

CARACTERIZACIÓN GENÓMICA DE LOS SITIOS DE INTEGRACIÓN DE PROVIRUS HTLV-I EN LINFOCITOS DE INDIVIDUOS ASINTOMÁTICOS DEL SUROCCIDENTE

GENOMIC CHARACTERIZATION OF OF HTLV-I PROVIRUS INTEGRATION SITES IN LYMPHOCYTES FROM ASYMPTOMATIC INDIVIDUALS IN SOUTHWEST

Mercedes Salcedo-Cifuentes¹, Jesús Cabrera², Yesid Cuesta Astroz¹, Martha C. Domínguez¹, Adalberto Sánchez¹, Irene Tischer³, Felipe García-Vallejo¹

¹Laboratorio de Biología Molecular y Patogénesis. Departamento de Ciencias Fisiológicas. Escuela de Ciencias Básicas. Facultad de Salud. Universidad del Valle. Cali. Colombia.

²Facultad de Ciencias. Universidad de Nariño. Pasto. Colombia.

³Grupo de Bioinformática. Escuela de Ingeniería de Sistemas. Facultad de Ingenierías. Universidad del Valle. Cali. Colombia.

RESUMEN

Aunque algunas proteínas virales juegan un papel importante en la patogénesis de las enfermedades asociadas con el HTLV-I, su estilo integracional permanece parcialmente sin resolver. En este trabajo se evaluaron características de la cromatina en una región de 100 Kbp alrededor de provirus HTLV-I en sujetos seropositivos asintomáticos. A partir de una biblioteca genómica humana de clones IPCR proveniente de cuatro individuos seropositivos asintomáticos del Suroccidente colombiano, se secuenciaron 27 clones. Mediante el uso del programa BLAT del Genome Browser UCSC, se calculó el contenido de C/G y las diferencias estadísticas del número de islas CpG, de elementos repetidos, de genes referenciados y otras covariables genómicas, mediante pruebas no paramétricas. El contenido promedio de C/G en una región de 50 bp del sitio de integración proviral, fue de 57.6%. Las integraciones se distribuyeron con mayor frecuencia en las porciones subteloméricas de los cromosomas 1, 2, 3, 4, 5, 9, 10, 11, 14, 16, 17, and 19. El 89.9% de las secuencias correspondieron a elementos repetidos, 9.65% a islas CpG y solamente 0.46% a genes ($p < 0.001$). SINE and LINE fueron las secuencias repetidas mas abundantes. AluSx fue el elemento Alu más frecuente (70.9%) ($p < 0.01$). Se concluye que la integración del HTLV-I en el estado asintomático ocurre en áreas de la cromatina celular con elevada densidad de elementos SINE/LINE que estarían favoreciendo estados conformacionales variables de la cromatina, además de eventos epigenéticos.

Palabras claves: Integración Retroviral. HTLV-I. Seropositivos Asintomáticos. Expansión Clónica. PCR Inverso. Genoma Humano. Bioinformática.

Recibido: Agosto 30 de 2009

Aceptado: Noviembre 16 de 2009

Correspondencia: Departamento de Ciencias Fisiológicas.

Escuela de Ciencias Básicas. Facultad der Salud. Universidad del Valle. Sede San Fernando, Edificio 120 oficina 508. Calle 436-00. Cali. Colombia. E-mail: labiomol@gmail.com.

ABSTRACT

Although some viral proteins play important functions in the associated diseases, the role of HTLV-I proviral integration style remains partially unsolved. In this work we evaluated several characteristics of the genomic regions associated with HTLV-I provirus in asymptomatic carriers. From a human genome library of Inverse PCR DNA from four HTLV-I asymptomatic seropositive subjects of southwest pacific of Colombia twenty seven IPCR clones were sequenced. Using UCSC BLAT Genome Browser, mapping and characterization of several genomic covariables in a window of 100 Kbp around the integration sites were performed. The average of G/C content in a 50 bp flanking proviral region was $57.6\% \pm 6.93$, which was significantly higher than 60 random genome control sequences. IPCR sequences were distributed in several loci of human chromosomes 1, 2, 3, 4, 5, 9, 10, 11, 14, 16, 17, and 19. The 89.9% of sequences were repetitive elements, 9.65% CpG island and only 0.46% genes ($p < 0.001$). SINEs (42.5%) and LINEs (23.7%) were the most abundant but AluSx constituted the 70.9% of all Alu types that were recorded ($p < 0.01$). We propose that genome integration of HTLV-I during asymptomatic status, occur in chromatin areas with a high density of non coding repetitive elements (SINE/LINE) that would favored variable conformational states of chromatin and epigenesis.

Key words Retrovirus integration. Human T-Lymphotropic Virus I. Asymptomatic carriers. Clonal expansion. Inverse PCR. Human Genome. Bioinformatics.

Introducción

La infección por el Virus Linfotrópico Humano tipo I (HTLV-I), de la cual se calculan alrededor de 20 millones de personas infectadas en el mundo, constituye un problema serio de salud pública en muchas áreas del planeta incluyendo la región Caribe y Suramérica (1). Varios estudios epidemiológicos muestran que las regiones suroccidente y del Caribe de Colombia, son áreas de media y elevada seroprevalencia del virus. Aunque la gran mayoría de los infectados permanece en un estado asintomático, el 2 al 3% desarrolla una leucemia de células T en el adulto

(ATL) (2, 3), mientras que del 1 al 2% una condición neurodegenerativa denominada Paraparesia Espástica Tropical/Mielopatía Asociada al HTLV-I (PET/MAH) (4)

La integración del cDNA retroviral en el genoma del hospedero constituye una etapa crucial de su ciclo de vida puesto que está directamente implicada en la expresión eficiente de su progenie (5, 6). Se ha estudiado la relación entre la modalidad de integración del virus y su asociación con la patogénesis de enfermedades asociadas. Aunque todavía se conoce poco, en la Leucemia de células T en adultos (ATL), se ha observado un

patrón de integración monoclonal (7,8), en el que los provirus se localizan dentro o cerca a unidades de transcripción de genes reguladores del ciclo celular y transducción de señal, estas características podrían explicar la transformación y carcinogénesis celular (9).

