

201019006A

厚生労働科学研究費補助金

第3次対がん総合戦略研究事業

ヒトT細胞白血病ウイルス1型関連疾患における感受性遺伝子多型の同定と
発症危険群へのアプローチ

(H21-3次がん一般-004)

平成22年度 総括・分担研究報告書

研究代表者 松岡 雅雄
(京都大学ウイルス研究所・教授)

平成23 (2011) 年 3月

目 次

I. 総括研究報告

ヒト T 細胞白血病ウイルス 1 型関連疾患における感受性遺伝子多型の 同定と発症危険群へのアプローチ-----	1
研究代表者・松岡 雅雄（京都大学ウイルス研究所・教授）	

II. 分担研究報告

1. 松田 文彦（京都大学医学研究科・教授）-----	4
2. 齊藤 峰輝（琉球大学医学部・准教授）-----	6
3. 野坂 生郷（熊本大学大学院医学薬学研究部・助教）-----	9

III. 研究成果の刊行に関する一覧表・刊行物・別刷

1. 松岡 雅雄（京都大学ウイルス研究所・教授）-----	11
2. 松田 文彦（京都大学医学研究科・教授）-----	52
3. 齊藤 峰輝（琉球大学医学部・准教授）-----	84
4. 野坂 生郷（熊本大学大学院医学薬学研究部・助教）-----	101

ヒト T 細胞白血病ウイルス 1 型関連疾患における感受性遺伝子多型の同定と
発症危険群へのアプローチ

研究代表者：京都大学ウイルス研究所 教授 松岡雅雄

研究要旨

ヒト T 細胞白血病ウイルス 1 型 (HTLV-1) は成人 T 細胞白血病 (ATL), HTLV-1 関連脊髄症 (HAM) の原因ウイルスである。ATL, HAM 発症には家族内集積性があることが知られており遺伝的素因の関与が強く疑われている。本研究は全ゲノム関連解析 (GWAS) を用いてこれらの疾患感受性遺伝子多型を明らかにすることを目的としている。ATL, HAM 検体の収集を進めると共に、全ゲノム関連解析を進め、それぞれの疾患に関して候補遺伝子多型を同定しつつある。

A. 研究目的

本邦における HTLV-1 感染者数は約 108 万人と推定されており世界最大の蔓延地域であり年間約 1000 名のキャリアが ATL を発症している。ATL は予後不良であり HAM も有効な治療法が無い。ATL では家族内発症が高頻度に認められ、その発症に遺伝的背景が密接に関連していることが明らかであるが、その解析は進んでいない。本研究では ATL および HAM 発症に関与する遺伝子多型を解析し発症危険群を同定することで将来の発症予防法確立に繋げる。松岡・野坂は ATL でプロウイルスマイナス鎖にコードされる *HBZ* 遺伝子が ATL 発症に重要であることを明らかにした (PNAS, 2006)。松田は様々な病原体と疾患感受性遺伝子について解析を進め HIV-1 感染者における遺伝子多型と AIDS への進行との関連 (J Infect Dis 2006; PNAS, 2007)、デングウイルスレセプターである DCSIGN-1 プロモーター多型とデング熱重症度の関連を報告した (Nat Genet 2005)。齊藤は HLA ハプロタイプや IL-10 などが HAM 発症に関与することを報告した。これまで ATL 感受性遺伝子多型に関する報告は少なく、HAM の発症機序に関しても限られた情報しかない。既に平成 22 年度の研究においてイルミナ社の Human610k-Quad BeadChip を用いて ATL 患者 448 検体、HAM 患者 458 検体、無症候性 HTLV-1 キャリア 314 検体の網羅的タイピングをおこなった。タイピングデータを用いたメタ解析により、潜在的有意差を示す多型を同定している。本研究では以下の点を明らかにする。

1) これまで全ゲノム関連解析を行ったデータを解析し有意差の得られた

SNP に関してサンプルを更に集め再現性の検証を行う。

2) 同定された SNP から疾患発症に関連する遺伝子型を決定した後、連鎖不平衡地図ならびにハプロタイプ地図を作成し、相関解析、ハプロタイプ解析を実施して最も信頼性の高い遺伝子座を明らかにする。

3) 関連する遺伝子多型の疾患発症への関与を解析する。

B. 研究方法

SNP を用いた全ゲノム関連解析を行い、疾患関連 SNP を同定する。さらに、疾患関連 SNP の生物学的活性の解析：同定された疾患感受性 SNP が発症を促進する機序を明らかにするため、SNP がその分子に付与する機能について解析する。

(倫理面への配慮)

本研究は全ての倫理指針に基づいた倫理委員会での審議、許可を得ている。また、検体は全て匿名化する。

C. 研究結果

1) 全ゲノムジェノタイピング

ATL, HAM 症例、キャリアサンプルの収集を継続し、解析数を増やしている。我々はイルミナ社の Human610k-Quad BeadChip を用いて熊本で収集された ATL 患者 258 検体と無症候性 HTLV-1 感染者 182 検体 (Set1)、鹿児島で収集された HAM 患者 296 検体と無症候性 HTLV-1 感染者 100 検体 (Set2) の GWAS を行った。タイピング結果の品質管理として患者、対照群両群に関してアッセイ

成功率が90%以上、マイナーアレル頻度が0.01以上、対照群のジェノタイプ分布がハーディー・ワインバーグ平衡に達しているもの ($p > 10^{-7}$) を選び、条件を満たした477,250 SNPsのみを解析にまわした。サンプルの品質管理後(タイピング成功率が90%以上、さらにPLINKで遺伝的近縁度を推定後、近縁度が高いペアは片方を除去し、条件を満たした検体のみを統計にまわした)、最終的にATL患者238検体と無症候性HTLV-1感染者165検体(Set1)、HAM患者278検体と無症候性HTLV-1感染者97検体(Set2)を統計解析した。選択されたSNPsに関して、患者と対照群の間で、アレル頻度、ジェノタイプ分布を考慮に入れた統計解析を行い、患者集団内での頻度が統計学的有意に達するSNPがあるか検討した。また、患者を病態で分類した複数のサブグループ間での比較解析に着手し、HTLV-1感染と疾患の発症のどの段階でどのような遺伝子が関与しているかを詳細に調べている。

