'-LTR was common and observed in at least one time point in all participants (Fig. 2). Moreover, 5'-LTR methylation showed a dynamic behavior over time. Overall, two general methylation patterns were observed: methylation remained present throughout the follow-up period (constant methylation group; e.g., participant TP31) or methylation significantly increased at some point (mainly around month 42) along the study period (methylation increase group; e.g., participant TP25) (Fig. 3). Examining individual variants, differences in the degree of methylation between different time points frequently coincided with alternation in predominance of heavily methylated vs. unmethylated variants (Fig. 3 and Additional file 3:

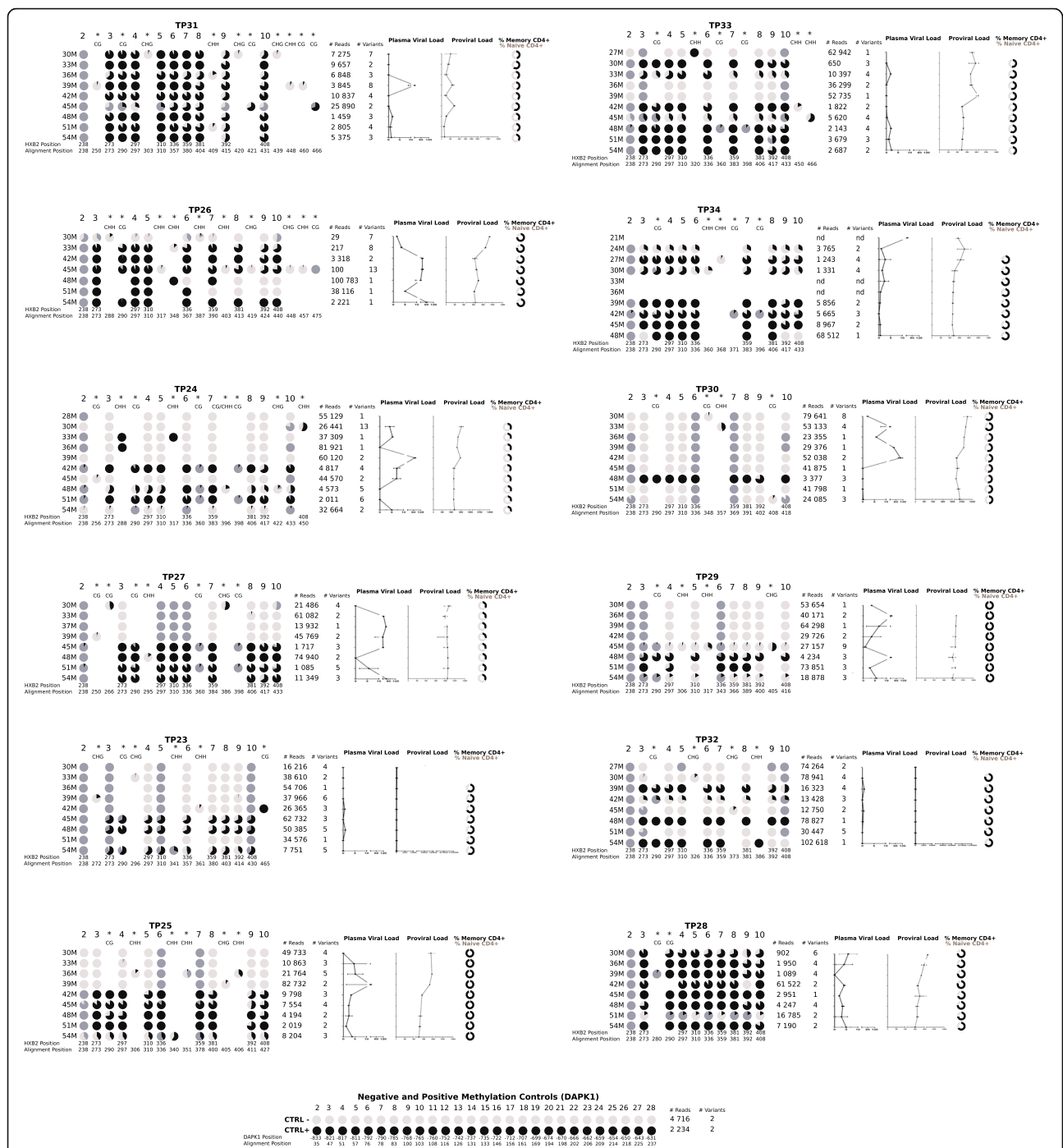


Fig. 2 Overall longitudinal HIV 5'-LTR methylation profiles from total CD4⁺ T cells of 12 participants with short-term antiretroviral treatment. Each group represents a single participant on short-term antiretroviral treatment (24 to 30 months), followed up for 2 years at 3-month intervals. Each line depicts a summary of all variants observed per time point. CpG sites are shown as pie charts indicating the proportion of methylated (black), unmethylated (light gray) or mutated (dark gray) variants in the sample. Canonical HXB2 CpG positions are numbered. Additional CpG, CHG, and CHH methylation-susceptible sites, not observed in the HXB2 sequence, are also included (marked with *). The number of reads and the total number of variants obtained per time point, after filtering and aligning, are shown. Additionally, measurements of residual plasma viral load, proviral load and proportion of memory (CD3⁺/CD4⁺/CD28⁺/CD95⁺; black) and naïve (CD3⁺/CD4⁺/CD45RO⁻/CCR7⁺/CD28⁺/CD95⁻; light gray) CD4⁺ T cells per time point are shown

Table 1 Baseline characteristics of the short-term constant methylation and methylation increase groups compared to the long-term ART group

	Short-Term ART group Combined <i>n</i> = 12	<i>p</i> value ^a	Short-Term ART group Constant methylation <i>n</i> = 5	<i>p</i> value ^b	Short-Term ART group Methylation increase <i>n</i> = 7	<i>p</i> value ^c	<i>p</i> value ^d	Long-term ART group <i>n</i> = 10
Sex (Male/Female)	11:1		5:0		6:1			9:1
Age (years)	39 (33–47)	<i>0.013</i>	41 (38–45)	0.910	34 (31–48)	<i>0.004</i>	< <i>0.001</i>	45 (39–56)
Nadir CD4 ⁺ T cell count (cells/mm ³)	71 (7–251)	0.294	128 (12–264)	0.843	13 (4–212)	0.106	0.075	96 (47–234)
Pre-ART pVL (RNA copies/mL)	194,077 (67,293–285,046)	0.077	144,952 (90,736–243,202)	0.174	248,219 (59,479–285,552)	<i>0.048</i>	0.205	47,800 (7581–154,809)
Delta CD4 ⁺ T cell count at 48 months (cells/mm ³)	388 (174–502)	< <i>0.001</i>	330 (130–412)	0.843	455 (214–580)	<i>0.029</i>	< <i>0.001</i>	291 (193–341)
Delta CD4 % at 48 months	13 (11–20)	0.163	12 (10–15)	0.190	13 (12–32)	0.023	<i>0.041</i>	11 (6–14.5)
Delta CD8 ⁺ T cell count at 48 months (cells/mm ³)	172 (–104, 315)	0.053	102 (–109, 246)	0.751	202 (–89, 325)	0.105	<i>0.023</i>	–85 (–405, 331)
Delta CD8 % at 48 months	–15 (–24, –5)	0.268	–13 (–17, 3)	<i>0.015</i>	–24 (–25, –10)	0.162	< <i>0.001</i>	–32 (–46, –9)
Delta CD4/CD8 at 48 months	0.40 (0.24–0.80)	0.058	0.33 (0.23–0.35)	0.668	0.53 (0.28–0.84)	<i>0.036</i>	< <i>0.001</i>	0.35 (0.25–0.55)
Days to undetectable pVL	183 (101–368)	0.287	186 (99–384)	0.551	180 (107–344)	0.771	0.205	148 (95–379)
Third drug in current ART regimen (NNRTIs / PIs)	83% / 17%	0.624	80% / 20%	1	86% / 14%	0.603	1	70% / 30%
Number of ART regimen changes	0 (0–1)	<i>0.027</i>	0 (0–0)	<i>0.003</i>	1 (0–1)	0.102	<i>0.008</i>	1 (0.8–2)

For each variable, median (interquartile range) is shown

ART antiretroviral therapy, pVL plasma viral load, NNRTI non-nucleoside reverse transcriptase inhibitors, PI protease inhibitors. Significant *p* values (*p* < 0.05) are shown in italic

^aShort-term ART group vs. long-term ART group

^bConstant methylation group vs. long-term ART group

^cMethylation increase group vs. long-term ART group

^dConstant methylation group vs. methylation increase group

Figure S2). In general, variants lacking methylated sites showed loss of one or more CpG sites by mutation, compared to HXB2 (e.g., Fig. 3, TP25 36M). Variants showing eight methylated canonical CpG sites (the CpG2 site was usually mutated or unmethylated) were common, observed in all participants in at least one time point.

Overall, the number of CpG sites in the studied 5'-LTR fragment ranged from 1 to 10 across the study participants and different time points. Over time, we observed mutations leading to both loss and gain of canonical CpG sites in predominant variants. Additionally, all participants gained potentially methylation-susceptible CpG, CHG, or CHH sites, originally absent in the reference HXB2 in at least one time point (Fig. 2).

Taken together, our results show that HIV 5'-LTR methylation is variable along time in total CD4⁺ T cells from blood of individuals under short-term ART and that methylation can be observed from early stages of ART initiation. Furthermore, at least two different temporal methylation variation patterns were observed among the participants.