La secuenciación del genoma humano ha permitido, identificar en términos cuantitativos cuales son las características del ambiente genómico que rodea los sitios de integración de retrovirus. Adicionalmente, los estudios de integración con base en los análisis bioinformáticos en varios retrovirus han mostrado que la integración no es aleatoria, y que incluso difiere dependiendo del tipo de virus (10, 11). Por ejemplo Mitchell et al 2004 (10) indicó que el VIH-1 prefiere integrarse dentro de unidades génicas, mientras que el MLV lo hace preferencialmente en regiones ricas en islas CpG o en regiones de inicio de transcripción. Estas diferencias en patrones de integración y en las características del ambiente genómico que rodea los provirus parecen estar influenciada por factores tanto (12), como celulares (13, 14).

Los análisis a nivel genómico sobre la integración del HTLV-I en linfocitos, se han centrado en pacientes ATL y PET/MAH, además de algunos estudios in vitro. Sin embargo es muy poco lo que se conoce acerca de las características genómicas de la

integración durante el estado asintomático. En este estudio se evaluó, mediante secuenciación de regiones de ADN humano adyacentes a provirus HTLV-I, tanto la expansión clonal en linfocitos de sangre periférica, como los microambientes genómicos asociados con la integración proviral en una extensión de 100 Kpb; para ello, se incluyeron 4 individuos seropositivos asintomáticos procedentes del Suroccidente Colombiano. Mediante secuenciación de 27 clones recombinantes construidos a partir de amplicones generados por PCR inverso, se analizó el contenido de GC, el locus cromosómico, genes referenciados, elementos repetidos y otras características estructurales del genoma asociado con la integración. Los resultados mostraron que éstas, son regiones que contienen algunos genes de función esencial, una elevada frecuencia de secuencias repetitivas (LINE y SINE/Alu), además de ser ricas en islas CpG.

Materiales y Métodos

Muestras y extracción de ADN de linfocitos de sangre periférica

En el estudio se incluyeron 4 individuos asintomáticos con diagnóstico confirmado de infección con HTLV-I+ por ELISA, Western Blot y PCR (15) provenientes del suroccidente colombiano (tabla 1). A cada uno de ellos, se les tomó, bajo consentimiento informado, una muestra de 5 mL de sangre total;

mediante gradientes de Ficoll Hypaque, se colectaron linfocitos de sangre periférica que fueron criopreservados a -156°C hasta su utilización (16). El ADN linfocitario se extrajo utilizando el estuche

Wizard™ Genomic DNA Purification (Promega, Madison, WI). El grado de pureza y la concentración de los ADN extraídos se evaluó por espectrofotometría (47).

Tabla 1. Descripción de las características de edad, sexo y seropositividad de individuos asintomáticos seropositivos confirmados para la infección por el HTLV-I del suroccidente colombiano

Individuo	Edad	Sexo	IE ^a	WB ^b	PCR ^c	Bandas IPCR ^d
ColAC-13	33	M	+	+	++	3
ColAC-16	62	M	+	+	++	6
ColAC-17	36	M	+	+	++	7
ColAC-23	45	M	+	+	++	5

a. Inmunoensayo enzimático; b. Western blot; c. Reacción en Cadena de la Polimerasa; d. Reacción en Cadena Inversa de la Polimerasa

Amplificación del ADN genómico por PCR inverso (IPCR)

La amplificación de secuencias de ADN celular adyacentes a provirus HTLV-I se efectuó siguiendo un protocolo ya descrito (17,18), que se muestra en la figura 1A. Todo el proceso incluyó las siguientes etapas: Una digestión inicial del ADN extraído con AluI. Una segunda etapa incluyó un proceso de autoligación de los extremos cohesivos de cada uno de los fragmentos obtenidos en la etapa anterior usando la enzima ADN ligasa del fago T4. Una tercera etapa los fragmentos autoligados se sometieron a digestión con SstII, endonucleasa de restricción que reconoce específicamente un sitio de corte dentro del gen gag del provirus; mediante este tratamiento se

conservaron los anillos de ADN que contenían porciones de genoma del hospedero más un fragmento de la región U5 del LTR proviral. Después de una etapa de calentamiento a 93°C por 30 minutos, los círculos relajados fueron amplificados por PCR. Los oligonucleótidos cebadores sintetizados a partir de secuencias específicas de la región U5 del provirus HTLV-I y corresponden fueron HTLV-009, 5'-AAGCCGGCAGTCAGTCGTGA-3' (8946-8927) y el HTLV-010, 5'-AAGTACCGGCAACTCTGCTG-3' (8958-8977). El alineamiento de los cebadores se llevó a cabo a 52°C por 30 segundos, seguida de una etapa de extensión a 72°C por 90 segundos. El número de ciclos de amplificación IPCR fue de 35 en total.