2) 二次スクリーニング(再現性の検証) 有意差の得られたSNPsについては、それを含まハプロタイプの頻度に差があるか否かを検討し、その結果Set1に関しては、 $p = 2.50 \times 10^{-7}$ をはじめとして、 $p < 1 \times 10^{-5}$ のSNPが29個、Set2に関しては $p = 2.00 \times 10^{-6}$ をはじめとして、 $p < 1 \times 10^{-5}$ のSNPが52個得られた。一次スクリーニングで得られた結果の再現性を確認するため、候補SNPsについて新たな検体を用いて二次スクリーニングを行う(Taqman法)。詳細な臨床情報を伴うATL、HAM、および無症候性HTLV-1感染者のDNAが各153、67、210検体、既に収集されているが、さらに検体数の増加に努める。結果をデータベースに取り込み統計解析を行い、一次スクリーニングと同じ基準でデータの品質管理を行う。

3) 疾患感受性遺伝子/多型の同定 二次スクリーニングでも関連性が確認されたSNPの連鎖不平衡(r^2 および D')を調べ、さらに、我々が既に所有するデータベースおよび国際HapMap計画データベースから、連鎖不平衡ブロック(LDブロック)を特定する。特定したLDブロック内に存在する(可能な限りの)全多型を同定するため、日本人32検体を用いてSNP同定を行う。遺伝子の調節領域、エクソンおよびその周辺領域内を特異的プライマーによって増幅

し塩基配列を再決定する。得られた塩基配列のデータをSNP同定のソフトウェアで解析し、SNP同定を行う。同定された新しいSNPの情報は、当研究室の遺伝子データベースに収め、遺伝子内の各多型の位置、頻度、多型間の連鎖不平衡、遺伝子毎のハプロタイプと集団における推定頻度などのデータを蓄積する。その結果を用いて詳細な多型地図を作成し各候補遺伝子の主要ハプロタイプ(頻度5%以上)を区別可能なSNPs(タグSNPs)を選択する。シークエンス法、もしくはTaqMan法を用いて患者および対照群全て(既存の検体に加え、追加で収集した検体全てを用いる)のジェノタイプングを実施する。遺伝子型を決定した後、連鎖不平衡地図ならびにハプロタイプ地図を作成し、相関解析、ハプロタイプ解析を実施して最も信頼性の高い遺伝子座を明らかにする。

D. 考察

HTLV-1によって起こるATLとHAMは家族内発症が多いことが、疾患発見当初から報告されていた。この臨床的観察は両疾患の発症に遺伝的なバックグラウンドが重要な役割を担っていることを示している。これまでATL、HAMに関連するSNPに関して、いくつかの報告があったが解析数も限られており十分な解析がなされていなかった。今回の研究では既に500例以上のATLサンプル、300検体のHAMサンプルを集めている。今年度の解析で既に疾患感受性遺伝子候補を同定できた。これらの多型は、疾患感受性遺伝子の候補となるものであるが、独立した集団の検体を用いた再現性検証が必須である。そのため、新たに収集された検体を用いた候補多型のタイピングを行い、結果の比較を通して、真の感受性遺伝子の同定を目指す。

E. 結論

今年度の解析でATL、HAMの疾患感受性と関連するゲノム上の多型の候補を複数個同定した。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

1. Fan J, Ma G, Nosaka K, Tanabe J, Satou Y,

- Koito A, Wain-Hobson S, Vartanian JP, Matsuoka M. APOBEC3G generates nonsense mutations in HTLV-1 proviral genomes *in vivo*. **J Virol**, 84: 7278-7287, 2010.
2. Sato H, Oka T, Shinnou Y, Kondo T, Washio K, Takano M, Takata K, Morito T, Huang X, Tamura M, Kitamura Y, Ohara N, Ouchida M, Ohshima K, Shimizu K, Tanimoto M, Takahashi K, Matsuoka M, Utsunomiya A, Yoshino T. Multi-step aberrant CpG island hyper-methylation is associated with the progression of adult T-cell leukemia/lymphoma (ATLL). **Am J Pathol**, 176: 402-415, 2010.
 3. Satou Y, Matsuoka M. HTLV-1 and the host immune system: How the virus disrupts immune regulation, leading to HTLV-1 associated diseases. **J Clin Exp Hematop**, 50:1-8, 2010.
 4. Matsuoka M. HTLV-1 bZIP factor gene: its roles in HTLV-1 pathogenesis. **Molecular Aspects of Medicine**, 31: 359-66, 2010.
 5. Satou Y, Yasunaga J, Zhao T, Yoshida M, Miyazato P, Takai K, Shimizu K, Ohshima K, Green PL, Ohkura N, Yamaguchi T, Ono M, Sakaguchi S, Matsuoka M. *HTLV-1 bZIP factor* induces T-cell lymphoma and systemic inflammation *in vivo*. **PLoS Pathog**, 7: e1001274, 2011.
 5. 佐藤賢文、Jun Fan、Guangyong Ma、野坂生郷、田邊順子、松岡雅雄：APOBEC3G generates nonsense mutations in HTLV-1 proviral genomes *in vivo*: 第 72 回日本血液学会学術集会、横浜、2010 年 9 月 24-26 日
 6. 菅田謙治、佐藤賢文、原英樹、光山正雄、松岡雅雄：HBZ gene expression in CD4+T cells impairs cell-mediated immunity against *Listeria monocytogenes*: 第 72 回日本血液学会学術集会、横浜、2010 年 9 月 24-26 日
 7. 松岡雅雄:HTLV-1 アクセサリー遺伝子 HBZ の機能と意義:第 58 回日本ウイルス学会学術集会、徳島、2010 年 11 月 7-9 日
 8. 佐藤賢文:HTLV-1 bZIP factor による T 細胞発がん機構:第 58 回日本ウイルス学会学術集会、徳島、2010 年 11 月 7-9 日
 9. 菅田謙治、佐藤賢文、松岡雅雄：HBZ 発現は IFN γ 産生を抑制し、HSV-2 感染に対する細胞性免疫を障害する:第 58 回日本ウイルス学会学術集会、徳島、2010 年 11 月 7-9 日
 10. 田口奈々絵、佐藤賢文、Miyazato Paola、吉田美香、松岡雅雄：HTLV-1 bZIP factor による慢性炎症惹起機構:第 58 回日本ウイルス学会学術集会、徳島、2010 年 11 月 7-9 日
 11. 馬広勇、Fan Jun、柳川伸一、松岡雅雄：DAPLE, a novel HBZ-binding protein, regulates Wnt signaling pathway in ATL cells: 第 58 回日本ウイルス学会学術集会、徳島、2010 年 11 月 7-9 日
 12. Masao Matsuoka. Pathogenesis and therapeutic strategies of human retroviruses:HTLV-1 and HIV. Mahidol-Kyoto Universities International Symposium 2010. Bangkok, Thailand. December, 16-17, 2010.

