

Epidemiología y bioinformática en el estudio de la Leucemia Linfoma de Células T del Adulto asociada a la infección con VLHT-1

Using epidemiology and bioinformatics for studying HTLV-1 infection-associated adult T-cell leukemia

Mercedes Salcedo-Cifuentes¹, Oscar Restrepo² y Felipe García-Vallejo¹

1 Laboratorio de Biología Molecular y Patogénesis. Departamento de Ciencias Fisiológicas. Escuela de Bacteriología y Laboratorio Clínico. Escuela de Ciencias Básicas. Facultad de Salud. Universidad del Valle. Cali. Colombia. mercysal2003@yahoo.com y labiomol@gmail.com.

2 Escuela de Ingeniería de Sistemas. Facultad de Ingenierías. Universidad del Valle. Cali. Colombia. orestrepov@gmail.com.

Recibido 31 Mayo 2010/Enviado para Modificación 20 Enero 2011/Aceptado 5 Febrero 2011

RESUMEN

Objetivos Establecer la relación entre el número de provirus VLHT-1 y las características de la cromatina adyacente en casos de Leucemia Linfoma de Células T del Adulto.

Metodología Se realizó una revisión sistemática y un metaanálisis de la literatura pública que considero como variables de estudio los provirus por cromosoma y características estructurales y funcionales de la cromatina adyacente a los sitios de integración. La concordancia entre los resultados de la evaluación que emitieron dos expertos fue evaluada con el coeficiente de Spearman Rho. Se evaluó el sesgo de publicación mediante el gráfico de embudo y el estadígrafo Egger. De acuerdo con los resultados de la evaluación de la heterogeneidad se aplicó el modelo de efectos fijos para la combinación de los resultados de las integraciones que ocurrieron en: secuencias codificantes y secuencias codificantes de acuerdo con su función molecular.

Resultados La concordancia entre expertos evaluadores fue de 0,7. No se encontró sesgo de publicación. Se determinó homogeneidad entre los estudios seleccionados ($p > 0,05$). El provirus VLHT-1 se integró en secuencias en regiones teloméricas y subteloméricas. La combinación de los resultados mostró una integración sitio dirigida hacia regiones codificantes del genoma humano ($p < 0,05$).

Conclusión En su conjunto los resultados permiten concluir que la integración proviral no es al azar en LCCTA; ésta ocurrió en regiones reguladoras o de control; que explicarían algunos de los procesos moleculares involucrados en leucemogénesis.

Palabras Clave: Virus Linfotrópico de Células T Humanas Tipo 1, integración viral, genoma humano, epidemiología, bioinformática, Leucemia-Linfoma de Células T del Adulto (*fuentes: DeCS, BIREME*).

ABSTRACT

Objectives Establishing a correlation between the number of HTLV-1 provirus and the characteristics of the genomic environment in ATL cases.

Methodology A systematic search was made of publications as well as a meta-analysis of the pertinent literature considering proviruses per chromosome and structural and functional characteristics of flanking chromatin regions as variables. The concordance of experts' study was evaluated by Spearman Rho correlation. Publication bias was analysed by funnel plot and the Egger statistigrapher. A fixed effects model was applied according to heterogeneity evaluation to combine the results of integration occurring in coding sequences as well as coding sequences according to their molecular function.

Results The expert concepts' Kappa index was 0.7 and no publication bias was observed. The meta-analysis result was homogeneous ($p>0.05$). HTLV-1 integration was directed towards several chromosomes' telomeric and subtelomeric regions. The combination of published results in the articles which were analysed supported the hypothesis of integration events being site-directed towards coding regions of the human genome ($p<0.05$). Moreover, the groups of genes having enzymatic and receptor functions was statistically significant.

Conclusion The results led to concluding that HTLV-I integration in the ATLL cases analysed here was not random but was directed towards regulatory regions. Such results could help to explain the role of some processes involved in leukemogenesis.

Keys Words: HTLV-1, proviral integration, human genome, genomic epidemiology, bioinformatics, ATLL (source: MeSH, NLM).