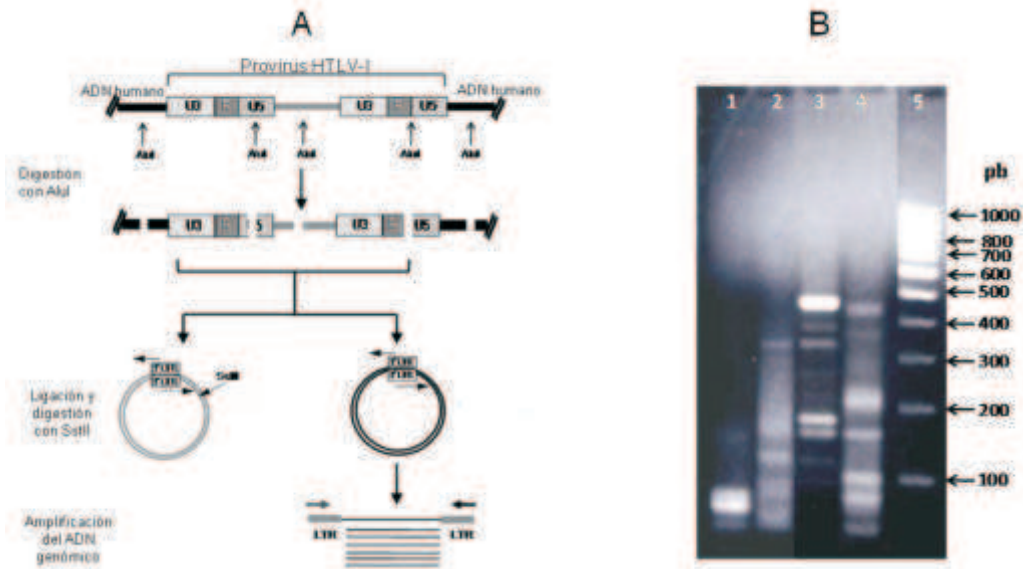


Figura 1. Determinación del grado de expansión clonal del HTLV-I en linfocitos de portadores asintomáticos. (A). Esquema empleado para realizar el PCR inverso con el fin de amplificar segmentos de genoma humano adyacentes a provirus HTLV-I. (B) Electroforesis en gel de agarosa del 2% que muestra los diferentes amplicones generados por IPCR. Los números a la derecha del gel corresponden a un patrón de tamaño molecular de una escalera de 100pb. (pb) par de bases.

Todos los productos de la IPCR se separaron, de acuerdo con su tamaño molecular mediante electroforesis en geles de agarosa al 2%. La visualización de las bandas de ADN se registró por la fluorescencia del Bromuro de Etidio unido al ADN. El grado de expansión clonal y el número de clones se calculó por conteo directo sobre el gel del número de bandas de ADN de los amplicones IPCR; cada banda representaba un clon celular.

Protocolos de clonación y secuenciación

Los fragmentos de ADN obtenidos por IPCR se eluyeron de los geles de agarosa usando el estuche comercial QIAEX II Gel Extraction® (Qiagen, Valencia, CA). Los fragmentos de ADN eluidos fueron ligados al sitio Hind III del plásmido vector pCR2.1 previamente linearizado con la misma enzima. Los plásmidos recombinantes se transformaron en células de *Escherichia coli* cepa DH5 siguiendo las instrucciones del estuche comercial TA Cloning System™ (In VitroGen, Carisbad, CA). Los recombinantes obtenidos por

extracción de ADN en lisis alcalina (19), fueron digeridos con EcoR I para liberar el fragmento de ADN celular clonado. En total se obtuvieron 24 clones IPCR del HTLV-I. El ADN celular clonado se secuenció empleando el método de secuenciación cíclica en un secuenciador ABI PRISM 310; (PE Applied Biosystems, Norwalk, CON).

Alineamientos y asignación de secuencias IPCR

Con el fin de eliminar la contaminación de porciones de LTR virales, las secuencias IPCR obtenidas en este trabajo, se analizaron mediante el programa de Vector Screen (www.ncbi.nlm.nih.gov/vecscree/vecscree.html). Ya editadas se alinearon con porciones de ADN humano, empleando el programa BLAT del “Genome Browser” de la Universidad de California en Santa Cruz (<http://genome.ucsc.edu/>) (20). Para los distintos análisis, se utilizaron solamente aquellas secuencias cuyo tamaño varió de 50 a 500 pb que tuvieron un score mayor de 95% de similitud consignadas hasta mayo de 2006. Adicionalmente la información se complementó a partir del programa Blast N del Instituto Nacional de Biotecnología e Información (NCBI) de los Estados Unidos (<http://www.ncbi.nih.gov>).

Se exploró una región de 100Kbp localizada alrededor de cada uno de los sitios de integración registrados. En

cada caso, utilizando las herramientas computacionales disponibles en el “Genome Browser” de la Universidad de California en Santa Cruz (<http://genome.ucsc.edu/>), en el el Gencard(<http://www.genecards.org/cgi-bin/carddisp.pl>); Gene-Entrez (<http://www.ncbi.nlm.nih.gov/ncbi/genentrez>) y Gene Ontology (GO) (<http://www.geneontology.org/index.shtml>), se obtuvo la información sobre el locus cromosómico, los genes anotados, los elementos repetidos y otras características de la cromatina adyacente a los provirus HTLV-I (orientación, proceso biológico en el que intervienen, función molecular y ubicación celular),

Análisis Estadísticos

Los análisis estadísticos se realizaron utilizando Stata versión 8.0 (30) y el software Statgraphics plus 5.1 para Windows. Las diferencias entre la composición nucleotídica de las secuencias adyacentes a los sitios de integración se calcularon usando la prueba t de student. Las diferencias estadísticas entre el tipo de secuencia adyacente al ADN proviral y la frecuencia de éstas en el genoma humano obtenidas por Venter et al-2001 (21) y las del IHGSC-2001 (“International Human Genome Sequencing Consortium”) (22), se calcularon aplicando la prueba de Chi²; un valor de $p < 0,05$ fue considerado como estadísticamente significativo. Se ejecutó una prueba de Kruskal-Wallis para determinar

humano alrededor de los alineamiento de las secuencias IPCR, se determinó que el 70,6% (12/17) de ellos ocurrieron en porciones no codificantes ($2=7,58$, $p<0.05$), mientras que el 29,4% de los restantes coincidieron con secuencias codificantes de genes anotados en las bases Gene Entrez o Gene Card. De acuerdo con el grado de repetición, las secuencias tipo LINE/L2 fueron las mas frecuentes (46,2%). Se determinaron diferencias estadísticamente significantes entre la distribución de las secuencias de los sitios de integración y aquellas del genoma total tomando como base de comparación los datos reportados por

Venter et al-2001(50) y los de Consorcio Internacional para la secuenciación del Genoma Humano (IHGSC) (51), tanto las diferencias encontradas, como los valores p de la pruebas estadísticas realizadas, se consignan tabla 2.