2. 学会発表

1. 趙鉄軍、佐藤賢文、今村健志、松岡雅雄：Human T-cell Leukemia virus Type 1 bZIP factor enhances TGF-beta signaling through p300/CBP coactivators: 第 69 回日本癌学会学術総会、大阪、2010 年 9 月 22-24 日
2. 中西梓、佐藤賢文、松岡雅雄：HTLV-1 bZIP factor inhibits expression of a proapoptotic factor, Bim:implication in leukemogenesis: 第 69 回日本癌学会学術総会、大阪、2010 年 9 月 22-24 日
3. 佐藤賢文、安永純一朗、趙鉄軍、大島孝一、松岡雅雄：HTLV-1 bZIP Factor transgenic mice induce chronic inflammations and T-cell lymphoma: 第 69 回日本癌学会学術総会、大阪、2010 年 9 月 22-24 日
4. 松岡雅雄：The HTLV-1 bZIP factor gene is responsible for leukemogenesis of adult T-cell leukemia: 第 72 回日本血液学会学術集会、横浜、2010 年 9 月 24-26 日

H. 知的財産権の出願・登録状況

特になし

1. 特許取得
2. 実用新案登録
3. その他

ヒト T 細胞白血病ウイルス 1 型関連疾患における感受性遺伝子多型の同定と
発症危険群へのアプローチ

研究代表者：京都大学ウイルス研究所 松岡 雅雄

研究分担者：京都大学医学研究科附属ゲノム医学センター 松田 文彦

研究要旨

ヒト T 細胞白血病ウイルス 1 型 (HTLV-1) 感染により惹起される成人 T 細胞白血病 (ATL)、HTLV-1 関連脊髄症 (HAM) の感受性遺伝子多型の同定を、SNP を用いた全ゲノム関連解析 (GWAS) を用いて試みた。その結果、それぞれの疾患に関して、候補遺伝子多型を複数個同定した。

A. 研究目的

ヒト T 細胞白血病ウイルス 1 型 (HTLV-1) 感染により惹起される成人 T 細胞白血病 (ATL)、HTLV-1 関連脊髄症 (HAM) の感受性遺伝子多型を解析し発症危険群の同定を目的とする。

B. 研究方法

SNP を用いた全ゲノム関連解析を行い、疾患関連 SNP を同定する。さらに、疾患関連 SNP の生物学的活性の解析：同定された疾患感受性 SNP が発症を促進する機序を明らかにするため、SNP がその分子に付与する機能について解析する。

(倫理面への配慮)

本研究は全て国の倫理指針に基づいた倫理委員会での審議、許可を得ている。また、検体は全て匿名化する。

C. 研究結果

前年度申請者らは、イルミナ社の Human610k アレイを用いて、熊本で収集された ATL 患者 258 検体と無症候性 HTLV-1 感染者 182 検体、鹿児島で収集された HAM 患者 296 検体と HTLV-1 感染者 100 検体のゲノムスキャンを実施した。本年度はそれに加えて、長崎県の ATL 患者 200 検体と HAM 患者 13 検体、鹿児島県の HAM 患者 82 検体、東京都の HAM 患者 61 検体と HTLV-1 感染者 42 検体、沖縄県の HAM 患者 8 検体のゲノムスキャンを実施した。結果の品質管理は、検体については、タイピング成功率が 90% 以上、近縁度が高いペアの片方を除外、また SNP についてはタイピング成功率が 95% 以上のもののみを解析に使用した。そして、前年度の結果と合わ

せて ATL 患者 415 検体、HAM 患者 429 検体と無症候性 HTLV-1 感染者 308 検体を用いて、患者と対照群の間で、アリル頻度、ジェノタイプ分布を統計解析した。その結果、ATL に関しては、 $p < 1 \times 10^{-5}$ の SNP が 4 個、HAM に関しては $p < 1 \times 10^{-5}$ の SNP が 15 個得られた。

D. 考察

これらの多型は、疾患感受性遺伝子の候補となるものであるが、独立した集団の検体を用いた再現性検証が必須である。そのため研究全体を通して検体数の増加につとめる。新たに収集された検体を用いて候補多型のタイピングを行い、結果の比較を通して、真の感受性遺伝子の同定を目指す。

E. 結論

ATL に関しては、 $p < 1 \times 10^{-5}$ の SNP が 4 個、HAM に関しては $p < 1 \times 10^{-5}$ の SNP が 15 個得られ、それぞれの疾患感受性と関連するゲノム上の多型の候補が複数個同定された。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