La infección por el Virus Linfotrópico Humano tipo 1 (VLHT-1), es un problema actual de salud pública global (1), ya que infecta de 15 a 20 millones de personas en todo el mundo (2), siendo Centroamérica, América del Sur y el Caribe áreas con una alta prevalencia de la infección en las cuales se observan conglomerados de regiones endémicas (3). Colombia registra seroprevalencias del 1,0 % al 2,0 %; aunque algunas zonas del Pacífico y de la región Caribe se han obtenido valores de 7,5 % y 10,0 % respectivamente (4). Este retrovirus se puede transmitir de dos maneras: horizontalmente (sexual, sangre y hemoderivados) (5) y verticalmente (madre al niño) principalmente por la leche materna (6).

Existe una amplia gama de manifestaciones clínicas asociadas a la infección por el VLHT-1 entre las que se incluyen infecciones asintomáticas, además de desórdenes linfoproliferativos y neurológicos (2,4). La Leucemia Linfoma de Células T en Adulto (LLCTA) asociada epidemiológicamente con la infección, es un proceso de tumorigénesis en el cual se ha logrado demostrar que la carga proviral y el patrón de integración del cADN viral a

nivel cromosómico determinan la aparición de los estadios más severos de la enfermedad (7,8).

Inicialmente se determinó que aquellas regiones del ADN humano con Kco/t más bajos eran las preferidas para la integración del cADN viral (9). Posteriormente, se identificó que el provirus se integraba con preferencia en las bandas R, con mayor frecuencia en ciertos cromosomas y loci que diferían según la condición clínica y enfermedad asociada (10,11). DE los estudios de secuencias de nucleótidos de aquellas regiones del genoma humano adyacentes al provirus VLHT-1, se observó una alta densidad de islas CpG y elevado contenido en G:C (4,12). En su conjunto, todos estos resultados han permitido plantear la hipótesis que la integración proviral no es un evento al azar, si no que es un proceso en el cual, algunas características de la cromatina estarían condicionando la selección de los sitios de integración (13-16).

Esta hipótesis ha sido acogida por varios investigadores quienes además de análisis experimentales han apoyado sus hallazgos en estudios bioinformáticos realizados con la información registrada en las bases de datos de la secuencia del genoma humano de dominio público (17-19). La información actual es que el provirus tiene una tendencia a integrarse en los cromosomas de gran tamaño (14,17,18,21), en regiones con alta densidad de islas CpG y de genes involucrados en el control del ciclo celular, apoptosis y transducción de señales (14,16-18,20). Sin embargo, el poder de estos estudios ha sido limitado por el pequeño número de sitios de integración estudiados. Adicionalmente, para algunos investigadores (19,20) la distribución genómica de los sitios de integración no ha sido ampliamente estudiada en el contexto de la infección persistente in vivo, en donde las células infectadas se someten a fuerzas de selección adicionales, tales como la respuesta inmune, lo cual podría influenciar la residencia de la partícula viral en la región seleccionada inicialmente para su integración. Por estas razones se planteó realizar, una revisión sistemática y un meta-análisis, con el fin de analizar la relación existente entre el número de integraciones provirales en regiones específicas del genoma humano y las características de su ambiente genómico en casos de Leucemia Linfoma de Células T del Adulto.

MATERIAL Y MÉTODO

Revisión de Literatura

La búsqueda y selección de los artículos fue conducida simultáneamente

para estudios que combinaran en su metodología tanto técnicas moleculares como de análisis bioinformáticos con el objetivo de caracterizar las regiones adyacentes a la integración del VLHT-1 en casos Leucemia Linfoma de Células T del Adulto (LLCTA). Las palabras claves que orientaron la búsqueda bibliográfica en las bases de datos de PubMed, OVID y MEDLINE desde el 2002 hasta el 2008 fueron: HTLV-1, Viral Integration, Adult T-cell leukemia/lymphoma, "Ligation Mediated Polymerase Chain Reaction" - LMPCR- e "Inverse Polymerase Chain Reaction -IPCR- dos metodologías comúnmente utilizadas para amplificar regiones de genoma adyacentes a LTR provirales.