Tabla 2. Características de los sitios de integración del HTLV-I en el genoma de paciente colombianos con LLCTA con relación al genoma humano. Los datos de porcentajes de secuencias génicas y repetitivas en el genoma humano fueron tomadas de Venter et al 2001 (28) y del Consorcio Internacional para la Secuenciación del Genoma Humano (IHGSC) (31).

Clase de secuencias	Número de potenciales sitios de integración (porcentaje)	Número en miles (porcentaje en el genoma humano)	P*
<i>Unidades de transcripción</i>			
Genes	6 (31,6%)	30-40 (1.4%)	$P<0.05^S$
<i>Elementos repetitivos</i>			
LINES	6 (31,6%)	868 (21%)	$P<0,05^S$
SINES	0	1558 (13%)	NC
R S	5 (26,3%)	762(8%)	$P<0.05^S$
LTR	2 (10,5%)	737 (11%)	$P>0.5^{NS}$

La secuencia IPCR del clon AC1611 se localizó en el intrón 7 del gen que codifica para la proteína G acoplada al receptor 125 (GRP125) (figura 2). Éste es un megagen mapeado en el locus 4p15.31 con un tamaño de 128.674 pares de bases y contiene 19 exones; después del procesamiento, el mRNA transcrito final, es de 3.966

nucleótidos. De acuerdo con la información disponible en el Entrez Gene, la proteína GRP125, es un receptor huérfano localizado en la membrana plasmática de linfocitos de sangre periférica. Específicamente se une a proteínas y regula los pasos de señalización de neuropéptidos.

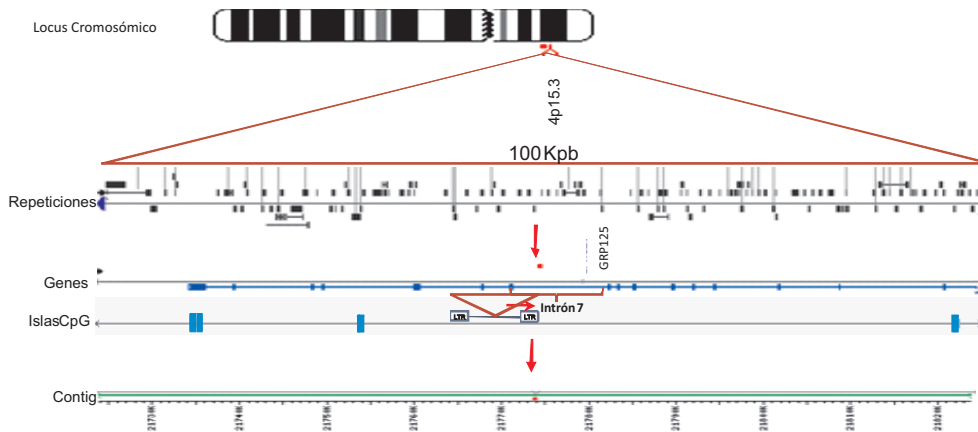


Figura 2. Integración de un provirus HTLV-I dentro del intrón 7 del gen GRP125 que codifica para la proteína G acoplada al receptor 125 localizado en el locus cromosómico 4p15.3 humano. En el gráfico se muestra también la distribución de repeticiones, el mega gen GRP125 y las islas CpG.

Adicionalmente dentro de la región genómica correspondiente a otras integraciones, en la ventana de observación 500 pb, se localizaron tres genes: el que codifica por la Proteína Cinasa C, alfa (PKC), que puede ser activado por calcio y por diacilglicerol como segundo mensajero y participa en varios procesos de señalización celular. El de la subunidad alfa del factor 1 de hipoxia inducible (HIF1A) el cual es un factor de transcripción que se activa por bajos niveles de oxígeno en células cultivadas, manteniendo la homeostasis celular bajo condiciones de hipoxia; y el gen de la Glicoproteína sináptica 2 (GPSN2), el cual cataliza, como Oxidorreductasa, la reducción del trans-2,3-enoil-CoA a acyl-CoA saturado.

El microambiente genómico de integración del HTLV-I

Con el objetivo de identificar

covariables genómicas que puedan definir características comunes de microambientes dentro del genoma humano que favorezcan la integración del HTLV-I en los asintomáticos seropositivos, se analizó la distribución de genes codificantes, islas CpG and elementos repetidos –RE- (SINE, LINE, LTR, DNA transposons and repeticiones simples) en una extensión de 100Kpb alrededor de los sitios de integración viral previamente identificados

Se determinó que el 90,2% (2243/2458) de las secuencias localizadas en las zonas de integración proviral correspondieron a elementos repetidos; un 8,9% (220/2458) a Islas CpG y solamente un 0,9% a genes referenciados en las bases de datos consultadas. Del análisis estadístico realizado aplicando la prueba de Kruskal-Wallis y Friedman se obtuvo un valor de $p < 0,001$ (figura 3A).