1. Nalpas, B., Laviaille-Meziani, R., Plancoulaine, S., Jouanguy, E., Nalpas, A., Munteanu, M., Charlotte, F., Ranque, B., Patin, E., Heath, S., Fontaine, H., Vallet-Pichard, A., Pontoire, D., Bourlière, M., Casanova, J. L., Lathrop, M., Bréchet, C., Poynard, T., Matsuda, F., Pol, S. and Abel, L. Interferon- γ receptor 2 gene variants are associated with liver fibrosis in

- patients with chronic hepatitis C infection. *Gut*. 59:1120-1126, 2010.
2. Hirose, K., Kawaguchi, T., Matsuda, F. and Yamada, R. Estimation of P-value of MAX test with double triangle diagram for 2 x 3 SNP case-control tables. *Genet. Epidemiol.* 34:543-551, 2010.
 3. Ohmura, K., Terao, C., Maruya, E., Katayama, M., Matoba, K., Shimada, K., Murasawa, A., Honjo, S., Takasugi, K., Tohma, S., Matsuo, K., Tajima, K., Yukawa, N., Kawabata, D., Nojima, T., Fujii, T., Yamada, R., Saji, H., Matsuda, F. and Mimori, T. Anti-citrullinated peptide antibody-negative RA is a genetically distinct subset: a definitive study using only bone-erosive ACPA-negative rheumatoid arthritis. *Rheumatology (Oxford)*. 49:2298-2304, 2010.
 4. Murakami, Y., Toyoda, H., Tanaka, M., Kuroda, M., Harada, Y., Matsuda, F., Tajima, A., Kosaka, N., Ochiya, T. and Shimotohno, K. The Progression of Liver Fibrosis Is Related with Overexpression of the miR-199 and 200 Families. *PLoS One*. 6:e16081, 2011.

2. 学会発表

特になし

H. 知的財産権の出願・登録状況

1. 特許取得
2. 実用新案登録
3. その他

ヒト T 細胞白血病ウイルス 1 型関連疾患における感受性遺伝子多型の同定と
発症危険群へのアプローチ

研究代表者：京都大学ウイルス研究所 松岡 雅雄

研究分担者：琉球大学大学院医学研究科 齊藤 峰輝

研究要旨： HTLV-1 関連疾患感受性遺伝子多型解析に用いる検体を可能な限り多く確保し、多型同定後の病態における役割の解明に資するため、昨年度に引き続き HAM 患者の臨床情報収集と検体バンクの整備を行った。一方、HTLV-1 マイナス鎖にコードされる新規遺伝子 HBZ の HAM 病態形成における意義を明らかにするため、各種抗 HBZ 抗体（ウサギ抗 HBZ ポリクローナル抗体、ラット抗 HBZ モノクローナル抗体、ヒト抗 HBZ モノクローナル抗体）を作製した。HAM 患者末梢血単核球、血清を用いて HBZ および Tax 発現を定量し、各種検査所見、病勢との関連を検討した。炎症反応が強い進行期の HAM 患者では、慢性期の患者と比較して HBZ が高発現しており病勢との相関が認められたが、Tax 発現と病勢との相関は見いだせなかった。特に脊髄と大脳深部白質に多巣性病変を認めた重症例において、ATL 患者レベルの極めて高い HBZ 遺伝子発現を認めたが、Tax 遺伝子は検出感度以下であった。HBZ トランスジェニックマウスの解析から、HBZ が *in vivo* において T 細胞増殖促進作用と皮膚・肺胞への CD4 陽性 T リンパ球浸潤作用を示すことが報告されているが、実際の HAM 患者においても HBZ 遺伝子が病態形成に極めて重要であることが示唆された。

A. 研究目的

ヒト T 細胞白血病ウイルス 1 型 (HTLV-1) は世界ではじめてヒト疾患との関連が見いだされたレトロウイルスであり、成人 T 細胞白血病 (ATL) および HTLV-1 関連脊髄症 (HAM) の原因ウイルスである。HTLV-1 感染者の分布には地理的な偏在が認められ、日本、アフリカ、カリブ海諸国、南アメリカ等に集積している。一昨年国立感染症研究所から報告された約 20 年ぶりの全国調査によると、我が国にはいまだに先進国で最多の約 108 万人もの HTLV-1 感染者が存在しており、従来多かった九州・沖縄では減少しているものの、都市部では逆に増加していることが明らかになっている。ほとんどの HTLV-1 感染者が生涯にわたって未発症の無症候性キャリアーとして経過し、HAM や ATL を発症するのは感染者全体の 5% 前後ではあるが、最も予後不良の白血病の一つである ATL は死亡者数が年間 1000 人を超え、HAM 患者では約 40% が経過中に歩行不能となり生活の質が著しく障害される。よって、HTLV-1 感染症の制圧は我が国の公衆衛生上・医療上の緊急の課題である。本研究の目的は、HAM 患者の疾患感受性遺伝子多型解析に用いる検体を可能な限り多く確保し、多型同定後、その HAM 病態における役割の解明に資するための研究資

源と病態解析法を整備することである。

B. 研究方法

鹿児島大学病院脳神経センター神経内科、琉球大学病院神経内科の協力のもと、昨年度に引き続き HAM 患者の臨床情報を収集した。同時に、十分な説明と書面による同意を得たのちに末梢血を採取し、血漿、リンパ球 (viable stock)、ゲノム DNA、cDNA を分離・保存した。リンパ球は Histopaque (Sigma) を用いて密度勾配遠心法にて分離し、AllPrep™ DNA/RNA Mini Kit (QIAGEN) を使用してゲノム DNA と total RNA を同時に抽出した。さらに PrimeScript RT reagent Kit (Takara) を用いて逆転写反応を行い鋳型 cDNA を合成した。

一方、HTLV-1 マイナス鎖にコードされる新規遺伝子 HBZ の HAM 病態形成における意義を明らかにするため以下の実験を行った。1) HBZ の C 末端ペプチド抗原を用いてウサギ抗 HBZ ポリクローナル抗体およびラット抗 HBZ モノクローナル抗体を作製した。2) 大腸菌で発現させ精製した HBZ を抗原としてファージディスプレイ法によりヒト抗 HBZ モノクローナル抗体を作製した。3) 56 例の HAM 患者末梢血単核球、血清を用いて HBZ および Tax 発現