Temporal patterns of 5'-LTR methylation

Considering all participants with short-term ART, we observed significant 5'-LTR methylation increase at month 42 with respect to month 30 (median CpG Methylation Index: 74.7% vs. 0%, *p* = 0.025) (Table 2, Fig. 4a). This methylation increase was evident in a subset of the participants (methylation increase group), while the rest maintained fairly high and constant methylation along the follow-up period (constant methylation group) (Fig. 2, 3, and 4b). Indeed, the methylation increase group was significantly different from the constant methylation group in that the starting 5'-LTR CpG methylation level (at 30 months on ART) in the former group was significantly lower compared to the latter (median CpG Methylation Index: 0 vs. 65, *p* = 0.007). This difference was also observed for methylation in non-canonical, additional CpG sites (median Additional CpG Methylation Index: 0 vs. 21.7 *p* = 0.030) (Table 3). Interestingly, we observed an overall decrease in proviral load at month 42 compared to month 30 (median proviral load: 3638 vs. 14146 copies/million CD4⁺ T cells, *p* = 0.005) (Table 2, Fig. 4c). Residual pVL was stable in

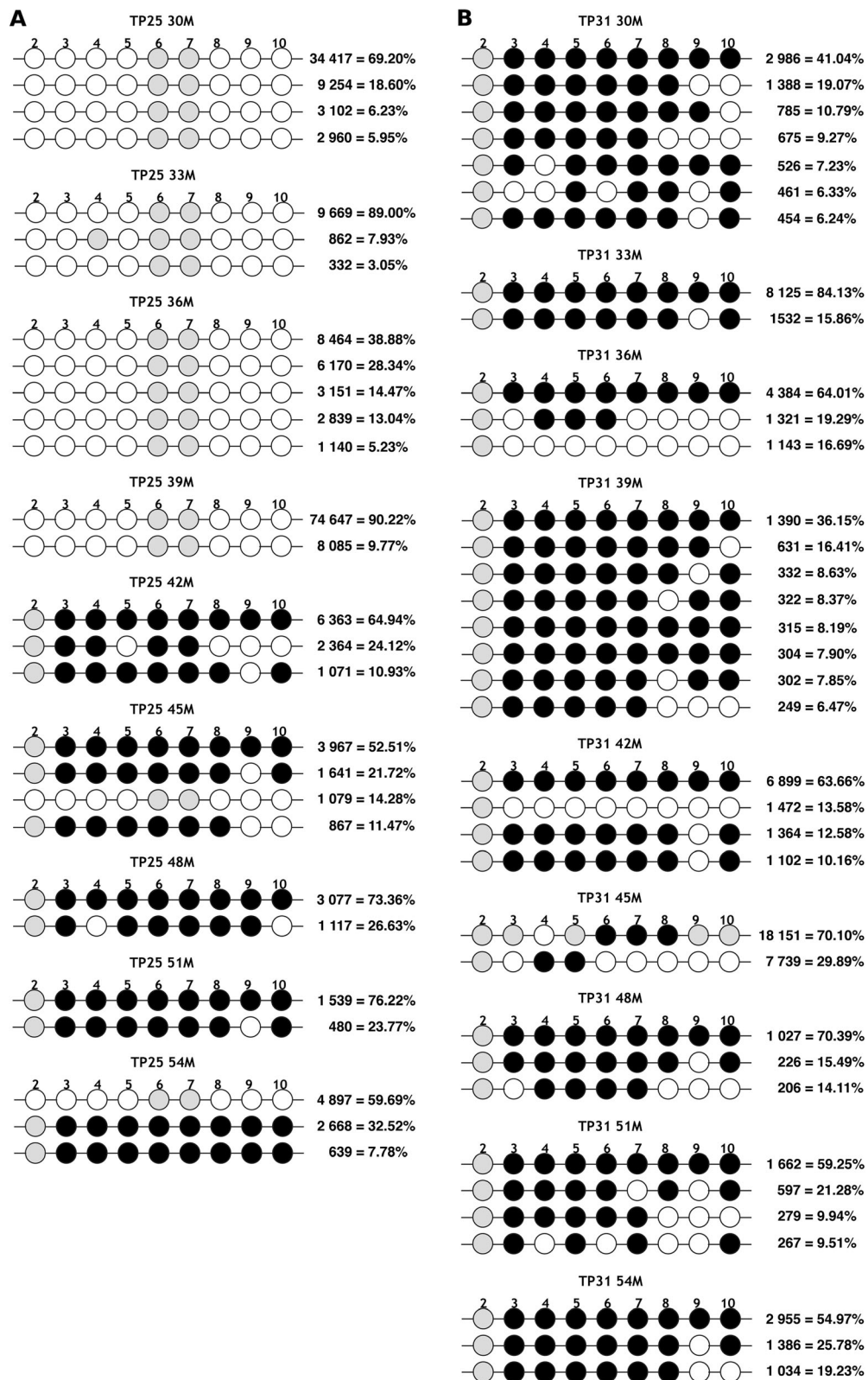


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Detailed time trends of HIV 5'-LTR methylation in two model participants. Detailed HIV variant composition per time point (30, 33, 36, 39, 42, 45, 48, 51, and 54 months) is shown for participants **a** TP25, representative of the methylation increase group, and **b** TP31, representative of the constant methylation group. Each horizontal line represents a variant. The number of reads representing that variant and its proportion in the sample are shown on the right. Each circle represents a canonical CpG site, numbered according to the HXB2 reference: black—methylated, white—non-methylated, gray—mutated

most participants during follow-up, with occasional blips: TP30 39M (43 copies/mL), TP30 42M (99 copies/mL), TP24 39M (71 copies/mL), TP27 54M (42 copies/mL), and TP26 54M (60 copies/mL) (Fig. 2). No significant trends in time were observed for residual pVL. However, the constant methylation group showed overall less dispersion than the methylation increase group (except at month 54 when participant TP26 had a blip) (Fig. 4d).

A group of participants with a long history on ART was included to explore methylation associations with residual pVL and proviral load in the long term. This cross-sectional group included persons with median time on ART of 12 years (IQR 7–15), mostly male (90%), with median age of 45 years (IQR 39–56), and median nadir CD4⁺ T cell count of 96 cells/mm³ (IQR 47–234) (Table 1). Participants on long-term ART showed significantly more ART regimen changes than persons on short-term ART (8/10 participants changed ART regimen at least once), mostly for optimization or switches to newer drugs. The current

ART regimen of participants in this group was tenofovir + emtricitabine + efavirenz for 7/10 and tenofovir + emtricitabine and a protease inhibitor for 3/10 (Additional file 2: Table S1). Persons with a long history of suppressive ART showed overall high degree of methylation with a median CpG Methylation Index of 78% (IQR 71–87%) (Table 2). Variants with mutated CpG sites were rare in this group and presence of new methylation-susceptible sites was infrequent, compared to individuals followed up longitudinally (Additional file 4: Figure S3).

Comparing the longitudinal follow-up group of participants at 30 months on ART with the cross-sectional group on long-term ART, we observed a significant difference in overall methylation (median CpG Methylation Index: 0% vs. 78%, $p = 0.005$), but this difference was lost at 42 months on ART (median CpG Methylation Index: 75% vs. 78%, $p = 0.273$) (Table 2). More specifically, the difference in early methylation (30 months vs. long-term ART) was observed in persons in the methylation increase

Table 2 Early changes in methylation patterns and associated clinical variables

	Short-Term ART group (30 months)	<i>p</i> value ^a	Short-term ART group (42 months)	<i>p</i> value ^b	<i>p</i> value ^c	Long-term ART group
CpG methylation Index	0 (0–58.4)	0.005	74.7 (0–85.3)	0.273	0.025	77.9 (70.8–86.7)
Additional CpG methylation Index ^d	0.8 (0–21.0)	0.155	27.0 (4.3–45.1)	0.163	0.074	20.0 (13.8–20.0)
Non-CpG methylation Index ^e	1.3 (0–9.5)	0.388	0 (0–0)	0.332	0.093	0 (0–2.2)
Residual pVL (RNA copies/mL)	1.4 (0–3.6)	0.737	3.4 (0–7.1)	0.305	0.176	0.4 (0–2.9)
Proviral load (DNA copies/million CD4 ⁺ T cells)	14146 (517–29277)	0.526	3638 (228–10216)	0.260	0.005	9780 (4 057–13716)
Naive CD4 ⁺ T cells (%)	25.2 (20.6–52.6)	0.944	29.1 (18.3–49.1)	0.888	0.114	39.4 (11.2–45.4)
Memory CD4 ⁺ T cells (%)	74.8 (47.4–79.4)	0.944	70.9 (50.9–81.7)	0.888	0.114	60.6 (54.6–88.8)
T _{SCM} (%)	5.6 (3.0–6.9)	0.049	4.2 (2.9–7.1)	0.035	0.508	9.4 (5.7–15.9)
T _{CM} (%)	23.2 (15.3–31.2)	0.673	21.7 (12.7–33.0)	0.439	0.169	23.8 (19.5–40.9)
T _{EM} (%)	12.5 (2.1–20.4)	0.398	10.1 (3.2–12.5)	0.205	0.333	6.3 (4.8–9.3)
T _{TM} (%)	15.0 (11.3–25.4)	0.091	13.8 (11.4–31.5)	0.049	0.799	8.3 (4.5–16.6)
T _{NEW} (%)	1.1 (0.6–1.5)	0.003	2.5 (1.6–3.6)	0.526	0.005	2.0 (1.5–3.9)
T _{TE} (%)	2.3 (0.3–11.1)	0.181	5.0 (1.2–9.8)	0.944	0.285	4.8 (2.7–7.8)

In all cases, medians (interquartile ranges) are shown. Significant *p* values ($p < 0.05$) are shown in italic