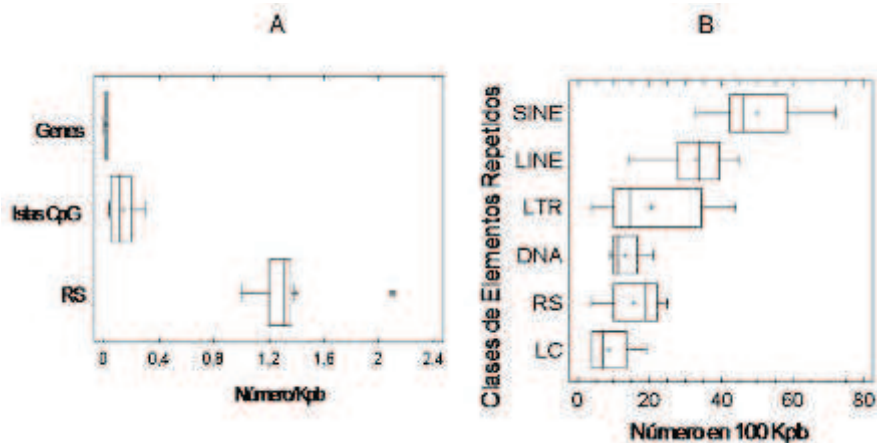


Figura 3. Graficas de cajas y bigotes que representan las diferencias en la distribución de distintos elementos del genoma humano localizados en una extensión de 100 Kbp adyacentes a secuencias provirales del HTLV-I. (A)

Diferencias de medias entre los genes, las islas CpG y las secuencias repetidas (RS). (B) Diferencias entre los diferentes elementos repetidos no codificantes del genoma humano. (SINE), Elemento inter espaciador corto. (LINE). Elementos inter espaciador largo. (LTR). Secuencias que contienen repeticiones terminales largas equivalentes a retrovirus endógenos. (RS). Repeticiones simples. (LC). Repeticiones de baja complejidad. (Kpb). Kilopar de bases.

Dentro de una extensión de 100 Kpb del genoma humano homologa a las

distintas secuencias IPCR, fue posible localizar e identificar once genes (tabla 3). De estos, cinco tiene como función unión a moléculas y participan en procesos que regulan la homeostasis celular. De estos genes y tal como se describe en la tabla 3, se localizan tanto en el núcleo, en el citoplasma y en las membranas; sin embargo el gene TTL11, el cual es un miembro 11 de la familia de ligasa de tirosina a Tubulina, se localiza especialmente en los cilios y cumple la función de ligar tirosinas a la Tubulina una de las proteínas componentes de los cilios.

Tabla 3. Principales características de los genes asociados con regiones genómica de integración de provirus HTLV-I, en una extensión de 100 Kpb, en individuos seropositivos asintomáticos del Sur Occidente colombiano

Símbolo *	Código *	Locus	Tamaño pb	Exones	Función	Proceso	Localización
NEK7	140609	1q31	162.576	10	Actividad de Transferasa	Fosforilación de proteínas	Citoplasma
GPR125	166647	4p15.31	128.674	19	Unión a proteína	Señalización	Membrana
EEFSEC	60678	3q21.3	255.177	7	Factor de Elongación en la traducción	Síntesis de proteínas	Citoplasma
DNAJB8	165721	3q21.3	8.495	6	Chaperona molecular	Unión a proteínas de choque térmico	Citoplasma
GATA2	2624	3q21.3	8.495	6	Unión de iones de zinc	Transcripción	Núcleo
ERCC4	2072	16p13.12	32192	11	Unión a DNA de cadena única	Estabilidad de telómeros	Núcleo
DAB2IP	153090	16p13	1.634.818	17	Activador de actividad de GTPasa	Transducción de señales	Membrana
TTLL11	158135	9q33.2	271681	8	Actividad de ligasa de tirosinas a Tubulina	Modificación de proteínas	Citos
A2BP1	54715	16p13.2	1.634.818	12	Unión a proteínas	Procesamiento de mRNA	Núcleo

(a). (NEK7 NIMA), Cinasa 7 del gen "nunca en mitosis". (GPR125), Receptor 125 acoplado a proteína G. (EEFSEC), Factor de elongación eucariótico específico de tRNA de seleniocisteína. (DNAJB8), Miembro 8 de la subfamilia B DNAJ homólogo a Hsp40. (GATA2), Proteína 2 de unión a GATA. (ERCC4), Proteína del grupo de complementación 4 de deficiencia de reparación por excisión. (DAB2IP), Proteína de Interacción DABB2. (TTLL11), Miembro 11 de la familia de ligasa de tirosina a Tubulina. (A2BP1), Proteína 1 de unión a Ataxina.)
(b) Código de acceso al Entrez Gene (<http://www.ncbi.nlm.nih.gov/sites/entrez>)

Una comparación entre las medianas de los valores de cada uno de los miembros de esta subfamilia Alu, reveló diferencias estadísticamente significantes entre ellos ($p < 0,001$). Para los otros grupo de elementos como LINE, LTR, Transposones DNA y repeticiones simples y de baja complejidad, no se registraron diferencias estadísticamente significante excepto en el caso de las secuencias LINE, L1 y L2 ($p < 0.001$).

Discusión

En este trabajo se realizó un análisis molecular fino y a nivel genómico de la integración del HTLV-I en linfocitos de sangre periférica de individuos seropositivos asintomáticos. De los cuatro sujetos incluidos, el de menor edad fue el que tuvo la expansión clonal más baja que los otros. A partir de las características de la integración observada en los asintomáticos

incluidos en el estudio, se concluye que la integración ocurrió en zonas del genoma con baja densidad de genes y una elevada frecuencia de secuencias SINE. En su conjunto, los resultados obtenidos, permiten postular que ciertas características de la dinámica de la expansión clonal y de las regiones del genoma humano asociado con integraciones provirales, son covariables útiles para definir un estado asintomático.