を定量し、各種検査所見、病勢との関連を検討した。

(倫理面への配慮)

本研究は参加各施設の倫理委員会の承諾を得た後に施行した。十分な説明と同意のもと、書面による研究協力承諾書が得られた被験者から採取した検体のみを用い、完全に匿名化した後に行った。臨床情報と検体とは非連結匿名化した。

C. 研究結果

昨年度に引き続き、今年度は新たに 15 症例の HAM 患者から主治医の協力のもと匿名で臨床情報を収集した。同時に各患者から末梢血を採取した後に、血漿・リンパ球・ゲノム DNA・cDNA を分離、保存した。リンパ球は viable stock とした。病歴等の臨床情報に加え、各種一般臨床検査所見・免疫学的検査所見を含めた臨床情報データベースを作成した。

一方、HBZ 蛋白を特異的に検出可能な以下の抗体を得た。

抗原：HBZ の C 末端ペプチド

(HBZ 193-206 : CVNYWQGRLEAMWLQ)

抗体：ウサギ抗 HBZ ポリクローナル抗体
ラット抗 HBZ モノクローナル抗体 1 クローン (clone:4B12, Rat IgG2b)

抗原：大腸菌で発現させ精製した recombinant HBZ (HBZ 96-206)

抗体：2 価ヒト抗 HBZ Fab 抗体 17 クローン (Myc-His タグ付き)

これらを用いて HBZ 抗原、抗 HBZ 抗体を検出する ELISA 系を確立した。得られた抗 HBZ モノクローナル抗体は、HBZ を過剰発現させた 293T 細胞および HTLV-1 感染細胞株中の HBZ を蛍光抗体法、Western blot、フローサイトメトリーにより検出可能であったが、HAM 患者末梢血単核球 (PBMC) 中の HBZ は検出できなかった。また、HBZ 蛋白、抗 HBZ 抗体を検出する ELISA 系を作成したが、患者末梢血単核球中の HBZ、患者血漿中の抗 HBZ 抗体ともに検出できなかった。

一方、HAM 患者の病態と HBZ 発現との関連の解析結果から、炎症反応が強い進行期の HAM 患者は慢性期の患者と比較して HBZ mRNA が高発現しており病勢との有意な相関が認められたが、Tax mRNA 発現と病勢との有意な相関は見いだせなかった。特に脊髄と大脳深部白質に多巣性病変を認

めた重症例において、ATL 患者レベルの極めて高い HBZ mRNA 発現を認めたが、Tax mRNA は検出感度以下であった。

D. 考察

これまでに、HBZ 遺伝子が全ての ATL 症例で発現していること、HBZ 遺伝子の発現抑制により ATL 細胞の増殖が抑制されること、HBZ 遺伝子を発現するトランスジェニックマウスにおいて T リンパ腫のみならず CD4 陽性 T リンパ球の皮膚、肺への浸潤も認められることが報告されており、HBZ は HTLV-1 関連疾患の病態形成に極めて重要な遺伝子である。よって、実際の HTLV-1 感染者における HBZ 遺伝子の動態と病態との関連を明らかにすることは、HTLV-1 関連疾患の発症予防法、治療法の開発、病態形成機構の解明においてきわめて重要である。今回の結果は、HBZ 遺伝子が ATL のみならず HAM の病態形成においても極めて重要であること、作成した抗 HBZ 抗体が今後 HTLV-1 関連疾患の病態を解明する上で非常に有用なツールとなりうることを示している。HAM 患者の PBMC 中の HBZ 蛋白が検出できず、血漿中の抗 HBZ 抗体も感度以下であったことから、HBZ 蛋白発現は感染者の体内で強く抑制されている、HBZ 蛋白は限られた局所でのみ発現する、HBZ は主に RNA として機能する等の可能性がある。いずれにせよ宿主の免疫応答から巧みに逃避して機能していることを示しており、きわめて興味深い。

E. 結論

新たに 15 例の HAM 患者から臨床情報および末梢血を採取し、血漿、リンパ球 (viable stock)、ゲノム DNA、cDNA を分離、保存した。HBZ は HAM の病態形成に積極的な役割を果たしており、有望な治療標的分子となりうることを明らかにした。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

[1] Saito M. Immunogenetics and the pathological mechanisms of human T-cell leukemia virus type 1- (HTLV-1)-associated myelopathy/tropical spastic paraparesis (HAM/TSP). **Interdisciplinary Perspectives**

on Infectious Diseases. 2010 Article ID
478461.

[2] Saito K, Saito M, Taniura N, Okuwa T,
Ohara Y : Activation of the PI3K-Akt pathway
by human T cell leukemia virus type 1
(HTLV-1) oncoprotein Tax increases Bcl3
expression, which is associated with enhanced
growth of HTLV-1-infected T cells. **Virology**
403(2): 173-180, 2010.

[3] 齊藤峰輝
HAM/TSP 病態研究の最近の進歩
血液・腫瘍科 60(5): 642-650, 2010.

2. 学会発表

[1] 第 51 回日本神経学会総会 2010, 5. 東京
齊藤峰輝、田中礼子、松崎敏男、梅原藤雄、
田中勇悦: HTLV-1 関連脊髄症における OX40
陽性 T 細胞の意義

[2] 第 3 回 HTLV-1 研究会・合同班会議
2010, 8. 東京
齊藤峰輝、田中礼子、樋口雄二郎、末原雅
人、田中勇悦: HTLV-1 bZIP Factor (HBZ) に
対する抗体作製と HAM 臨床検体を用いた解
析.