ART antiretroviral therapy, pVL plasma viral load, T_{SCM} stem cell memory T cells, T_{CM} central memory T cells, T_{EM} effector memory T cells, T_{TM} transitional memory T cells, T_{NEW} T new, T_{TE} terminally differentiated effector T cells

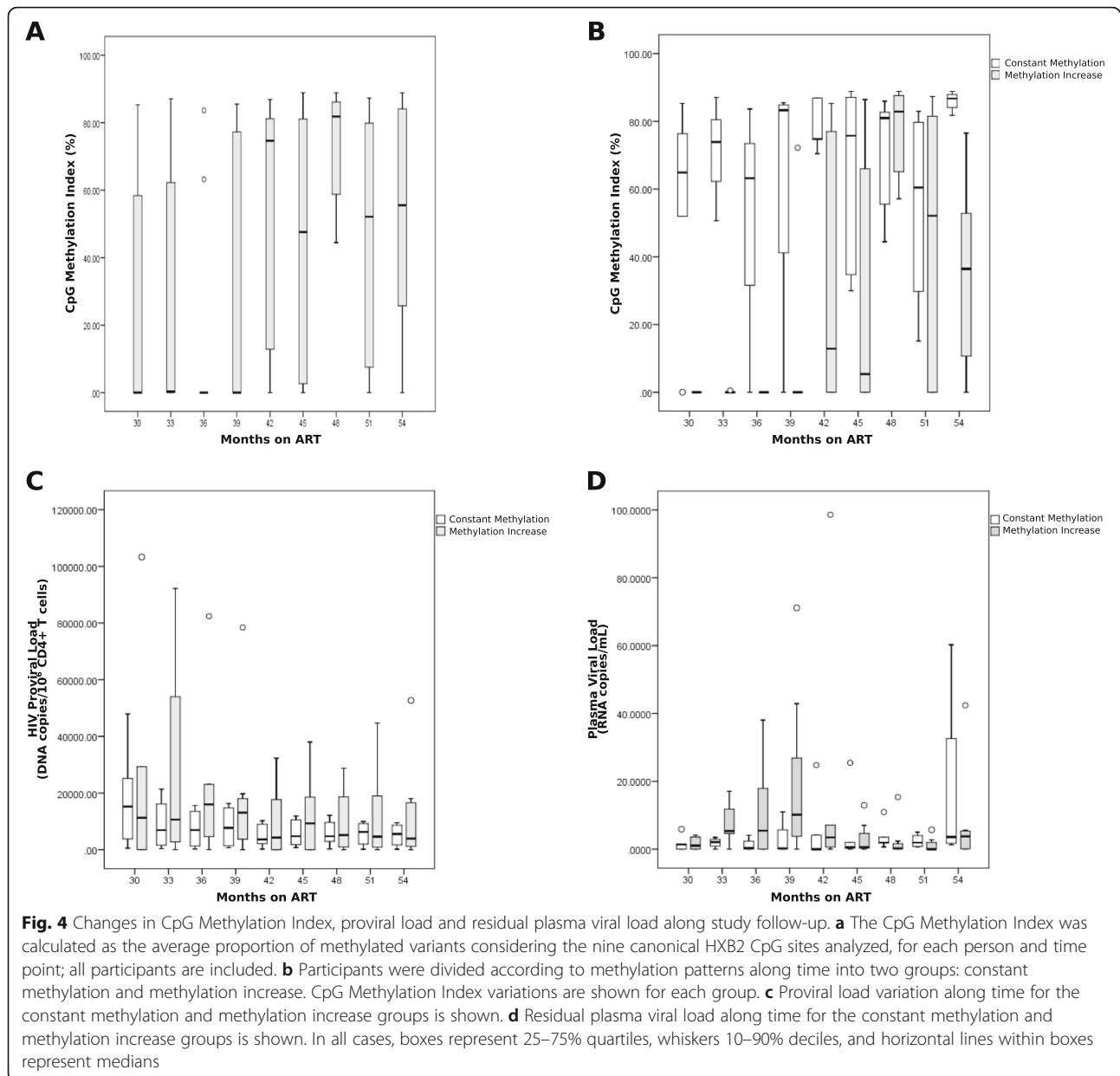
^aShort-term ART group (30 months) vs. long-term ART group

^bShort-term ART group (42 months) vs. long-term ART group

^cShort-term ART group (30 months) vs. short-term ART group (42 months)

^dConsiders methylation of CpG positions not present in the reference HXB2

^eConsiders methylation of non-CpG positions only



group ($p = 0.001$), but not in the constant methylation group ($p = 0.244$) (Table 3). No differences were observed in residual pVL or proviral load, comparing individuals on short-term ART (both at 30 or 42 months of follow-up) and long-term ART (Tables 2 and 3).

In order to be more comprehensive in the methylation analysis, we studied non-canonical methylation patterns along the follow-up period, establishing a measure of methylation on CpG sites different to the canonical ones, present in the HXB2 reference (additional CpG methylation index), and methylation in sites other than CpG (non-CpG methylation index). Similar to our observations for canonical CpG methylation, a significant

difference in additional CpG methylation was observed in persons in the methylation increase group ($p = 0.003$), but not in the constant methylation group (30 months; $p = 0.282$) when compared to persons on long-term ART (Table 3). No differences were observed in non-CpG methylation comparing the short-term (both at 30 and 42 months of follow-up) and long-term ART groups (Tables 2 and 3).

Considering CD4⁺ T cell subpopulation proportions, three general patterns were seen among participants: (1) individuals with higher proportion of naive CD4⁺ T cells (17%): TP24 and TP27; (2) individuals with higher proportion of memory CD4⁺ T cells (58%): TP26, TP23,

Table 3 Differential variation of clinical variables along time between the constant methylation and the methylation increase groups

	Short-term ART group (30 months)				Short-term ART group (42 months)				Long-term ART group
	Constant methylation group	<i>p</i> value ^a	Methylation increase group	<i>p</i> value ^b	Constant methylation group	<i>p</i> value ^a	Methylation increase group	<i>p</i> value ^b	
CpG methylation Index	64.9 (52.0–76.3)	0.244	0 (0–0)	<i>0.001</i>	74.8 (74.7–86.9)	0.902	12.9 (0–77.0)	0.071	77.9 (70.8–86.7)
Additional CpG methylation Index ^c	21.7 (9.4–34.9)	0.282	0 (0–1.7)	<i>0.003</i>	27 (8.6–40.6)	0.282	27.0 (0–50.0)	0.219	20.0 (13.8–20.0)
Non-CpG methylation Index ^d	2.6 (0–8.3)	0.352	0 (0–11.2)	0.585	0 (0–3)	0.560	0 (0–0)	0.356	0 (0–2.2)
Residual pVL (RNA copies/ml)	1.4 (0–1.5)	0.699	1.0 (0–3.6)	0.862	0 (0–4.2)	0.842	3.5 (0.6–7.1)	0.164	0.4 (0–2.9)
Proviral load (DNA copies/million CD4 ⁺ T cells)	15241 (3800–25166)	0.462	11250 (22–29277)	0.745	3638 (2155–9012)	0.221	4264 (10–17670)	0.515	9780 (4057–13716)
Naive CD4 ⁺ T cells (%)	25.2 (23.9–45.7)	0.713	28.0 (4.2–64.1)	0.828	29.1 (27.7–49.1)	0.713	26.3 (4.8–47.0)	0.588	39.4 (11.2–45.4)
Memory CD4 ⁺ T cells (%)	74.8 (54.3–76.2)	0.713	72.1 (35.9–95.8)	0.828	70.9 (50.9–72.3)	0.713	73.7 (53.0–95.2)	0.588	60.6 (54.6–88.8)
T _{SCM} (%)	6.4 (4.7–6.6)	0.221	4.5 (1.0–6.9)	0.051	4.2 (2.9–7.4)	0.111	4.7 (3.4–6.2)	0.065	9.4 (5.7–15.9)
T _{CM} (%)	23.2 (23.2–28.8)	0.713	22.3 (15.3–35.1)	0.745	21.7 (12.7–33.0)	0.540	21.9 (15.5–32.5)	0.515	23.8 (19.5–40.9)
T _{EM} (%)	12.5 (5.4–17.2)	0.270	9.2 (2.1–26.5)	0.745	9.0 (7.6–10.1)	0.462	12.1 (3.2–13.8)	0.193	6.3 (4.8–9.3)
T _{TM} (%)	13.3 (12.6–15.1)	0.221	15.2 (11.3–25.4)	0.129	11.6 (11.4–19.5)	0.142	17.4 (11.7–31.5)	0.083	8.3 (4.5–16.6)
T _{NEW} (%)	1.4 (0.6–1.56)	0.075	1.1 (0.6–1.2)	<i>0.002</i>	2.5 (1.6–4.1)	0.624	2.7 (2.3–3.3)	0.588	2.0 (1.5–3.9)
T _{TE} (%)	4.0 (0.3–11.8)	0.713	1.5 (0.3–4.5)	0.083	3.9 (1.2–9.6)	0.806	6.3 (2.2–9.8)	0.914	4.8 (2.7–7.8)

In all cases, medians (interquartile ranges) are shown. Significant *p* values (*p* < 0.05) are shown in italic
 ART antiretroviral therapy, pVL plasma viral load, T_{SCM} stem cell memory T cells, T_{CM} central memory T cells, T_{EM} effector memory T cells, T_{TM} transitional memory T cells, T_{NEW} T new, T_{TE} terminally differentiated effector T cells

^aConstant methylation group vs. long-term ART group

^bMethylation increase group vs. long-term ART group

^cConsiders methylation of CpG positions not present in the reference HXB2

^dConsiders methylation of non-CpG positions only.