Aunque la infección por el HTLV-I ha sido epidemiológicamente asociada con ATL y PET/MAH, existen todavía aspectos que se conocen muy poco con relación a la evolución de la infección viral en asintomáticos seroconvertidos y portadores de largo plazo o en aquellas personas que podrían desarrollar otro tipo de condiciones patológicas por causa de la infección viral (23). En estudios previos se

postula que la infección progresa mediante expansión clonal de linfocitos infectados, aunque son muy pocos los datos que ligan esta variable con la carga proviral medida como el número de copias de ADN proviral/103 linfocitos, y cuyo comportamiento es variable dependiendo del modo de infección y el progreso hacia la transición a ATL (24)

El progreso de la condición de asintomático hacia algunas de las patologías asociadas con la infección viral, es largo y se presenta con una frecuencia baja en la población de infectados. Uno de los mecanismos responsables por este hecho, tiene que ver con el estado inmunológico y la expresión de la proteína viral tax que promueve, vía interacción con otras proteínas celulares, una desregulación de la homeostasis del ciclo celular linfocitario y de la respuesta celular a reparación de daños en el ADN, esta propuesta es reforzada en este trabajo pues los cuatro sujetos asintomáticos analizados, tuvieron en el plasma, anticuerpos anti-Tax circulantes detectados, mediante western blot (24,25). Previamente, se ha demostrado que linfocitos CD4+/CD8+ infectados con HTLV-I, tienen patrones similares de expansión clonal entre seropositivos sin fenotipo de malignidad. Sin embargo, en la expansión espontánea de clones CD8+ que ocurre en algunos asintomáticos, existe una asociación con los niveles de expresión de mRNA de la proteína viral tax (26). Estos hallazgos podrían explicar dos de los efectos de la acción viral, uno de ellos es que la infección previene la muerte celular en el seropositivo asintomático mientras se

reclutan las clonas, en la última fase del ciclo celular; el otro caso sería que el ciclo celular y no la apoptosis promovería la acumulación de defectos genéticos asociados que desencadena el fenotipo leucémico. Como conclusión general de este estudio y otros, es posible postular que la infección por el HTLV-I juega un papel crucial en la expansión de linfocitos infectados y en la supervivencia incrementada de células CD4+ in vivo; además sugieren que la expansión clonal está contrabalanceada por factores que limitan la carga viral; ello podría ocurrir vía una subregulación mediada por tax de la proliferación de linfocitos infectados y estaría asociada con otras proteínas virales como rex y p30II (27-30). Así los linfocitos sobrevivientes pueden ser susceptibles de cambios genéticos que conllevarían hacia una condición patológica

A pesar de que los marcadores inmunológicos y virales ya descritos, son útiles en el momento de definir el estado asintomático, no revelan a nivel genómico la dinámica del proceso de integración proviral. De acuerdo con nuestros resultados, la mayor frecuencia de eventos de integración en los asintomáticos incluidos en este trabajo, ocurrió en zonas del genoma humano altamente ricas en secuencias repetidas y con baja densidad de genes e islas CpG. Es importante destacar que el 90,2% de las secuencias del genoma humano asociadas a integraciones virales en una extensión de 100 Kpb correspondieron a elementos repetidos no codificantes en los que AluSx junto con L2 fueron las más frecuentes. Una consecuencia de este estilo integracional, sería la conservación de la estructura de

aquellos genes localizados en la vecindad de provirus, impidiendo que ocurran con elevada frecuencia, eventos de mutación integracional, y de esta forma promover una regulación de la homeostasis celular. Sin embargo y a pesar de que este hecho se comprobó para la gran mayoría de los eventos de integración de los sujetos que se estudiaron, en uno de ellos fue posible determinar una integración proviral dentro del intrón siete del gen que codifica por el receptor 125 acoplado a proteína G, la proteína juega un papel fundamental en la transducción de señales vía proteína G; el mRNA transcrito de este gene, se sobre expresa nódulos linfáticos y en células provenientes de leucemia crónica mielogénica (K562) y leucemia promielocítica (hl60) (datos tomados de “GNF Expression Atlas 2” a partir de los análisis de expresión en los chips de ADN U133A and GNF1H); Con base en estos análisis, es posible proponer que teóricamente este evento de mutación integrativa produciría una alteración de la expresión del mRNA correspondiente en las células antes descritas. Es posible entonces hipotetizar que en este asintomático, como consecuencia de mutación insercional, se produciría un disturbio en la regulación de aquellos procesos celulares asociados con la transducción de señales mediada por proteínas G. Se podría postular que en este caso, se estaría incrementando el riesgo a desarrollar condiciones patológicas asociadas con la infección por el HTLV-I.

Estudios recientes sobre la integración del HTLV-I in vitro, mostraron que las características de los microambientes genómicos adyacentes a este provirus

son muy semejantes a los del virus de la leucemia aviar (ASLV) (31-34). Las regiones comparadas correspondieron a secuencias génicas y a regiones promotoras de la transcripción ricas en islas CpG; sin embargo con relación al HIV-1, se observaron diferencias (33); aunque estos resultados parecen no concordar con los que se obtuvieron en este trabajo, es importante destacar que los datos nuestros provienen de linfocitos naturalmente infectados, mientras que aquellos in vitro, se obtuvieron mediante cotransfección con vectores gamarretrovirales o derivados de MLV. El análisis comparativo entre estas características en su conjunto con relación a su distribución en el genoma humano, mostró diferencias estadísticamente significativas (35-37).