Criterios de inclusión de los trabajos preseleccionados

Se incluyeron en esta investigación: trabajos originales publicados en revistas de reconocida trayectoria registrados en bases de datos en inglés y español. Células mononucleares de sangre periférica o aspirado de nódulos recolectados a partir de pacientes con diagnóstico confirmado de LLCTA, las secuencias vecinas al sitio de integración del provirus, fueron la fuente de las unidades de análisis. Los métodos experimentales usados para la extracción de las secuencias adyacentes al provirus debían ser IPCR o LMPCR. Además, los análisis bioinformático debían considerar como fuente de información las bases de datos del GeneCard (<http://www-bimas.cit.nih.gov/cards//index.shtml>) y de Genoma Browser (<http://www.genome.ucsc.edu/>). Como criterio final para la inclusión de los estudios en el meta-análisis, se consideró que los resultados bioinformáticos partieron de homologías ≥ 90 % en una ventana de apertura de 100 a 500 Kb.

Obtención y análisis de los trabajos

De forma independiente, dos expertos evaluaron la pertinencia de la preselección calificando, mediante un formulario estándar, la calidad científica, la comparabilidad de las técnicas experimentales empleadas por cada grupo de autores y la de las muestras origen de las secuencias en estudio considerando la clasificación de adecuada, inadecuada o incierta, de acuerdo con los criterios del "lymphoma study groups Japonés"⁷. Aplicando el coeficiente de Spearman Rho ($p < 0,05$), se evaluó la concordancia entre los resultados de las valoraciones individuales de cada uno de los expertos.

Evaluación del sesgo de publicación y heterogeneidad

El sesgo de publicación se evaluó de acuerdo con la asimetría del gráfico en embudo (funnel plots). Éste se complementó con el estadígrafo de Egger de StatDirects® que valora la asimetría del gráfico, considerando significativo

un valor de p menor de 0,05. Se analizó la homogeneidad entre los estudios mediante el diagrama de bosque y la prueba de I² de StatDirects®. No se halló heterogeneidad estadísticamente significativa, por lo cual se usó el modelo de efectos fijos para combinar los resultados en el meta-análisis.

Variables sometidas a estudio y procesamiento de datos

La variable de efecto valorada fue: número de integraciones provirales. Las variables independientes fueron secuencias codificantes y no codificantes; secuencias codificantes según función molecular; secuencias repetitivas y asignación cromosómica. El análisis estadístico se llevó a cabo usando el paquete estadístico StatDirects®.

RESULTADOS

Con la estrategia de búsqueda y selección de los trabajos preestablecida se encontraron cinco artículos (Tabla 1). Uno de ellos, el de Miyazaki-2007 (21), fue eliminado para el análisis final por considerar que el número de sitios de integración tenía una representación muy baja en el total de integraciones combinadas (7/139). La concordancia en la valoración de cada uno de los expertos, mediante el Índice Kappa Kohen fue de 0,7.

Tabla 1. Características de los estudios preseleccionados por los dos evaluadores.

Autor	Año de publicación	Fuente celular de las secuencia	IISC §	Técnicas experimental utilizadas	Uso de las secuencias
Hanai	2004	CMNSP ¥	20/25 (0,8)	LMPCR £	Caracterización de la secuencia por análisis bioinformático
Osawa	2004	CMNSP	15/35 (0,43)	LMPCR	Caracterización de la secuencia por análisis bioinformático
Doi	2005	CMNSP	20/59 (0,34)	IPCR ‡	Caracterización de la secuencia por análisis bioinformático
Miyazaki*	2007	CMNSP y NL £	15/35 (0,43)	IPCR	Caracterización de la secuencia por análisis bioinformático
Salcedo-Cifuentes	2008	CMNSP y NL	31/118 (0,26)	IPCR	Caracterización de la secuencia por análisis bioinformático

(*)Trabajo eliminado del meta-análisis por número de integraciones poco relevante. (¥) CMNSP (Células mononucleares de sangre periférica). (£) NL (Nódulo linfático). (§) IISC (Índice de integración en secuencias codificantes = Número de integraciones en secuencias codificantes/Número total de integraciones). (E) LMPCR ("Ligation Mediated Polymerase Chain Reaction").(‡) IPCR ("Inverse Polymerase Chain Reaction")