TP25, TP34, TP29, TP32, and TP28; and (3) individuals with equal proportion of naive and memory CD4⁺ T cells (25%): TP31, TP33, and TP30; however, these groups did not show specific associations with methylation patterns (Fig. 2). Although no significant differences in the proportion of naive and overall memory CD4⁺ T cells were observed between individuals with short-term and long-term ART, considering memory T cell subpopulations, a lower prevalence of the T_{SCM} and T_{new} was observed in individuals with 30 months on ART belonging to the methylation increase group (*p* = 0.051 and 0.002, respectively) compared to the long-term ART group, but not in the constant methylation group (*p* = 0.221 and 0.075, respectively) (Table 3), suggesting early differences in CD4⁺ T cell subpopulation composition between persons in the constant methylation and the methylation increase groups.

Baseline and follow-up characteristics of persons with distinct 5'-LTR methylation patterns

Next, we looked for differences in the baseline characteristics of persons belonging to the constant methylation and the methylation increase groups (Table 1). Persons in the methylation increase group were significantly younger (median age: 34 vs. 41 years, *p* < 0.001), had higher CD4⁺ count gain at 48 months after ART initiation (delta CD4⁺ T cell count: 455 vs. 330, *p* < 0.001; delta CD4 %: 13 vs. 12, *p* = 0.041), higher CD8⁺ count change at 48 months (delta CD8 count: 202 vs. 102, *p* = 0.023; delta CD8 %: -24 vs. -13, *p* < 0.001), and larger CD4/CD8 ratio change at 48 months (delta CD4/CD8: 0.53 vs. 0.33, *p* < 0.001) (Table 1). These results suggest that different behaviors on 5'-LTR methylation could be associated with baseline clinical characteristics of persons starting ART.

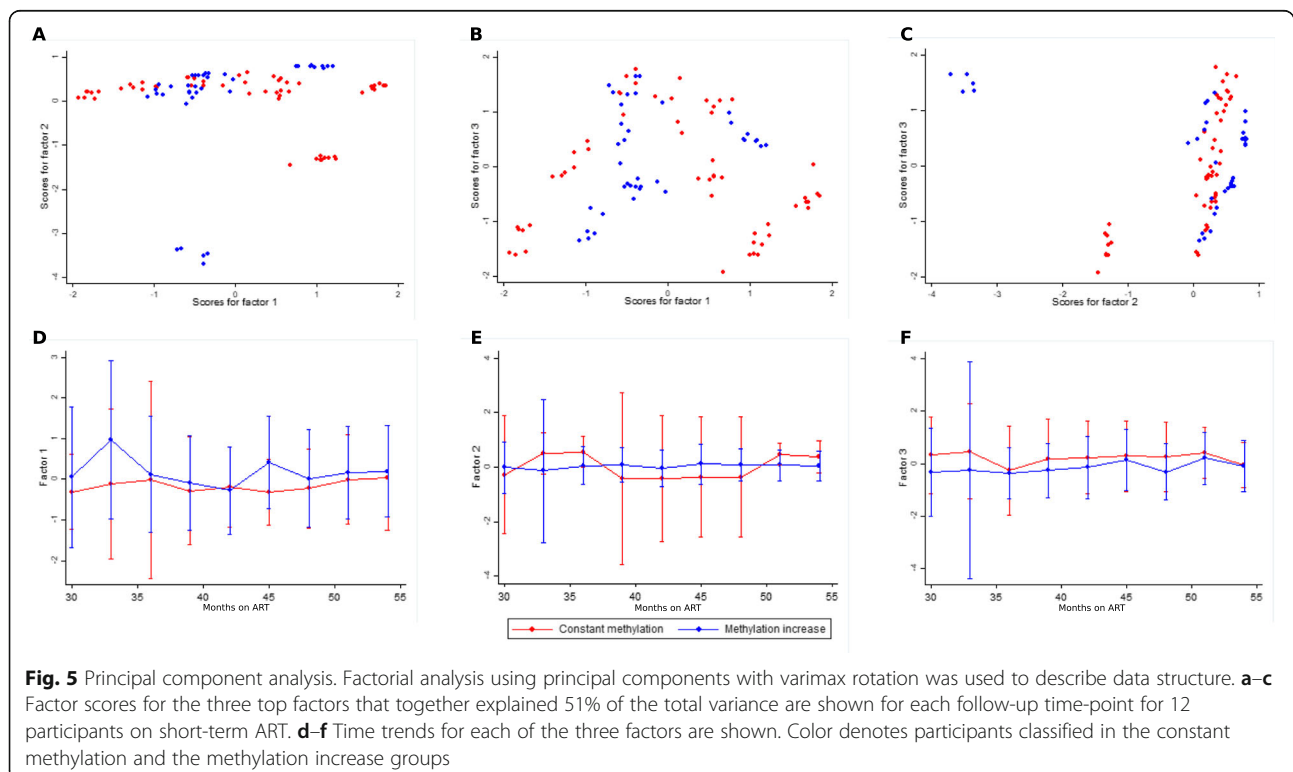
We applied factorial analysis using principal components (PCA) with varimax rotation in order to explain the structure of the data in an unbiased way. We included baseline and follow-up clinical variables (Fig. 5, see “Methods”). Altogether, the first three factors explained 51% of the total variance. Factor 1 explained 21% of the total variance and encompassed ten variables (age, CD4⁺ T cell number and % gain, CD8⁺ T cell number and % change, CD4/CD8 change, % CD4⁺ memory, namory, CM , and T_{TM} cells), factor 2 16% and encompassed seven variables (CD4⁺ T cell number gain, CD8⁺ T cell number change, time to undetectable pVL, time on ART: 30- vs. 54-months follow-up, third drug class in ART regimen: NNRTI vs. PI), and factor 3 13% and encompassed five variables (nadir CD4⁺ T cell count, CD8⁺ T cell count change, % T_{EM} , % T_{TE} , and number of ART regimen changes). Of note, CpG methylation index and pre-ART pVL were only included up to factor 4; ART regimen backbone in factor 5; and residual viral load, proviral load, and presence of blips were included up to factor 6. The constant methylation and methylation increase groups showed low evidence of significant separation along time (Wilcoxon test for factor 2: $p = 0.04$; factors 1 and 3: $p > 0.05$) (Fig. 5).

In all, these observations suggest that temporal differences in 5'-LTR methylation patterns could be associated with a combination of baseline and follow-up clinical characteristics, mainly age, CD4⁺ T cell gain,

CD8⁺ T cell decrease, time on ART, class of third drug in ART regimen, and CD4 T cell subpopulation composition.

Discussion

Many HIV cure efforts have focused on developing strategies to reverse viral latency that, in theory, could allow viral gene expression leading to virus recognition and eradication. The contribution of methylation of the HIV 5'-LTR promoter region to HIV latency remains elusive, with previous studies showing contrasting results. Our study shows for the first time that HIV 5'-LTR methylation is dynamic along time, with different methylation variation patterns possibly associated with baseline clinical characteristics of persons starting ART. These observations could explain some of the differences between previous studies showing contrasting results. Strengths and novel aspects of this study include implementation of a next generation sequencing-based methodology, able to extensively analyze the composition of 5'-LTR variants and their methylation status with high sensitivity, in contrast with previous studies using traditional clonal Sanger sequencing. Additionally, we present for the first time a longitudinal assessment of 5'-LTR methylation changes on CD4⁺ T cells of a cohort of individuals on short-term ART, and further compared these results with a cross-sectional group of persons with long-term ART. Our study shows that HIV 5'-LTR



methylation in total blood CD4⁺ T cells is dynamic along time, and that at least two different methylation patterns exist that show association with both baseline and follow-up clinical characteristics of individuals, mainly age, CD4⁺ T cell gain, CD8⁺ T cell decrease, and CD4⁺ T cell subpopulation proportions.

It has been previously suggested that latent HIV resides mainly in resting CD4⁺ T cells [35, 36] and that DNA methylation may be a late event that reinforces silencing of the latent provirus [37, 38]. Indeed, CpG hypermethylation within the HIV 5'-LTR has been associated with silencing of the HIV promoter in latently infected Jurkat cells and resting CD4⁺ T cells from aviremic patients under ART [13, 14]. However, other studies analyzing the latent viral reservoir of aviremic patients on ART have found very low or no methylation on the HIV 5'-LTR in resting memory CD4⁺ T cells or peripheral blood mononuclear cells (PBMCs), questioning the importance of methylation as a maintenance mechanism of the latent reservoir [15, 17]. There are several factors that could account for contrasting observations on the role of 5'-LTR methylation on HIV latency, including the length of HIV infection, ART regimen used, time on ART, cell type studied, CD4⁺ T cell subpopulation composition, rate of replacement of latently infected CD4⁺ T cells, proportion of defective proviruses, and HIV integration sites [16, 17, 19]. Our study identified associations between several of these variables and distinct longitudinal methylation change patterns, including age, CD4⁺ T cell number and % gain, CD8⁺ T cell number and % change, proportions of CD4⁺ T cell subpopulations, third drug class in the current ART regimen, and time to undetectable viral load. Moreover, as recently suggested [39], other rather methodological factors could also account for differences and limit conclusions on methylation analyses including the use of cell lines versus primary CD4⁺ T cells in vitro, frozen versus fresh samples, the type of bisulfite treatment, sampling bias, polymerase chain reaction (PCR) amplification strategy and primers used, and sequencing technique (Sanger versus next generation sequencing). Using PCA, our study identified CD4⁺ T cell gain (at 48 months after ART initiation), CD8⁺ T cell change, CD4/CD8 rate change, and proportion of naïve, TCM, and TTM cells as the most important variables, accounting for 21% of the total variance in the study participants. Interestingly, time on ART, time to undetectable pVL, and third drug class in the ART regimen were also identified as important variables explaining total variance in the second eigenvector. Since DNA methylation could be a late event that strengthens the silencing of already latent viruses, rather than driving latency from the beginning, infection time and source cell type could be very important factors in methylation analyses. Indeed,