El genoma de los eucarióticos superiores está literalmente plagado de copias de elementos entre los que se incluyen Transposones ADN, elementos interesparadores cortos y largos (SINE y LINE) y pseudogenes procesados. Sin embargo de los SINE humano los más frecuentes son las de tipo Alu, que es uno de los más exitosos elementos móviles pues tiene aproximadamente 1 millón de copias en el genoma humano. Algunas de ellas, intervienen dinámicamente remodelando el genoma humano y además son la causa de mutación insercional que es la base molecular de muchas enfermedades. (38,39). La elevada frecuencia de elementos SINE y especialmente los AluSx, podría significar que la integración ocurre en zonas cuya densidad de genes es baja y que los efectos de inserción proviral no provocarían más que una

reconfiguración local de la estructura de la cromatina asociada con los eventos de integración registrados en este estudio. Esta explicación se sustenta en el hecho que tanto la expansión clonal como la carga proviral, son covariables que podrían regular la integración vía desregulación del ciclo celular, lo que produciría efectos de mediano y corto rango sobre la cromatina asociada con la integración (40-45).

En este trabajo se mostró como la integración de provirus HTLV-I ocurrió en regiones del genoma humano cuya densidad de genes es baja y con una riqueza de elementos repetidos del tipo SINE y LINE. Aproximadamente el 0,27% de las enfermedades humanas causadas por mutación insercional se atribuyen a elementos retrotransponibles. Los múltiples mecanismos de alteración vía mutación insercional incluyen la recombinación, la retrotransposición mediada y la conversión génica mediada por delección, favorecen la oncogenicidad. A la luz de nuestros resultados, es posible postular que tanto la densidad de AluSx y como de VSA-B (un elemento que tiene aproximadamente 3000 copias en el genoma humano asociado esporádicamente con algunas enfermedades en humanos (46), se conviertan en covariables que permitan seguir el progreso de la dinámica de la integración hacia estados patológicos asociados.

Se estudió la integración retroviral del HTLV-I a nivel genómico como una aproximación al entendimiento mas detallado de este mecanismo crucial

que es para asegurar el establecimiento de una infección activa. Los datos pueden compararse con aquellos que provengan de pacientes con alguna condición patológica asociada con el virus. Puesto que provienen de linfocitos naturalmente infectados, son fuente de comparación con equivalentes patológicos con el fin de lograr entender de manera más detallada, como es el efecto de la integración vs las enfermedades asociadas como la ATL y la PET/MAH.

Agradecimientos: A los pacientes que contribuyeron con sus muestras de sangre a la realización del estudio. Este trabajo se realizó con fondos del proyecto cofinanciado COLCIENCIAS) y la Universidad del Valle in Cali, Colombia (Contrato RC-542-2002) y el proyecto de la convocatoria de la Vicerrectoria de Investigaciones de la Universidad del Valle (contrato CI-1576)

BIBLIOGRAFÍA

1. Edlich RF, Arnette JA, . Global epidemic of human T-cell lymphotropic virus type-I (HTLV-I): an update. J Long Term Eff Med Implants. 2003;13(2):127-40
2. Yoshida M, Seiki M, Yamaguchi K, Takatsuki K. Monoclonal integration of human T-cell leukemia provirus in all primary tumors of adult T-cell leukemia suggests causative role of human T-cell leukemia virus in the disease. Proc Natl Acad Sci USA 1984;81:2534-7
3. Gessain A, Barin F, Vernant JC, et al. Antibodies to human T-lymphotropic virus type-I in

- patients with tropical spastic paraparesis. *Lancet* 1985;2:407-10
4. Osame M, Janssen R, Kubota H, et al. Nationwide survey of HTLV-I-associated myelopathy in Japan: association with blood transfusion. *Ann Neurol* 1990;28:50-6.
 5. Goff SP, Integration of retroviral DNA into the genome of the infected cell *Cancer Cells* 1990;2:172-8
 6. Goff SP. Genetics of retroviral integration *Annu Rev Genet* 1992;26:527-44.
 7. Cereseto A, Giacca M. Integration site selection by retroviruses. *AIDS Rev* 2004;6:13-21
 8. Hanai S, Nitta T, Shoda M, et al. Integration of human T-cell leukemia virus type 1 in genes of leukemia cells of patients with adult T-cell leukemia. *Cancer Sci* 2004;95:306-10
 9. Bushman F, Lewinski M, Ciuffi A, et al. Genome-wide analysis of retroviral DNA integration. *Nat Rev Microbiol* 2005;3:848-58
 10. Mitchell RS, Beitzel BF, Schroder AR, et al. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences *PLoS Biol* 2004;2:E234
 11. Lewinski MK, Yamashita M, Emerman M, et al. Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathog* 2006;2:e60
 12. Miller MD, Bushman FD, Orgel LE. Target-sequence preferences of HIV-1 integration complexes in vitro. *Virology* 1996;222:283-8
 13. Taganov KD, Cuesta I, Daniel R, et al. Integrase-specific enhancement and suppression of retroviral DNA integration by compacted chromatin structure in vitro. *J Virol* 2004;78:5848-55
 14. Tsukasaki K, Koeffler P, Tomonaga M. Human T-lymphotropic virus type 1 infection. *Baillieres Best Pract Res Clin Haematol* 2000;13:231-43
 15. Osame, M. Review of WHO Kagoshima meeting and diagnostic guidelines for HAM/TSP. In *Human Retrovirology: HTLV*. (Blattner W, Ed) Raven Press, New York 1990, pp 191.
 16. de Almeida MC, Silva AC, Barral A, Barral-Netto M. A Simple Method for Human Peripheral Blood Monocyte Isolation. *Mem Inst Oswaldo Cruz* 2000; 95: 221-223
 17. Takemoto S, Matsuoka M, Yamaguchi K, Takatsuki K. A novel diagnostic method of adult T-cell leukemia: monoclonal integration of human T-cell lymphotropic virus type I provirus DNA detected by inverse polymerase chain reaction. *Blood* 1994;84:3080-5
 18. Cavois M, Wain-Hobson S, Wattel E. Stochastic events in the amplification of HTLV-I integration sites by linker mediated PCR. *Res Virol* 1995;146, 179-84.
 19. Birnboim HC, Doly J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* 1979;7:1513-23.
 20. Kuhn RM, Karolchik D, Zweig AS, et al. The UCSC Genome Browser Database: update 2009. *Nucleic Acids Res* 2009;37:D755-61
 21. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science* 2001;291:1304-51