La proporción combinada de integraciones en secuencias codificantes fue del 33,4 % (IC 95 % = 27,6-39,4 %) (Tabla 2). El gráfico en embudo ("funnel plot") (Figura 1) no mostró sesgo de publicación, este resultado fue corroborado con la prueba de asimetría de Egger, (7,44 IC95 %: -9,40-24,28; $p=0,1976$). Se observó una baja heterogeneidad en la proporción de integraciones provirales en secuencias codificantes ($I^2=88,9$ %; CI95 % = 70,4 % a 93,9 %), por lo cual se aceptó trabajar con el modelo de efectos fijos para evaluar la proporción combinada (Figura 2). Adicionalmente, los resultados del meta-análisis según función molecular permitieron establecer que había homogeneidad entre las integraciones provirales orientadas hacia secuencias codificantes con función enzimática y receptora en los cuatro trabajos finalmente seleccionados (Figura 3). La distribución cromosómica de las integraciones globales mostró una tendencia hacia los cromosomas grandes y medianos (81,4 % del total de las integraciones combinadas). Varias de las integraciones se ubicaron dentro o vecinos a: cuatro clúster de histonas (cromosomas 1, 2, 6 y 22); dos clúster de oncogenes localizados en los cromosomas 6 y 8; un clúster de genes reguladores (cromosoma 14), genes codificantes para interferón y proteínas estimuladoras del interferón (cromosoma 16) y el gen de la ubiquitina (cromosoma 21). Este último corresponde a un gen que participa en la remodelación de la cromatina por mecanismo epigenético.

Tabla 2. Distribución de las proporciones de integraciones provirales del VLHT-1 según autor y proporción global combinada

Autor	Proporción de integraciones en secuencias codificantes	IC _{95%}	% pesos (fijos, al azar)
Hanai	0,80	0,59-0,93	10,79-22,83
Ozawa	0,43	0,26-0,61	14,94-24,20
Doi	0,34	0,22-0,47	24,90-26,81
Salcedo-Cifuentes	0,26	0,19-0,35	49,38-27,15
Proporción global	0,36	0,30-0,43	-

Al comparar la distribución de integraciones del cADN VLHT-1 dentro de secuencias codificantes con aquella de secuencias codificantes del genoma humano según Venter (22), no se observaron diferencias estadísticamente significativas ($p<0,05$).

El mayor porcentaje de secuencias adyacentes a provirus reportadas por los cuatro autores correspondieron a secuencias que intervienen en transducción de señal (25,45 %), metabolismos celular (18,2 %), división y ciclo celular (11,8 %) y apoptosis celular (11,0 %) (Figura 4). Con base en la

coincidencia de los hallazgos de los cuatro autores incluidos en el análisis final, se hizo una búsqueda detallada en la base de datos del GeneCard de las características de los genes identificados por medio de análisis bioinformático en las regiones adyacentes a provirus (Tabla 3).

Figura 1. Evaluación del sesgo de publicación a través del gráfico en embudo (funnel plot)

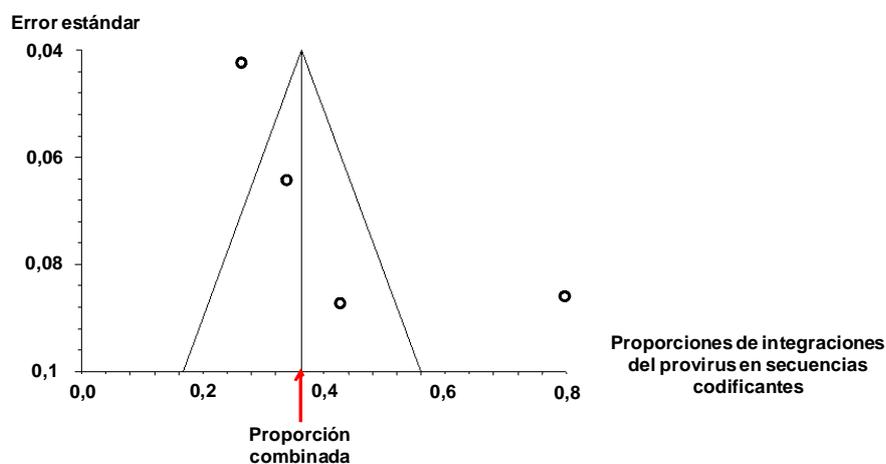
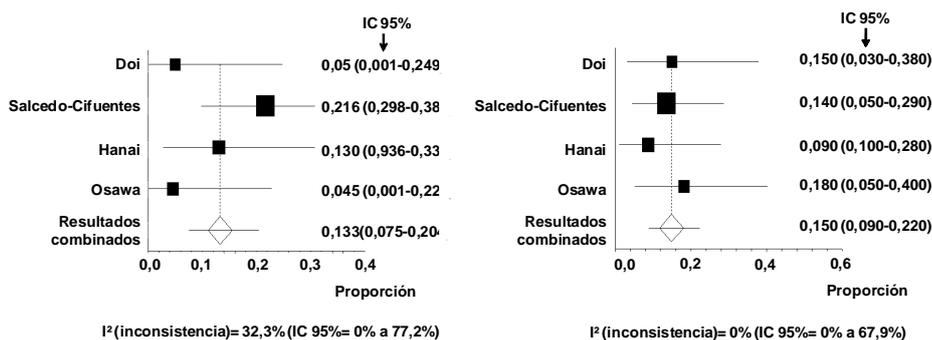