CpG methylation may be a more common latency mechanism in memory CD4⁺ T cells that are more prone to proliferation than in resting CD4⁺ T cells [40]. Late selection of latent proviruses in the memory cell subpopulation could result in the accumulation of proviruses with methylated promoters as compared to unmethylated proviruses found in the resting cell subpopulation [40]. Recently, Trejbalova et al. showed a positive correlation between time on ART and the degree of CpG methylation in the HIV 5'-LTR in resting CD4⁺ T cells [19]. However, it is hard to discriminate if the correlation was due to time on ART only, time of HIV infection before ART initiation, or both. Our study contributes with additional information to the field, showing that 5'-LTR methylation appears early after ART start in at least two different patterns that can be associated with the baseline clinical characteristics of each person. Interestingly, while a group of participants maintained fairly high and constant methylation levels across the follow-up period, a significant increase in methylation burden occurred in a subset of participants (58%) at around 42 months (3.5 years) after ART initiation. This time point coincided with an overall significant decrease in proviral load, an observation that has been reported before in a large cohort [41]. The methylation increase subgroup had overall different baseline characteristics compared to the constant methylation subgroup as shown by PCA, and was composed mainly of younger individuals presenting late to clinical care (nadir CD4⁺ T cell count: 13 cells/mm³), and showing larger CD4⁺ cell gains. These observations could help to explain conflicting observations of previous studies and could be important for cure strategies.

Our study is unique in the sense that it used a highly effective bisulfite conversion system with low DNA degradation, a nested PCR strategy to selectively amplify and analyze bisulfite-converted 5'-LTR, and a highly sensitive next generation sequencing strategy that detects low frequency variants more efficiently than clonal Sanger sequencing strategies. As 5'-LTR methylation analyses were made on total CD4⁺ T cells, we accounted for the CD4⁺ T cell subpopulation composition in each sample as a possible confounder. Finally, our study design was longitudinal, which made it possible to study the dynamics of 5'-LTR methylation over time in the same individual.

With our experimental design, we were able to identify between 1 and 13 different variants (median: 3) per time point with median coverage of 16270 reads, and additionally estimate the prevalence of each variant within the sample, allowing to weight the contribution of each variant to the overall proviral landscape and methylation burden. This greatly exceeds the average number of clones analyzed per sample in other studies using Sanger-based sequencing strategies. Our controls showed a highly

efficient bisulfite-conversion process (100% conversion), which combined with our primer design (that considered bisulfite-converted cytosines) (Additional file 5: Table S2) and filtering strategy, selected for PCR products derived from bisulfite-converted templates.

Some weaknesses of our study are worth mentioning: as we exclusively sequenced the 5'-LTR and not the complete HIV genome, we were not able to assess the role of 5'-LTR methylation in replication-competent proviruses only, nor the role of methylation in other parts of the provirus or the integration site. Moreover, our analyses were made in total CD4⁺ T cells due to limitations in cell number and the possibility of amplification of the LTR region in bisulfite-treated DNA, and thus, it was not possible to assess methylation in specific cell subpopulations or in resting CD4⁺ T cells. Also, as proviral loads were variable among participants and time points, even when the same amount of DNA was used for bisulfite conversion and as input for PCR amplification, the number of provirus templates analyzed in different samples was variable. Finally, as previously suggested [39], proviral load and residual pVL may not be the most appropriate outputs to assess the role of 5'-LTR methylation in latency as defective and replication-competent proviruses are both studied and cannot be differentiated in our method. Nevertheless, they are good proxy variables, given our high sampling depth, which significantly increases sensitivity to include intact proviruses. Indeed, considering our median proviral load (8129 copies/million CD4⁺ cells), when starting with 900 ng of DNA for bisulfite conversion, we would be working with 136,364 human genomes, from which 1109 would contain HIV DNA. If we consider that 88% of proviruses are defective [7], we would have a total of 133 intact proviruses in the sample. As half of the volume of bisulfite-converted DNA was used for PCR, we would be analyzing 67 intact proviruses (without considering inherent DNA degradation due to bisulfite treatment). Furthermore, this study did not contemplate large blood draws or collection of leukapheresis samples, necessary to perform viral outgrowth assays [5, 42], given the low proportion of cells expected to contain latent HIV in blood in vivo (1 in 10⁶ resting CD4⁺ T cells) [5, 43] in comparison with in vitro latency models, and thus it was not possible to perform virus reactivation analyses.

Further insight on the observed alternation between unmethylated and methylated variants is warranted to understand the overall role of 5'-LTR methylation on HIV latency. This could be a complex phenomenon with several possible mechanisms in play. Regarding absence of or low-level 5'-LTR methylation in cells with latent infection or low-level HIV expression, important mechanisms could include (1) prevention of

de novo methylation [18, 44], (2) high rate of demethylation [44], or (3) deamination of 5mCpG or CpG (detection of a TpG site may be the result of deamination of either 5mCpG or unmethylated CpG, both of which would produce UpG, which would in turn be observed as TpG after PCR amplification). Indeed, it has been suggested that methylation could be the main mechanism responsible for depletion of CpG sites in the HIV genome [45]. This type of mutation could either originate replication-defective proviruses or remove transcriptional silencing [46]. Regarding enrichment of methylated variants, some possible mechanisms include (1) accumulation of proviruses with methylated promoters in memory cells through homeostatic proliferation [47, 48], (2) change in the proportion of the cellular subpopulations containing proviral DNA [18], (3) de novo methylation [19, 49], and (4) low-level cellular activation.

Conclusions

We show strong evidence that HIV 5'-LTR methylation can vary significantly along time in proviruses from circulating CD4⁺ T cells and that methylation is observed from early stages of ART initiation. We also show that 5'-LTR methylation can follow different temporal patterns that are associated with baseline clinical characteristics, and which may account for some differences observed between previous contrasting studies. In particular, different 5'-LTR methylation patterns can be associated with age, CD4⁺ T cell gain and CD8⁺ T cell change after ART initiation, changes in proportions of circulating T cell subpopulations (mainly T_{CM}, T_{TM}, and naïve), time on ART, time to undetectable pVL, and third drug class in the ART regimen.

Further studies in specific cell subpopulations and in tissues, the use of larger cohorts with longer follow-up times and stricter control of confounding variables (e.g., infection time) and the use of novel sequencing techniques, allowing assessment of methylation along the complete provirus and in the integration site are necessary to assess the role of viral LTR methylation in proviral silencing and the extent to which it might be clinically relevant.

Methods

Ethics statement

This study was approved by the Institutional Review Board of the National Institute of Respiratory Diseases (INER) (registry: DHHS IORG0003948; CONBIOETICA-09-CEI-003-20160427) in Mexico City (project code: B26-09), and was conducted according to the principles of the Declaration of Helsinki. All participants gave written informed consent before blood sample donation.

Patients

This study included two parallel designs, longitudinal and cross-sectional. Longitudinal blood samples from 12 HIV-infected persons with a short history of suppressive ART (21–30 months on ART and viral load 40 copies/mL) were obtained in 1-month intervals for 2 years. Cross-sectional blood samples from ten additional HIV-infected persons with a long history of suppressive ART (6–15 years) were obtained. All blood samples were obtained and processed at the Center for Research in Infectious Diseases (CIENI) of INER in Mexico City. Plasma and DNA of total CD4⁺ T cells were stored from each blood sample donation. Plasma viral loads for clinical follow-up were determined using the m2000 system (Abbott, Abbott Park, IL, USA). CD4⁺ and CD8⁺ T cell counts were assessed by flow cytometry using the TruCount kit in a FACSCanto II instrument (BD Bioscience, San Jose, CA, USA).

Residual plasma viral load measurement

Viral RNA was extracted from 9 mL of plasma. Plasma was centrifuged at 15,000 rpm for 3 h, at 4 °C. The supernatant was removed and the pellet re-suspended in 140 µL. Viral RNA was purified from the pellet using QIAamp viral RNA mini kit (QIAGEN, Valencia, CA, USA; resuspending in 41 µL of elution buffer) and stored at -80 °C until use. Residual plasma viral load (<40 RNA copies/mL) was determined by real-time PCR (qPCR) with TaqMan Fast Virus One-Step Master Mix (ThermoFisher, Waltham, MA, USA) using an in-house NL4-3-derived RNA standard. NL4-3 RNA was extracted from culture supernatants (QIAamp viral RNA mini kit, QIAGEN) and quantified (m2000 system, Abbott). The standard curve was generated performing two-fold dilutions ranging from 256 to 1 RNA copies. qPCR was performed in triplicate for standard curve points and duplicate for sample points using primers Fwd 5'-GGTCTCTCTGGTTAGACCAGAT-3' (HXB2 positions: 455–476), Rv 5'-CTGCTAGAGATTTCCACACTG-3' (635–614), and the TaqMan probe 5'-6FAM-AGTAGTGTGTGCCCGTCTGTT-TAMRA-3' (552–572), directed to the viral LTR, in 50 µL reactions containing 20.5 µL of standard or sample RNA, on a 7500 Real Time PCR System instrument (ThermoFisher). qPCR conditions were as follows: 50 °C for 50 min for reverse transcription, 95 °C for 20 s, followed by 40 cycles of 95 °C for 15 s and 60 °C for 60 s.