22. McPherson JD, Marra M, Hillier L, et al. A physical map of the human genome. *Nature* 2001;409:934-41
23. Wattel E, Vartanian JP, Panneitier C, Wain-Hobson S. Clonal expansion of human T-cell leukemia virus type I-infected cells in asymptomatic carriers without malignancy. *J Virol* 1995; 69:2863-8.
24. Tanaka G, Okayama A, Watanabe T, et al. The clonal expansion of human T lymphotropic virus type 1-infected T cells: a comparison between seroconverters and long-term carriers. *J Infect Dis* 2005;191:1140-7.
25. Etoh K, Tamiya S, Yamaguchi K, et al. Persistent clonal proliferation of human T-lymphotropic virus type I-infected cells in vivo. *Cancer Res* 1997; 57:4862-7.
26. Marriott SJ, Semmes OJ. Impact of HTLV-I Tax on cell cycle progression and the cellular DNA damage repair response. *Oncogene*. 2005 Sep 5;24(39):5986-95)
27. Zane L, Sibon D, Mortreux F, Wattel E. Clonal expansion of HTLV-1 infected cells depends on the CD4 versus CD8 phenotype. Clonal expansion of HTLV-1 infected cells depends on the CD4 versus CD8 phenotype. *Front Biosci*. 2009 Jan 1;14:3935-41
28. Fujii M, Niki T, Mori T, et al. HTLV-1 Tax induces expression of various immediate early serum responsive genes. *Oncogene* 1991; 6:1023-9
29. Copeland KF, Haaksma AG, Goudsmit J, Krammer PH, Heeney JL. Inhibition of apoptosis in T cells expressing human T cell leukemia virus type I Tax. *AIDS Res Hum Retroviruses* 1994; 10:1259-68.
30. Kannagi M, Matsushita S, Harada S. Expression of the target antigen for cytotoxic lymphocytes on adult T-cell-leukemia cells. *Int J Cancer* 1993; 54:582-8.
31. Wodarz D, Bangham CRM. Evolutionary dynamics of HTLV-I. *J Mol Evol* 2000; 50:448-55.
32. Wattel E, Cavrois M, Gessain A, Wain-Hobson S. Clonal expansion of infected cells: a way of life for HTLV-I. *J Acquir Immune Defic Syndr Hum Retrovirol* 1996; 13:S92-9
33. Derse D, Crise B, Li Y, Princler G, Lum N, et al. Human T-cell leukemia virus type 1 integration target sites in the human genome: comparison with those of other retroviruses. *J Virol* 2007;81: 6731-6741
34. Leclercq I, Mortreux F, Cavrois M, et al. Host sequences flanking the human T-cell leukemia virus type 1 provirus in vivo. *J Virol* 2000;74: 2305-2312
35. Ozawa T, Itoyama T, Sadamori N, et al. Rapid isolation of viral integration site reveals frequent integration of HTLV-1 into expressed loci. *J Hum Genet* 2004;49: 154-165
36. Hanai S, Nitta T, Shoda M, et al. Integration of human T-cell leukemia virus type 1 in genes of leukemia cells of patients with adult T-cell leukemia. *Cancer Sci* 2004;95: 306-310
37. Doi K, Wu X, Taniguchi Y, et al. Preferential selection of human T-cell leukemia virus type I provirus

- integration sites in leukemic versus carrier states. *Blood* 2005;106: 1048-53
38. Novick GE, Batzer MA, Deininger PL, Herrera R. The mobile genetic element Alu in the human genome. *Bioscience* 1996;46: 32-41
39. Smit AF. The origin of interspersed repeats in the human genome. *Curr Opin Genet Develop* 1996;6:743-748.
40. Dewannieux M, Esnault C, Heidmann T. LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet* 2003;35:41-8
41. Britten, RJ. Mobile elements inserted in the distant past have taken on important functions. *Gene* 1997;205: 177-82.
42. Englander EW, Howard BH. Nucleosome positioning by human Alu elements in chromatin. *J Biol Chem* 1995;270: 10091-96
43. Heller H, Kammer C, Wilgenbus P, Doerfler W.. Chromosomal insertion of foreign (adenovirus type 12, plasmid, or bacteriophage lambda) DNA is associated with enhanced methylation of cellular DNA segments. *Proc Natl Acad Sci USA* 1995;92: 5515-19
44. Cox GS, Gutkin DW, Haas MJ, Cosgrove DE. Isolation of an Alu repetitive DNA binding protein and effect of CpG methylation on binding to its recognition sequence. *Biochim Biophys Acta* 1998;1396: 67-87
45. Sandelin A, Bailey P, Bruce S, et al. Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. *BMC Genomics* 2004; 5:99.
46. Wang H, Xing J, Grover D, Hedges DJ, Han K, Walker JA, Batzer MA. SVA elements: a hominid-specific retroposon family *J Mol Biol.* 2005 ;354(4):994-1007)