[3] 第 58 回日本ウイルス学会学術集会
2010, 11. 徳島.
齊藤峰輝, 田中礼子, 田中勇悦. HTLV-1
関連脊髄症における OX40 陽性細胞の病因
的意義とその制御

[4] Saito M, Tanaka R, Matsuzaki T, Umehara
F, Tanaka Y : Enhanced expression of OX40
by HTLV-1 Tax and its roles in the
pathogenesis of HTLV-1-associated
myelopathy/tropical spastic paraparesis
(HAM/TSP). *The 10th International
Symposium on NeuroVirology*. 2010, 10. Milan,
Italy.

H. 知的財産権の出願・登録状況

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし

厚生労働科学研究費補助金（第3次対がん総合戦略研究事業）
分担研究報告書

ヒト T 細胞白血病ウイルス 1 型関連疾患における感受性遺伝子多型の同定と
発症危険群へのアプローチ

研究代表者：京都大学ウイルス研究所 松岡 雅雄
研究分担者：熊本大学医学部附属病院 野坂 生郷

研究要旨

ヒト T 細胞白血病ウイルス 1 型関連疾患の遺伝子多型の解析にあたり、特に成人 T 細胞白血病の臨床調査、検体収集を行い、発症に関与すると思われる制御性 T 細胞の機能についての解析を行った。

A. 研究目的

HTLV-I 感染によって惹起される成人 T 細胞白血病は、CD4 陽性の T 細胞リンパ系腫瘍であるが、その発症のメカニズムはまだ明らかになっていない。また、感染者の発症頻度は、5%以下と非常に低く、また家族内発症も報告されていることから、ウイルス側因子の違いや宿主側因子の遺伝的背景をはじめとした遺伝子多型の違いや宿主側の免疫機構の変化や異常により、HTLV-1 関連疾患への発症の可能性もある。成人 T 細胞白血病に焦点を当て、臨床調査も含め、免疫機構の異常について、特に制御性 T 細胞の働きについての解析も行った。

B. 研究方法

研究同意の得られた成人 T 細胞白血病患者の各病型の臨床調査を行い、また抽出した単核球を使用し、各病型における制御性 T 細胞の検出、ウイルス蛋白である Tax の発現と CCL22 などのサイトカイン産生の解析を行い、制御性 T 細胞への影響と HTLV-1 感染による免疫機能への関与も解析した。

（倫理面への配慮）

本研究を遂行するにあたり、ヘルシンキ宣言、臨床研究における倫理指針を遵守し、当施設の倫理委員会の承認を受け、患者に説明文書を使い、説明を行い、同意を得た。

C. 研究結果

成人 T 細胞白血病の臨床においては、ウイルス側因子の変異や欠失が急性型などの悪性の高い症例に比較的認められている。

また制御性 T 細胞のマーカーである FoxP3 の発現頻度は、HTLV-1 感染により増加するが、

HTLV-1 感染によりサイトカインの CCL22 の上昇が認められ、成人 T 細胞白血病では慢性型で上昇がみられた。急性型では上昇がみられていない。CCL22 上昇と FoxP3 陽性 CD4 陽性細胞の増加に相関が認められた。また、FoxP3 陽性 CD4 陽性細胞は成人 T 細胞白血病の T 細胞の増殖を抑制した。

D. 考察

HTLV-I 感染については、全く無症状で経過するキャリアの状態が大多数を占め、その中で HTLV-1 関連疾患の脊髄症や、T 細胞リンパ腫の成人 T 細胞白血病まで発症するがその頻度は著しく低い。そのため、遺伝的背景の違いや発症危険における免疫的な解明が必要である。FoxP3 陽性 CD4 陽性細胞が成人 T 細胞白血病細胞と異なる集団で存在し、その増殖にサイトカインの CCL22 により制御されていることがわかった。またこのことは、成人 T 細胞白血病の症例の T 細胞の増殖を抑制しており、病態の進行における何らかの役割を果たしている可能性もある。

E. 結論

FoxP3 陽性 CD4 陽性細胞が成人 T 細胞白血病細胞と異なる集団で存在し、その増殖にサイトカインの CCL22 により一部制御されており、成人 T 細胞白血病の発症の制御に関与している可能性もあると思われる。

F. 健康危険情報

臨床検体は数 mL の末梢血採取のみであり、貧血等が起る可能性はなく、また通常の検査採血と同時に行った。本研究における健康危険はないものとする。

G. 研究発表

1. 論文発表

Human T-lymphotropic virus type 1-induced CC chemokine ligand 22 maintains a high frequency of functional FoxP3+ regulatory T cells.

Toulza F, Nosaka K, et al J Immunol 185: 183-189, 2010.

APOBEC3G generates nonsense mutations in human T-cell leukemia virus type 1 proviral genomes in vivo.

Fan J, Ma G, Nosaka K, et al. J Virol. 84:7278-7287, 2010.

2. 学会発表

なし

H. 知的財産権の出願・登録状況

特になし。

1. 特許取得
2. 実用新案登録
3. その他

研究成果の刊行に関する一覧表

研究代表者：京都大学ウイルス研究所 松岡 雅雄

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Fan J, Ma G, N osaka K, Tanab e J, Satou Y, K oito A, Wain-H obson S, Vartan ian JP, <u>Matsuok a M.</u>	APOBEC3G generate s nonsense mutations in HTLV-1 proviral genomes <i>in vivo</i> .	J Virol	84	7278-7287	2010
Sato H, Oka T, Shinnou Y, Kon do T, Washio K, Takano M, Takata K, Mori to T, Huang X, Tamura M, Ki tamura Y, Ohar a N, Ouchida M, Ohshima K, Shimizu K, Ta nimoto M, Taka hashi K, <u>Matsu oka M.</u> , Utsuno miya A, Yoshin o T.	Multi-step aberrant C pG island hyper-meth ylation is associated with the progression of adult T-cell leuke mia/lymphoma (ATL L).	Am J Pathol	176	402-415	2010
Satou Y, <u>Matsu oka M.</u>	HTLV-1 and the host immune system: Ho w the virus disrupts i mmune regulation, le ading to HTLV-1 ass ociated diseases.	J Clin Exp H ematop	50	1-8	2010
<u>Matsuoka M.</u>	HTLV-1 bZIP factor gene: its roles in HT LV-1 pathogenesis.	Molecular As pects of Medi cine	31	359-366	2010
Satou Y, Yasun aga J, Zhao T, Yoshida M, M iyazato P, Taka i K, Shimizu K, Ohshima K, Green PL, Oh kura N, Yamag uchi T, Ono M, Sakaguchi S, <u>Matsuoka M.</u>	<i>HTLV-1 bZIP factor</i> induces T-cell lymph oma and systemic inf lamination <i>in vivo</i> .	PLoS Pathog	7	E1001274	2011