Proviral load measurement from total CD4⁺ cells

Peripheral blood mononuclear cells (PBMC) were obtained from 16 mL of blood by Ficoll gradient centrifugation. Total CD4⁺ T cells were purified by negative selection from 50 million PBMC, using EasySep Human CD4⁺ T Cell kit (StemCell Technologies, Vancouver, Canada), and DNA was purified using iPrep Purelink

gDNA Blood kit (ThermoFisher) on an iPrep Nucleic Acid Purification System (ThermoFisher). Proviral load was measured from total CD4⁺ T cells using TaqMan Gene Expression Master Mix (ThermoFisher). An in-house DNA standard derived from the ACH-2 cell line (obtained through the National Institutes of Health AIDS Reagent Program; catalogue 349) [50], containing two HIV proviral copies per cell, was used [51, 52]. To construct the standard curve, DNA was extracted from 1×10^6 ACH-2 cells (iPrep Purelink gDNA Blood kit; ThermoFisher), and quantified (NanoDrop 1000; Thermo Scientific). The standard curve was obtained from 12 1:3 serial dilution points starting from a DNA stock containing 400 ng, in triplicate. qPCR was performed from 400 ng total CD4⁺ T cell DNA, using the same primers and probe used for residual plasma viral load quantification (see above). RNase P quantification was used as internal reference (Taqman RNase P Control Reagents kit; ThermoFisher). PCR conditions were as follows: 95 °C for 10 min, followed by 45 cycles of 95 °C for 15 s and 60 °C for 1 min. All values were corrected for CD4⁺ T cell purity, assessed by multi-color flow cytometry as explained below, and adjusted to 1×10^6 genomes.

Viral LTR methylation analysis method

Bisulfite conversion

Methylation was assessed by bisulfite conversion using 900 ng of total CD4⁺ T cell DNA, with EpiTect Fast DNA Bisulfite kit (Qiagen), on a QIAcube instrument (Qiagen). Samples were eluted in 16 µL elution buffer. The complete volume of bisulfite-treated DNA was used immediately to perform first-round LTR-Gag PCR in duplicate, as described below. Human Methylated & Non-methylated DNA Set kit (Zymo Research, Irvine, CA, USA) was used as control to evaluate the efficiency of bisulfite conversion, according to manufacturer's specifications, using Platinum Taq DNA polymerase (ThermoFisher). Illumina adaptors for sequencing library construction were added to the primers recommended by the manufacturer (Additional file 5: Table S2).

5'-LTR amplification

The viral 5'-LTR region was amplified using nested PCR, with Platinum Taq DNA polymerase (ThermoFisher), obtaining U3-R amplicons from both bisulfite-treated and untreated samples. Illumina adaptors for sequencing library construction were included as part of the second-round primers (Additional file 5: Table S2). For bisulfite-untreated samples, primers LTR1-20Fw and LTR835-856Rv (Additional file 5: Table S2) were used for the first-round PCR, allowing amplification of an LTR-gag fragment, with the following conditions: 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 1 min, and a final step of 72 °C for 5 min. For

the second-round PCR, primers Nex-LTR213-236Fw and Nex-LTR471-495Rv were used, allowing amplification of the U3-R region of the 5'-LTR, with the following conditions: 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 1 min, and a final step of 72 °C for 5 min. For bisulfite-treated samples, primers LTR1-20TxFw and LTR835-856TxRv were used for the first-round PCR with the following conditions: 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 48 °C for 30 s, and 72 °C for 1 min, and a final step of 72 °C for 5 min. For the second-round PCR, primers Nex-LTR213-236TxFw and Nex-LTR471-495TxRv were used, with the following conditions: 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 47 °C for 30 s, and 72 °C for 1 min, and a final step of 72 °C for 5 min.

LTR next generation sequencing

Next generation sequencing (NGS) libraries were constructed from purified U3-R LTR amplicons (both from bisulfite-treated and -untreated DNA), using the amplicon approach with Nextera XT DNA sample preparation and index kits (Illumina, San Diego, CA, USA) (Fig. 1). Unique index combinations were added to each sample for multiplexing. Libraries were normalized to 60 nM and agarose gel-purified (338–440 bp bands, QIAquick gel extraction kit, Qiagen). After normalizing again, libraries were pooled, quantified (Qubit 3.0 Fluorometer; ThermoFisher), and run on a MiSeq instrument (Illumina), using v2 300-cycle pair-end kits (Illumina), according to manufacturer's instructions.

Read alignment and filtering

Reads were assembled and filtered using USEARCH program version 9.2.64 [53]. The following algorithm was used: (1) fastq R1 and R2 files were fused together, (2) reads were filtered by primer binding site presence (allowing a maximum of two nucleotide difference) and size (209–309 nucleotides for samples, 224–324 nucleotides for the DAPK methylation controls), (3) primer sequences were removed (24 and 25 nucleotides at the 5' and 3' end respectively for the samples, 32 and 35 for the controls), (4) reads were filtered by quality (maximum expected error = 1), (5) abundance of each sequence was determined, (6) singlets were eliminated, and (7) reads with abundance < 2% were filtered out. The reads obtained from USEARCH software showed 97.6% (range 95.1–100%) conversion rate of cytosines outside CpG sites for the samples and 100% for the controls, using the software Quma (Quantification Tool for Methylation Analysis) [54].

LTR methylation analysis calculations

Sequencing reads obtained for each time point for each individual were normalized to 100 and aligned to the HXB2

reference with Clustal Omega [55]. A position-specific score matrix (PSSM) was constructed using R and Seqinr [56]. The nine canonical CpG sites in the amplified HIV LTR region were located in the PSSM and the proportion of methylated (CpG), non-methylated (TpG), and mutated sequences was estimated for each one. The PSSM also provided information about additional CpG, CHG, and CHH methylation sites. These data were represented by multiple pie graphs using the program Raw Graphs [57].

We used Quma software [54], optimized for NGS reads analysis, to quantify and evaluate differences in global methylation, and estimated a "CpG Methylation Index" metric, averaging the proportion of methylated sequences for each of the nine canonical HXB2 LTR CpG sites in each sample. The CpG Methylation Index metric ranged from 0 (completely unmethylated) to 100 (fully methylated). Methylation differences between samples of different time-points were assessed using Mann-Whitney U test, using Quma. In order to be more comprehensive in the methylation analysis, we also studied non-canonical methylation patterns along the follow-up period, establishing a measure of methylation on CpG sites different to the canonical ones, present in the HXB2 reference (additional CpG methylation index), and methylation in sites other than CpG (non-CpG methylation index), averaging the proportion of methylated sequences for each relevant position in each category.

Immunophenotyping

A panel of 12 antibodies was selected to identify by flow cytometry CD4⁺ T lymphocyte subpopulation proportions from each sample time point for all participants included in the study. Antibodies used included the following: *Biolegend*: HLA-DR BV785 clone L243 (307641), CD45RO BV650 clone UCHL1 (304231), CD3 BV 570 clone UCHT1 (300435), CD95 PECy5 clone DX2 (305610), CCR7 APC Cy7 clone G043H7 (353211), CCR5 BV421 clone J418F1 (359117), PD1 PE clone EH12.2H7 (329906), and CD127 FITC clone A019D5 (351311). *BD*: CD28 BV711 clone CD28.2 (563131), CD38 AF700 clone HIT2 (560676), CD45 APC clone RUO (340943). *Invitrogen*: CD4 PE Cy5.5 clone S3.5 (MHCD0418). *Life technologies*: LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (L34957).

All antibodies were previously titrated using healthy donor CD4⁺ T lymphocytes. For staining, 1×10^6 CD4⁺ T cells were resuspended in 2 mL of staining buffer (1× PBS, 2% fetal bovine serum and 0.5% EDTA). Cells were centrifuged at 2000 rpm for 5 min and decanted. Staining was carried out in the remaining volume (approximately 50 µL), adding the anti-CCR7 antibody first. After 10-min incubation, the rest of the antibodies were subsequently added and the tube was incubated for 1 h

at room temperature in the dark. Cells were then washed with 3 mL of staining buffer, centrifuged at 2000 rpm for 5 min, decanted, and fixed with 0.3 mL of 1% paraformaldehyde. Cells were acquired on an LSR Fortessa cytometer (BD).

CD4⁺ T cell subpopulations were defined as follows: Naïve CD4⁺ T cells (CD3⁺/CD4⁺/CD45RO⁻/CCR7⁺/CD28⁺/CD95⁻), memory CD4⁺ T cells (CD3⁺/CD4⁺/CD28⁺/CD95⁺), stem cell memory (T_{SCM}: CD3⁺/CD4⁺/CD45RO⁻/CCR7⁺/CD28⁺/CD95⁺), central memory (T_{CM}: CD3⁺/CD4⁺/CD45RO⁺/CCR7⁺/CD28⁺), transitional memory (T_{TM}: CD3⁺/CD4⁺/CD45RO⁺/CCR7⁻/CD28⁺), effector memory (T_{EM}: CD3⁺/CD4⁺/CD45RO⁺/CCR7⁻/CD28⁻), terminal effector (T_{TE}: CD3⁺/CD4⁺/CD45RO⁻/CCR7⁻/CD28⁻), and T_{new} (T_{NEW}: CD3⁺/CD4⁺/CD45RO⁻/CCR7⁻/CD28⁺) using the program Cytobank [58]. For cell subpopulations with a limited number of captured events, we used a Poisson distribution to establish the minimal threshold for analysis in 100 positive events and 10% variation coefficient. The gating strategy is shown in Additional file 3: Figure S2.

Statistical analysis

Comparisons of clinical and methylation parameters of the longitudinal data were performed for each time-point against month 30 as reference (overall and both for the methylation increase and constant methylation groups), and against the long-term ART group (months 30 and 42 are shown in figures and tables). Medians with interquartile ranges were used for descriptive analysis. Variables from related samples (different time-points from the same participant) were compared using the Wilcoxon test and variables from independent samples were compared using the Mann-Whitney test. A $p < 0.05$ was considered significant. Statistical analyses were conducted using SPSS version 23.0 and STATA version 14.0.