Figura 2. Árbol de proporción del VLHT-1 de secuencias codificantes vs el número total de integraciones provirales en todas las secuencias adyacentes al provirus.



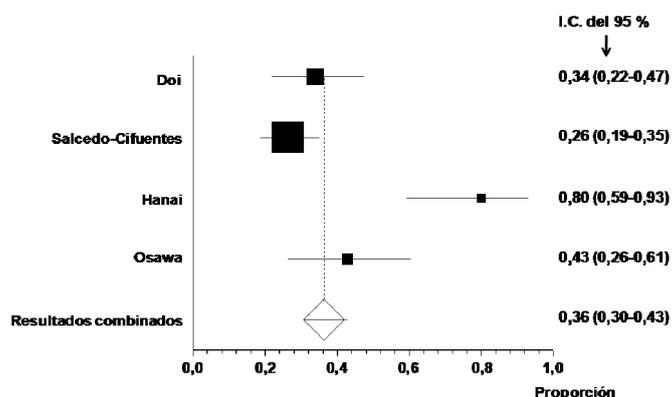
DISCUSIÓN

De acuerdo con hallazgos previos (8,14,23), tanto la integración como la transcripción del genoma viral inducirían cambios conformacionales en la estructura de la cromatina hospedera que podrían explicar algunos de los procesos patológicos de la leucemia. En este sentido se ha determinado que en las etapas tempranas de la infección, la transcripción en la célula hospedera es controlada por el efecto de la proteína Tax (8); mientras que la replicación viral tardía es controlada por proteínas accesorias del virus entre las que se incluye la p30 (24).

Tabla 3. Características de algunos de los genes identificados en las secuencias adyacentes al provirus VLHT-1 en una ventana de apertura entre 100 a 500 Kb

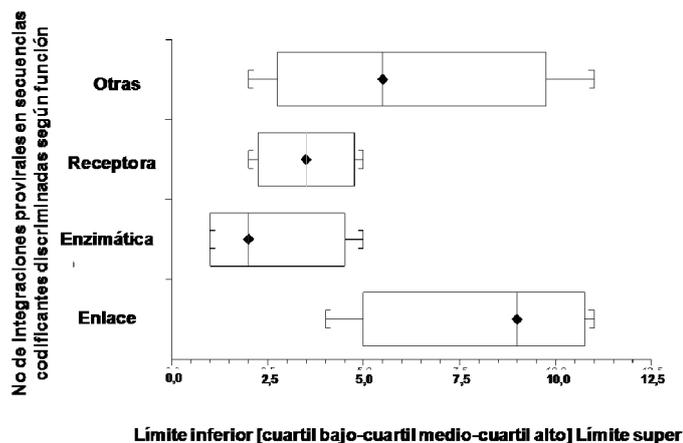
Símbolo del Gen	Locus	Descripción del proceso biológico	Orientación
TRIM33	1p13.1	Codifica para una proteína inhibitoria de la transcripción	Negativa
IRF8	16q24.1	Transcripción del interferon	Positiva
CD46	1q32	Codifica para una proteína que regula la activación del complemento	Positiva
PTPN5	2q14.2	Proteína que media la adhesión de caderinas	Positiva
GPHN	14q23.3	Proteína de enlace a nucleótido	Negativa
ANK1	8p11.1	Proteína constituyente estructural del citoesqueleto de eritrocitos	Negativa
SLC4A3	2q36	Proteína responsable del transporte de electrones	SD ^s
SDCBP	8q12	Proteína de enlace al receptor 5 de interleuquina	Positiva
IRF8	16q24	Proteína	Negativa
CDKN2A	9p21.3	Proteína que interviene en la fragmentación del ADN para la apoptosis celular	Positiva
USP25	21q21.1	Proteína que interviene en la degradación de proteínas	Negativa

Figura 3. Árbol de proporción de integraciones del VLHT-1 en: A) secuencias codificantes y B) Secuencias codificantes discriminadas según función enzimática y receptora.