APOBEC3G Generates Nonsense Mutations in Human T-Cell Leukemia Virus Type 1 Proviral Genomes *In Vivo*^{∇†}

Jun Fan,¹ Guangyong Ma,¹ Kisato Nosaka,² Junko Tanabe,¹ Yorifumi Satou,¹ Atsushi Koito,³ Simon Wain-Hobson,⁴ Jean-Pierre Vartanian,⁴ and Masao Matsuoka^{1*}

Laboratory of Virus Control, Institute for Virus Research, Kyoto University, 53 Shogoin Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan¹; Department of Hematology and Infectious Diseases, Kumamoto University School of Medicine, 1-1-1 Honjo, Kumamoto 860-0811, Japan²; Department of Retrovirology and Self-Defense, Faculty of Medical and Pharmaceutical Sciences, Kumamoto University School of Medicine, 1-1-1 Honjo, Kumamoto 860-0811, Japan³; and Molecular Retrovirology Unit, Institut Pasteur, 28 rue du Dr. Roux 75724 Paris Cedex 15, France⁴

Received 23 October 2009/Accepted 23 April 2010

Human T-cell leukemia virus type 1 (HTLV-1) induces cell proliferation after infection, leading to efficient transmission by cell-to-cell contact. After a long latent period, a fraction of carriers develop adult T-cell leukemia (ATL). Genetic changes in the *tax* gene in ATL cells were reported in about 10% of ATL cases. To determine genetic changes that may occur throughout the provirus, we determined the entire sequence of the HTLV-1 provirus in 60 ATL cases. Abortive genetic changes, including deletions, insertions, and nonsense mutations, were frequent in all viral genes except the *HBZ* gene, which is transcribed from the minus strand of the virus. G-to-A base substitutions were the most frequent mutations in ATL cells. The sequence context of G-to-A mutations was in accordance with the preferred target sequence of human APOBEC3G (hA3G). The target sequences of hA3G were less frequent in the plus strand of the *HBZ* coding region than in other coding regions of the HTLV-1 provirus. Nonsense mutations in viral genes including *tax* were also observed in proviruses from asymptomatic carriers, indicating that these mutations were generated during reverse transcription and prior to oncogenesis. The fact that hA3G targets the minus strand during reverse transcription explains why the *HBZ* gene is not susceptible to such nonsense mutations. HTLV-1-infected cells likely take advantage of hA3G to escape from the host immune system by losing expression of viral proteins.

Human T-cell leukemia virus type 1 (HTLV-1) is the causative agent of both adult T-cell leukemia (ATL) and inflammatory HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP) (9, 40). HTLV-1 is a complex retrovirus that carries regulatory genes, such as *tax* and *rex*, and accessory genes, such as *p12*, *p13*, and *p30* (6, 29). An accessory gene, *HTLV-1 bZIP factor* (*HBZ*), is transcribed from the 3' long terminal repeat (3'LTR) as an antisense transcript (10, 46). Previous studies have shown that Tax expression can immortalize T lymphocytes *in vitro* (1, 11), and *in vivo* the expression of Tax in transgenic mice causes various tumors depending on the tissue-specific promoter that expresses Tax (12, 14, 20). In more than half of human ATLs, the *tax* gene is not transcribed (41). Three mechanisms have been identified to inactivate Tax expression (29): (i) abortive genetic changes in the *tax* gene (7, 41), (ii) silencing by DNA methylation in the 5'LTR (18, 43), and (iii) deletion of the 5'LTR (31, 42). Since Tax is a major target of cytotoxic T-lymphocytes (CTLs) *in vivo* (17), the tumor cells might escape from the host immune system by suppressing Tax expression. However, the actual mechanism

used to create genetic changes in the *tax* gene remains to be elucidated. Nonsense mutations in the *tax* gene have also been observed in HTLV-1 carriers, and the mechanism to generate these mutations is similarly unknown (8).

As a host defense against retroviruses, mammalian cells employ the APOBEC3 cytidine deaminase family, which causes deamination during reverse transcription, resulting in nucleotide mutations (3, 26). Human APOBEC3G (hA3G) deaminates cytosine residues of single-stranded DNA during reverse transcription, resulting in high levels of plus-strand G-to-A mutations. The human immunodeficiency virus (HIV) nucleocapsid is critical for incorporation of hA3G into virions (2, 24), while its accessory protein, Vif, counteracts hA3G. Vif inhibits hA3G packaging into HIV-1 virion through ubiquitination and proteasomal degradation of hA3G (39, 49). HTLV-1 does not encode a protein analogous to Vif that inactivates hA3G. Instead, it has been demonstrated that a domain of the HTLV-1 nucleocapsid suppresses the incorporation of hA3G into the virion (4). Consistent with this finding, it has been reported that G-to-A mutations are rare in HTLV-1 carriers (25).

In this study, we analyzed the entire sequences of proviruses in ATL and carrier cells and found that most nonsense mutations in the proviruses were caused by deamination. The sites of nonsense mutation coincided with the preferred target sequences of hA3G. In ATL cells, nonsense mutations, deletions, and insertions were detected in most of the viral genes except the *HBZ* gene, supporting the critical role of the *HBZ* gene in ATL.