Exploratory factorial analysis using principal components (PCA) was used as a variable reduction method to describe data structure and summarize the variance in baseline and follow-up clinical variables into fewer dimensions. Indexes were constructed, followed by orthogonal (varimax) rotation. Variables included were age, nadir CD4⁺ T cell count, pre-ART pVL, delta CD4⁺ T cell count and % at 48 months, delta CD8⁺ T cell count and % at 48 months, CD4/CD8 ratio change at 48 months, time on ART, time to achieve undetectable pVL, residual pVL, proviral load, CpG methylation index, number of ART regimen changes, ART backbone (tenofovir + emtricitabine vs. abacavir + lamivudine), ART third drug (non-nucleoside RT inhibitors vs. protease inhibitors), presence of blips, and CD4⁺ T cell subpopulation proportions: naeoxide RT inhibitors v_{SCM} , T_{CM}, T_{EM}, T_{TM}, T_{NEW}, T_{TE}. Six factors showed an eigenvector value higher than 1 (cumulative variance

77%), from which three were selected (cumulative variance 51%) for further presentation and analysis. Factor scores were obtained and graphed. PCA was performed using STATA version 14.0.

Additional files

Additional file 1: Figure S1. Gating strategy to define CD4⁺ T cell subpopulations. (PDF 529 kb)

Additional file 2: Table S1. Antiretroviral treatment history of participants. (DOC 54 kb)

Additional file 3: Figure S2. Detailed HIV 5'-LTR variant composition and methylation analysis in the rest of participants. Detailed variant composition per time point is shown for ten participants with short-term antiretroviral treatment not included in Fig. 3 (TP23, TP24, TP27, TP29, TP30, TP32, TP26, TP28, TP33, TP34), and for the single time point available for the 10 participants with long-term antiretroviral therapy (TP35, TP36, TP37, TP38, TP40, TP41, TP43, TP44, TP45, TP46). Each horizontal line represents a variant. The number of reads representing that variant and its proportion in the sample are shown on the right. Each circle represents a canonical CpG site, numbered according to the HXB2 reference: black—methylated, white—unmethylated, gray—mutated. (PDF 16804 kb)

Additional file 4: Figure S3. Changes in 5'-LTR overall methylation patterns along time. A comparison of methylation patterns is shown for all participants with short-term antiretroviral therapy (ART) at 30 months (2.5 years) of follow-up (top), all participants with short-term ART at 42 months (3.5 years) of follow-up (center) and all participants with long-term ART (single time point available, 6 to 15 years) (bottom). Each line depicts a summary of all variants observed per time point. CpG sites are shown as pie charts indicating the proportion of methylated (black), unmethylated (light gray) or mutated (dark gray) variants in the sample. Canonical HXB2 CpG positions are numbered. Additional CpG, CHG and CHH methylation-susceptible sites, not observed in the HXB2 sequence, are also included (marked with *). The number of reads and the total number of variants obtained per time point, after filtering and aligning, are shown. Additionally, measurements of residual plasma viral load, proviral load and proportion of memory (CD3⁺/CD4⁺/CD28⁺/CD95⁺; black) and naïve (CD3⁺/CD4⁺/CD45RO⁻/CCR7⁺/CD28⁺/CD95⁻; light gray) CD4⁺ T cells per time point are shown. (PDF 207 kb)

Additional file 5: Table S2. Primers for HIV 5'-LTR and methylation control amplification and sequencing. (DOCX 17 kb)

Abbreviations

ART: Antiretroviral treatment; DAPK1: Death-associated protein kinase 1; HDACi: Histone deacetylase inhibitor; HIV: Human immunodeficiency virus; IQR: Interquartile range; LTR: Long terminal repeat; NGS: Next generation sequencing; NNRTI: Non-nucleoside reverse transcriptase inhibitors; NRTI: Nucleoside reverse transcriptase inhibitors; PBMC: Peripheral blood mononuclear cells; PCR: Polymerase chain reaction; PI: Protease inhibitors; PSSM: Position-specific score matrix; pVL: Plasma viral load; T_{CM}: Central memory T cells; T_{EM}: Effector memory T cells; T_{NEW}: T new; T_{SCM}: Stem cell memory T cells; T_{TE}: Terminally differentiated effector T cells; T_{TM}: Transitional memory T cells

Acknowledgments

We thank all participants in the study for their generosity, Diego Adrian Oliva-Rico for his technical advice on bisulfite conversion, Maribel Soto-Nava for supplying NL4-3 virus for residual pVL measurement, Robert Edgar for his advice on the use of USEARCH, Ramon Hernández-Juan for quantifying NL4-3 standards, Silvia del Arenal for sample inventory and management, and Olivia Briceparticipants in the study for their generosity, Diego A The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: ACH-2 Cells (cat# 349). CNCR was supported by a scholarship from Consejo Nacional de Ciencia y Tecnología (CONACyT). This paper constitutes partial fulfillment of the requisites of Programa de Doctorado en

Ciencias Biológicas, Universidad Nacional Autónoma de México (UNAM) for CNCR.

Authors' contributions

CNCR, GSMO, SAR, and GRT conceived and designed the study; CNCR performed the experiments, FJPG performed statistical analyses; MMF provided assistance with next generation sequencing; AMO, LKC contributed with participant enrolment and provided clinical files; APCA and CEO contributed with next generation sequencing data analysis; ICR assisted with database construction and management; and CNCR and SAR wrote the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by grants from the Mexican Government (Comisión de Equidad y Género de las Legislaturas LX-LXI y Comisión de Igualdad de Género de la Legislatura LXII de la Cámara de Diputados de la República Mexicana) to GRT, and Consejo Nacional de Ciencia y Tecnología (CONACyT SALUD-2017-01-289725) to SAR. The funders had no role in study design, data collection and analysis decision to publish, or preparation of the manuscript.

Availability of data and materials

The datasets used during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

This study was approved by the Institutional Review Board of the National Institute of Respiratory Diseases (INER) (registry: DHHS IORG0003948; CONBIOETICA-09-CEI-003-20160427) in Mexico City (project code: B26-09), and was conducted according to the principles of the Declaration of Helsinki. All participants gave written informed consent before blood sample donation.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Centre for Research in Infectious Diseases, National Institute of Respiratory Diseases, Tlalpan 4502, 14080 Mexico City, Mexico. ²National Institute of Public Health, Mexico City, Mexico. ³Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ⁴Adolfo López Mateos Hospital, ISSSTE, Mexico City, Mexico.