La tendencia a la integración subtelomérica y telomérica registradas en el meta-análisis, pueden ser determinantes de la estabilidad cromosómica y la carcinogénesis (26). En el LLCTA se detiene el acortamiento crítico por la activación de la Telomerasa y de sus proteínas de enlace (TRF1, TRF2 y TIN2) (27). Dicha activación es dirigida, en estadios tempranos de la infección, por la proteína Tax (8). En los procesos crónicos se ha observado que la activación es mediada por la expresión de Interleukina 2 (IL2) (28, 29), localizado en el locus 4q26.28 que en este estudio mostró varios sitios de integración proviral. Surge como una explicación probable a este hecho, que la presencia de estos provirus podría generar un efecto trans-activador a distancia (8,27).

Figura 4. Distribución del número de integraciones del provirus VLHT-1 según función molecular.



En el análisis sobre los genes localizados en aquellas regiones en las que se registró un elevado número de provirus, se pudo caracterizar el potencial papel de algunos de ellos en la oncogénesis asociada. El que codifica para la Isoforma 1 de la Ankirina Eritrocitaria (ANK1) es un gen cuya unión a la proteína tax podría inhibir la formación del complejo IKB-NF-KB promoviendo la traslocación nuclear de la actividad de este complejo que conllevaría a una estimulación de la proliferación de los linfocitos T.

Este proceso explicaría la acción de la proteína oncogénica del virus en etapas tempranas de la infección; sin embargo hay reportes que muestran como el mismo efecto parecería mantenerse por la acción de proteínas accesorias del genoma retroviral como p12 y p13 (30,31). Adicionalmente, el gen también podría involucrarse en la transmisión viral célula a célula por formación de sinsitios (32). Así mismo la presencia de otros genes adyacentes a provirus como el ARL15 ("ADP-ribosylation factor-like 5") es importante pues liga GTP que activa la expresión de las proteínas Ras (33); el gen de Sintenina (SDCBP), que codifica para proteínas que participan en los procesos de transmisión de señales a nivel de citoesqueleto (32). El factor 8 regulador del interferón (IRF-8), el cual pertenece a un grupo de proteínas reguladoras de la respuesta inmune (34,35); el gen USP25 relacionado no solo con la degradación de proteínas sino también con procesos de control celular, tráfico de proteínas y remodelación de la cromatina, entre otros (36); todos ellos estarían participando de redes de expresión de genes cruciales para el control de procesos asociados con la leucomogénesis y de descontrol del ciclo celular que se asocian con LLCTA.

Identificar las posibles redes alteradas por la integración del provirus así como las características del ambiente genómico local, proporcionaría información de la enfermedad que no es directamente discernible a partir de las secuencias únicas de los genes encontrados en los sectores cromatínicos preferidos por el provirus. Aquí está lo interesante de los estudios de la genómica aplicada, en donde el diseño de técnicas experimentales, computacionales y la aplicación de diferentes modelos epidemiológicos permiten extraer la información biológica contenida en la estructura y dinámica de la red o redes de genes y el ambiente asociado al síndrome linfoproliferativo maligno con el cual cursa aproximadamente el 4% de la población infectada por el VLHT-1 (2,4,8)♣

Agradecimientos: A todos los pacientes con Leucemia Linfoma de células T del Adulto (LLCTA)/VLHT-1 que participaron en los estudios incluidos en este subanálisis, quienes con su colaboración permitieron ampliar el conocimiento de la LLCTA en el contexto de la infección por el VLHT-1 a nivel mundial. Este trabajo se realizó en el marco de un proyecto financiado por Vicerrectoría de Investigaciones de la Universidad del Valle, bajo el acta de trabajo y compromiso CI-1576

Conflictos de interés: Ninguno.

REFERENCIAS

1. Proietti FA, Carneiro-Proietti AB, Catalan-Soares BC, Murphy EL. Global epidemiology of HTLV-I infection and associated diseases. *Oncogene*. 2005;24:6058-68..
2. Edlich RF, Arnette JA, Williams FM. Global epidemic of human T-cell lymphotropic virus type-I (HTLV-I). *J Emerg Med*. 2000; 18:109-19.
3. Blattner WA, Gallo RC: Epidemiology of HTLV-I and HTLV-II infection. In *Adult T-cell leukemia*. Edited by: Takahashi K. New York, Oxford University Press; 1994: 45-90.
4. García-Vallejo F. Caracterización molecular y genómica del proceso de integración del provirus del virus linfotrópico humano (HTLV) tipo I. *Rev Acad Colomb Cienc*. 2006; 30:155-70.
5. Moriuchi M, Moriuchi H, Seminal Fluid Enhances Replication of Human T-Cell Leukemia Virus Type 1: Implications for Sexual Transmission. *J Virol*. 2004; 78: 12709-711.
6. Nyambi PN, Ville Y, Louwagie J, et al. Mother-to-child transmission of human T-cell lymphotropic virus types I and II (HTLV-II/II) in Gabon : A prospective follow-up of 4 years. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1996; 12: 187-92.
7. Shimomaya M, Lymphoma Study Group (1984-87). Diagnostic criteria and classification of clinical subtypes of adult T-cell leukaemia-lymphoma. *Br J Haematol*. 1991; 79:428-37.
8. Yasunaga JI, Matsuoka M. Human T-Cell Leukemia Virus Type I Induces Adult T-Cell Leukemia: From Clinical Aspects to Molecular Mechanisms. *Cancer Control*. 2007; 14: 133-40.
9. Rynditch A, Zoubaka S, Tsybaa L, Tryapitsina-Guleya N., Bernardia G, The regional integration of retroviral sequences into the mosaic genomes of mammals. *Gene*. 1998; 22: 1-16.
10. Richardson JH, Rose NJ. Chromosomal positioning of human Tlymphotropic type I proviruses by fluorescent in situ hybridisation. *Jour Viro Meth*. 2001; 93:65-74.
11. Glukhova LA, Zoubak SV, Rynditch AV, Miller GG, Titova IV, Vorovyeva N, et al. Localization of HTLV-1 and HIV-1 proviral sequences in chromosomes of persistently infected cells. *Chromosome Res*. 1999; 7:177-83.
12. Zoubak S, Richardson J, Rynditch AV, Hollsberg P, Hafler DA, Boeri E, et al. Regional specificity of HTLV-I proviral integration in the human genome. *Gene*. 1994; 143:155-63.
13. Tsukasaki K, Tsushima H, Yamamura H, Hata T, Murata K, Maeda T, et al. Integration patterns of HTLV-I provirus in relation to the clinical course of ATL: frequent clonal change at crisis from indolent disease. *Blood*. 1997; 89: 948-56.
14. Leclercq I, Mortreux F, Gabet AS, Jonson CB, Wattel E. Basis of HTLV type I site selection. *AIDS Res Hum Retrovi*. 2000; 16:1653-59.
15. Vine AM, Witkover AD, Lloyd AL, Jeffrey KJM, Siddiqui A, Marshall SEF, et al. Polygenic control of human T Lymphotropic virus Type I (HTLV-I) provirus load the risk of HTLV-I Associated Myelopathy Tropical Spastic Paraparesis. *J Infec Dis*. 2002; 186: 932-39.
16. Hanai S, Nitta T, Shoda M, Tanaka M, Iso N, Mizoguchi I, 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.
17. Doi K, Wu X, Taniguchi Y, Yasunaga J, Satou Y, Okayama A, Nosaka K, Matsuoka M. Preferential selection of human T-cell leukemia virus type I provirus integration sites in leukemic versus carrier status. *Blood*, 2005; 106: 1048-53.
18. Derse D, Crise B, Li Y, Princler G, Lum N, Stewart CF, 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: 6731-41.
19. Mitchell RS, Beitzel BF, Schroder AR, Shinn P, Chen H, Berry CC, et al. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol*. 2004; 2:1127-37.