Received: 14 May 2019 Accepted: 30 August 2019

Published online: 13 September 2019

References

- Flexner C, Plumley B, Brown Ripin DH. Treatment optimization: an outline for future success. *Curr Opin HIV AIDS*. 2013;8:523–7.
- Furtado MR, Callaway DS, Phair JP, Kunstman KJ, Stanton JL, Macken CA, et al. Persistence of HIV-1 transcription in peripheral-blood mononuclear cells in patients receiving potent antiretroviral therapy. *N Engl J Med*. 1999;340:1614–22.
- Chun TW, Stuyver L, Mizell SB, Ehler LA, Mican JA, Baseler M, et al. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc Natl Acad Sci U S A*. 1997;94:13193–7.
- Chun TW, Engel D, Mizell SB, Hallahan CW, Fischette M, Park S, et al. Effect of interleukin-2 on the pool of latently infected, resting CD4+ T cells in HIV-1-infected patients receiving highly active anti-retroviral therapy. *Nat Med*. 1999;5:651–5.
- Eriksson S, Graf EH, Dahl V, Strain MC, Yukl SA, Lysenko ES, et al. Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. *PLoS Pathog*. 2013;9:e1003174.
- Cillo AR, Sobolewski MD, Bosch RJ, Fyne E, Piatak M Jr, Coffin JM, et al. Quantification of HIV-1 latency reversal in resting CD4+ T cells from patients on suppressive antiretroviral therapy. *Proc Natl Acad Sci U S A*. 2014;111:7078–83.
- Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DL, et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*. 2013;155:540–51.
- Rasmussen TA, Tolstrup M, Brinkmann CR, Olesen R, Erikstrup C, Solomon A, et al. Panobinostat, a histone deacetylase inhibitor, for latent-virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: a phase 1/2, single group, clinical trial. *The Lancet HIV*. 2014;1:e13–21.
- Du Chene I, Basyuk E, Lin YL, Triboulet R, Knezevich A, Chable-Bessia C, et al. Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. *EMBO J*. 2007;26:424–35.
- Keedy KS, Archin NM, Gates AT, Espeseth A, Hazuda DJ, Margolis DM. A limited group of class I histone deacetylases acts to repress human immunodeficiency virus type 1 expression. *J Virol*. 2009;83:4749–56.
- Imai K, Togami H, Okamoto T. Involvement of histone H3 lysine 9 (H3K9) methyltransferase G9a in the maintenance of HIV-1 latency and its reactivation by BIX01294. *J Biol Chem*. 2010;285:16538–45.
- Friedman J, Cho WK, Chu CK, Keedy KS, Archin NM, Margolis DM, et al. Epigenetic silencing of HIV-1 by the histone H3 lysine 27 methyltransferase enhancer of Zeste 2. *J Virol*. 2011;85:9078–89.
- Blazkova J, Trejbalova K, Gondois-Rey F, Halfon P, Philibert P, Guiguen A, et al. CpG methylation controls reactivation of HIV from latency. *PLoS Pathog*. 2009;5:e1000554.
- Kauder SE, Bosque A, Lindqvist A, Planelles V, Verdin E. Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathog*. 2009;5:e1000495.
- Palacios JA, Perez-Pinar T, Toro C, Sanz-Minguela B, Moreno V, Valencia E, et al. Long-term nonprogressor and elite controller patients who control viremia have a higher percentage of methylation in their HIV-1 proviral promoters than aviremic patients receiving highly active antiretroviral therapy. *J Virol*. 2012;86:13081–4.
- Doerfler W, Weber S, Kemal K, Weiser B, Korn K, Anastos K, et al. Epigenetic modifications of HIV proviral LTRs: potential targets for cure. *Retrovirology*. 2012;9:O4.
- Blazkova J, Murray D, Justement JS, Funk EK, Nelson AK, Moir S, et al. Paucity of HIV DNA methylation in latently infected, resting CD4+ T cells from infected individuals receiving antiretroviral therapy. *J Virol*. 2012;86:5390–2.
- Weber S, Weiser B, Kemal KS, Burger H, Ramirez CM, Korn K, et al. Epigenetic analysis of HIV-1 proviral genomes from infected individuals: predominance of unmethylated CpG's. *Virology*. 2014;449:181–9.
- Trejbalova K, Kovarova D, Blazkova J, Machala L, Jilich D, Weber J, et al. Development of 5' LTR DNA methylation of latent HIV-1 provirus in cell line models and in long-term-infected individuals. *Clin Epigenetics*. 2016;8:19.
- Yoon CH, Jang DH, Kim KC, Park SY, Kim HY, Kim SS, et al. Disruption of polycomb repressor complex-mediated gene silencing reactivates HIV-1 provirus in latently infected cells. *Intervirology*. 2014;57:116–20.
- Matsuda Y, Kobayashi-Ishihara M, Fujikawa D, Ishida T, Watanabe T, Yamagishi M. Epigenetic heterogeneity in HIV-1 latency establishment. *Sci Rep*. 2015;5:7701.
- Rafati H, Parra M, Hakre S, Moshkin Y, Verdin E, Mahmoudi T. Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency. *PLoS Biol*. 2011;9:e1001206.
- Van Duyn R, Guendel I, Narayanan A, Gregg E, Shafagati N, Tyagi M, et al. Varying modulation of HIV-1 LTR activity by Baf complexes. *J Mol Biol*. 2011;411:581–96.
- Harwig A, Das AT, Berkhout B. Retroviral microRNAs. *Curr Opin Virol*. 2014;7:47–54.
- Patel P, Ansari MY, Bapat S, Thakar M, Gangakhedkar R, Jameel S. The microRNA miR-29a is associated with human immunodeficiency virus latency. *Retrovirology*. 2014;11:108.
- Barichevsky S, Naidoo J, Mhlanga MM. Non-coding RNAs and HIV: viral manipulation of host dark matter to shape the cellular environment. *Front Genet*. 2015;6:108.
- Rice AP. Roles of microRNAs and long-noncoding RNAs in human immunodeficiency virus replication. *Wiley Interdiscip Rev RNA*. 2015;6:661–70.
- Suzuki K, Ahlenstiel C, Marks K, Kelleher AD. Promoter targeting RNAs: unexpected contributors to the control of HIV-1 transcription. *Mol Ther Nucleic Acids*. 2015;4:e222.
- Li E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet*. 2002;3:662–73.
- Christman JK, Weich N, Schoenbrun B, Schneiderman N, Acs G. Hypomethylation of DNA during differentiation of Friend erythroleukemia cells. *J Cell Biol*. 1980;86:366–70.
- Koiwa T, Hamano-Usami A, Ishida T, Okayama A, Yamaguchi K, Kamihira S, et al. 5'-Long terminal repeat-selective CpG methylation of latent human T-cell leukemia virus type 1 provirus *In Vitro* and *In Vivo*. *Journal of Virology*. 2002;76:9389–97.

32. Harbers K, Schnieke A, Stuhlmann H, Jahner D, Jaenisch R. DNA methylation and gene expression: endogenous retroviral genome becomes infectious after molecular cloning. *Proc Natl Acad Sci U S A*. 1981;78:7609–13.
33. Hejnar J, Plachy J, Geryk J, Machon O, Trejbalova K, Guntaka RV, et al. Inhibition of the rous sarcoma virus long terminal repeat-driven transcription by in vitro methylation: different sensitivity in permissive chicken cells versus mammalian cells. *Virology*. 1999;255:171–81.
34. Lavie L, Kitova M, Maldener E, Meese E, Mayer J. CpG methylation directly regulates transcriptional activity of the human endogenous retrovirus family HERV-K (HML-2). *J Virol*. 2005;79:876–83.
35. Persaud D, Pierson T, Ruff C, Finzi D, Chadwick KR, Margolick JB, et al. A stable latent reservoir for HIV-1 in resting CD4(+) T lymphocytes in infected children. *J Clin Invest*. 2000;105:995–1003.
36. Trono D, Van Lint C, Rouzioux C, Verdin E, Barre-Sinoussi F, Chun TW, et al. HIV persistence and the prospect of long-term drug-free remissions for HIV-infected individuals. *Science*. 2010;329:174–80.
37. Duverger A, Jones J, May J, Bibollet-Ruche F, Wagner FA, Cron RQ, et al. Determinants of the establishment of human immunodeficiency virus type 1 latency. *J Virol*. 2009;83:3078–93.
38. Karn J. The molecular biology of HIV latency: breaking and restoring the Tat-dependent transcriptional circuit. *Curr Opin HIV AIDS*. 2011;6:4–11.
39. LaMere SA, Chaillon A, Huynh C, Smith DM, Gianella S. Challenges in quantifying cytosine methylation in the HIV provirus. *MBio*. 2019;10.
40. Van Lint C, Bouchat S, Marcello A. HIV-1 transcription and latency: an update. *Retrovirology*. 2013;10:67.
41. Bachmann N, Von Siebenthal C, Vongrad V, Neumann K, Turk T, Beerwinkel N, et al. Determinants of HIV-1 reservoir size and long-term dynamics under suppressive ART In: 25th CROI Conference on Retroviruses and Opportunistic Infections; 2018 March 4–7; Boston, Massachusetts. United States of America; 2018. Available from: <http://www.croiconference.org/sessions/determinants-hiv-1-reservoir-size-and-long-term-dynamics-under-suppressive-art>.
42. Huang SH, Ren Y, Thomas AS, Chan D, Mueller S, Ward AR, et al. Latent HIV reservoirs exhibit inherent resistance to elimination by CD8+ T cells. *J Clin Invest*. 2018;128:876–89.
43. Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, et al. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science*. 1997;278:1295–300.
44. Pion M, Jordan A, Biancotto A, Dequiedt F, Gondois-Rey F, Rondeau S, et al. Transcriptional suppression of in vitro-integrated human immunodeficiency virus type 1 does not correlate with proviral DNA methylation. *J Virol*. 2003;77:4025–32.
45. Alinejad-Rokny H, Anwar F, Waters SA, Davenport MP, Ebrahimi D. Source of CpG depletion in the HIV-1 genome. *Mol Biol Evol*. 2016;33:3205–12.
46. Hernando-Herraez I, Garcia-Perez R, Sharp AJ, Marques-Bonet T. DNA Methylation: insights into human evolution. *PLoS Genet*. 2015;11:e1005661.
47. Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, Yassine-Diab B, et al. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med*. 2009;15:893–900.
48. Von Stockenström S, Odeval L, Lee E, Sinclair E, Bacchetti P, Killian M, et al. Longitudinal genetic characterization reveals that cell proliferation maintains a persistent HIV type 1 DNA pool during effective HIV therapy. *J Infect Dis*. 2015;212:596–607.
49. Rosca A, Anton G, Ene L, Iancu I, Temereanca A, Achim CL, et al. Immunoassay and molecular methods to investigate DNA methylation changes in peripheral blood mononuclear cells in HIV infected patients on cART. *J Immunoassay Immunochem*. 2017;38:299–307.
50. Clouse KA, Powell D, Washington I, Poli G, Strebel K, Farrar W, et al. Monokine regulation of human immunodeficiency virus-1 expression in a chronically infected human T cell clone. *J Immunol*. 1989;142:431–8.
51. Cannon P, Kim SH, Ulich C, Kim S. Analysis of Tat function in human immunodeficiency virus type 1-infected low-level-expression cell lines U1 and ACH-2. *J Virol*. 1994;68:1993–7.
52. Symons J, Chopra A, Malatinkova E, De Spiegelaere W, Leary S, Cooper D, et al. HIV integration sites in latently infected cell lines: evidence of ongoing replication. *Retrovirology*. 2017;14:2.
53. Edgar RC, Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*. 2015;31:3476–82.
54. Kumaki Y, Okano M: QUMA (QUantification tool for Methylation Analysis). <http://quma.cdb.riken.jp> (2019). Accessed 13 Feb 2019.
55. EMBL-EBI: Clustal Omega. <https://www.ebi.ac.uk/Tools/msa/clustalo/> (2018). Accessed 7 Apr 2018.
56. Charif D, Lobry JR. SeqinR 1.0-2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In: Bastolla U, Porto M, Roman E, Vendruscolo M, editors. Structural approaches to sequence evolution: Molecules, networks, populations. New York: Springer Verlag; 2007. p. 207–32.
57. Mauri M, Elli T, Caviglia G, Ubaldi G, Azzi M. RAWGraphs: a visualisation platform to create open outputs. In: Proceedings of CHIItaly '17, Cagliari, Italy; 2017. p. 1–5.
58. Kotecha N, Krutzik PO, Irish JM. Web-based analysis and publication of flow cytometry experiments. *Curr Protoc Cytom*. 2010;Chapter;10:1–40.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