20. Salcedo-Cifuentes M, Cabrera J, Cuesta Y, Carrascal E, Eizuru Y, Domínguez MC, et al. Caracterización Molecular, expansión clonal y estructura del DNA de los sitios de integración del provirus linfotrópico humano Tipo I (HTLV-I) en casos de ATL. *Biomédica*. 2009; 29: 218-31.
21. Miyazaki M, Yasunaga JI, Taniguchi Y, Tamiya S, Nakahata T, Matsuoka M. Preferential Selection of Human T-Cell Leukemia Virus Type 1 Provirus Lacking the 5' Long Terminal Repeat during Oncogenesis. *J Virol*. 2007; 81: 5714-23.
22. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science* 2001; 291: 1304-51.
23. Cavrois 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.
24. Sinha-Datta U, Datta A, Ghorbel S, Dodon MD, Nicot C. Human T-cell Lymphotropic Virus Type I Rex and p30 Interactions Govern the Switch between Virus Latency and Replication. *J Biol Chem*. 2007; 282:14608-15.
25. Osawa T, Itoyama T, Sadamori N, Yamada Y, Hata T, Tomonaga M, Isobe M. Rapid isolation of viral integration site reveals frequent integration of HTLV-1 into expressed loci *Journal of Human Genetics*. 2004; 49: 154-65.
26. Bellon M, Nicot Ch. Central role of PI3K in transcriptional activation of hTERT in HTLV-I-infected cells. *Blood*. 2008; 112: 2603-04.
27. Bellon M, Datta A, Brown M, Pouliquen JF, Couppie P, Kazanji M, Nicot C. Increased expression of telomere length regulating factors TRF1, TRF2 and TIN2 in patients with adult T-cell leukemia. *Int J Cancer*. 2006; 119:2090-97.
28. Nakamura Y, Moriuchi R, Nakayama D, Yamashita I, Higashiyama Y, Yamamoto T, et al. Altered expression of a novel cellular gene as a consequence of integration of human T cell lymphotropic virus type 1. *J Gen Virol*. 1994; 75: 2625-33.
29. Macera MJ, Szabo P, Verma RS. Chromosomal localization of HTLV-1 viral integration sites using in situ hybridization: detection of a novel IL2R fragment. *Mol Gen Genet*. 1992; 234: 466-74.
30. Albrecht B, Collins ND, Burniston MT, Nisbet JW, Ratner L, Green PL, Lairmore MD. Human T-lymphotropic virus type 1 open reading frame I p12(I) is required for efficient viral infectivity in primary lymphocytes. *J Virol*. 2000; 74: 9828-35.
31. Albrecht B, d'Souza CD, Ding W, Tridandapani S, Coggshall KM, Lairmore MD. Activation of nuclear factor of activated T cells by human T-lymphotropic virus type 1 accessory protein p12(I). *J Virol*. 2002; 76: 3493-501.
32. Ishizaki J, Okayama A, Kuroki M, Tsubouchi H. Detection of Human T-Lymphotropic Virus Type I (HTLV-I) Infection during Coculture of HTLV-I Infected and Uninfected Cells Using Inverse Long PCR. *Intervirology*. 2002; 45:164-71.
33. Taylor JM, Ghorbel S, Nicot Ch. Genome wide analysis of human genes transcriptionally and post-transcriptionally regulated by the HTLV-I protein p30 *BMC Genomics*. 2009; 10:311.
34. Takaoka A, Tamura T, Taniguchi T. Interferon regulatory factor family of transcription factors and regulation of oncogenesis *Cancer Science*. 2008. 99; 467- 78.
35. Tamura T, Tailor P, Yamaoka K, Kong, H J, Tsujimura H, O'Shea J J, Singh H, Ozato K. Regulatory Factor-4 and -8 Govern Dendritic Cell Subset Development and Their Functional Diversity. *J. Immunol*. 2005. 174: 2573-581.
36. Gao Y, Lecker S, Post MJ, Hietaranta AJ, Li J, r Volk R, Li M, Sato K, Saluja A K, Steer ML, Goldberg AL, Simons M. Inhibition of ubiquitin-proteasome pathway-mediated I β B? degradation by a naturally occurring antibacterial peptide. *J Clin Invest*. 2000; 106: 439-48.
37. Asquith B, Bangham CR. How does HTLV-I persist despite a strong cell-mediated immune response? *Trends Immunol*. 2008; 29: 4-11.