* Corresponding author. Mailing address: Institute for Virus Research, Kyoto University, 53 Shogoin Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan. Phone: 81-75-751-4048. Fax: 81-75-751-4049. E-mail: mmatsuok@virus.kyoto-u.ac.jp.

† Supplemental material for this article may be found at <http://jvi.asm.org/>.

[∇] Published ahead of print on 12 May 2010.

MATERIALS AND METHODS

Cell lines and clinical samples. MT-2 and MT-4 are HTLV-1-transformed cell lines. MT-1 and TL-Om1 were derived from leukemia cells. Jurkat cells were negative for HTLV-1. These cells were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum, 100 U/ml penicillin, 100 µg/ml streptomycin, and 2 mM L-glutamine at 37°C in a 5% CO₂ atmosphere. Clinical samples were collected from 60 ATL patients and 10 HTLV-1 carriers. Genomic DNA was extracted from peripheral blood mononuclear cells using standard phenol-chloroform methods. This study was conducted according to the principles regarding human experimentation expressed in the Declaration of Helsinki. The study was approved by the Institutional Review Board of Kyoto University (G204). All patients provided written informed consent for the collection of samples and subsequent analysis.

Sequencing the complete provirus genomes in ATL cells. The copy number of HTLV-1 provirus in cells from ATL patients was determined by inverse PCR (5), and ATL samples with one complete provirus were selected for direct sequencing. The complete provirus genome was amplified as two halves (fragments I and II) from genomic DNA samples derived from ATL patients. Fragment I was amplified using primers 5'-TGACAATGACCATGAGCCCCAAATATCC-3' and 5'-CGGCTATTAAGACCAGGAAG-3'. PCR conditions were as follows: an initial step of 5 min at 94°C; 30 cycles, with 1 cycle consisting of 30 s at 94°C, 30 s at 64°C, and 5 min at 72°C; and a final step of 10 min at 72°C. Fragment II was amplified using primers 5'-AGAAACAAGCTCAGAAGCTA-3' and 5'-TGTACTAAATTTCTCCTCGAGAGTGC-3'. PCR conditions were as follows: an initial step of 5 min at 94°C; 30 cycles, with 1 cycle consisting of 30 s at 94°C, 30 s at 60°C, and 5 min at 72°C; and a final step of 10 min at 72°C. PCR products were then subjected to nested PCR to amplify I-1 and I-2 from fragment I and II-1 and II-2 from fragment II. The primers for I-1 were 5'-TGACAATGACCATGAGCCCCAAATATCC-3' and 5'-GAGCTTAAAGTGATCTTGG-3'. The primers for I-2 were 5'-TTCCGATAGCCTTGTCTCA-3' and 5'-CGGCTATTAAGACCAGGAAG-3'. The primers for II-1 were 5'-TGTTATTATTTCAA GCTTCC-3' and 5'-AAATGCAGGAGTTGGGGATT-3'. The primers for II-2 were 5'-AACTGTCTAGTATAGCCATC-3' and 5'-TGTACTAAATTTCTCTC CTGAGAGTGC-3'. PCR conditions for nested PCR were as follows: an initial step of 5 min at 94°C; 28 cycles, with 1 cycle consisting of 30 s at 94°C, 30 s at X°C, and 2.5 min at 72°C; and a final step of 72°C for 10 min. X was 60°C for I-1 and II-2 and 58°C for I-2 and II-1. The final PCR products were gel purified, and the sequence was determined by direct sequencing.

Sequencing the *pol* and *tax* regions in HTLV-1 carriers. The proviral load was determined as previously reported (32). Herculase II fusion DNA polymerase (Stratagene, La Jolla, CA), a *Pfu*-based DNA polymerase, was used to amplify *pol* and *tax* regions from genomic DNA derived from HTLV-1 carriers. The primers used to amplify the *pol* region were 5'-TACACCTTGCATCCTATG G-3' and 5'-GCTAGGCTTGCCTAGATGGG-3'. PCR conditions were as follows: an initial step of 5 min at 95°C; 30 cycles, with 1 cycle consisting of 20 s at 95°C, 20 s at 60°C, and 50 s at 72°C; and a final step of 50 s at 72°C. The primers used to amplify the *tax* region were 5'-GGTCCCTCCGGGCATGACACA-3' and 5'-TCTCCACGCTTTTATAGACT-3'. PCR conditions were as follows: an initial step of 5 min at 95°C; 28 cycles, with 1 cycle consisting of 20 s at 95°C, 20 s at 68°C, and 50 s at 72°C; and a final step of 5 min at 72°C. The PCR products were purified and cloned, and at least 20 clones were sequenced. DNA polymerase fidelity was studied by using a clone carrying an amplified *tax* region as the template instead of genomic DNA. The *tax* region was amplified using the same PCR conditions described above and cloned, and 25 clones were sequenced.

Sequencing the HTLV-1 genome in nontumor cells derived from ATL samples. ATL patient-derived tumor cells are considered a mixture of a major population of monoclonal expanded tumor cells and a minor population of HTLV-1-infected nontumor cells. To amplify the HTLV-1 genome from infected nontumor cells, ATL samples carrying defective HTLV-1 genome based on direct sequencing were selected. Target regions were amplified by setting one of the primers at the deleted region. Sequences were determined by direct sequencing.

Mutation analysis. Nucleotide sequences of all ATL samples were aligned using GENETYX-MAC version 13.0 software. Mutations were determined after filtering for (i) polymorphisms, defined as a nucleotide substitution occurring in >5 cases; and (ii) linkers, defined as >7 common nucleotide substitutions occurring in >2 cases. With these criteria, the nucleotide sequence from an ATL case, ATL-25, showed the fewest mutations (3 single-nucleotide substitutions with no deletions or insertions) and was selected as a standard sequence (AB513134). The nucleotide positions of identified mutations were numbered relative to this standard sequence.

For carrier samples, the sequence of a given region from all clones amplified from a carrier was aligned and compared with the standard sequence described

above with mutated bases corrected. When a nucleotide substitution occurred in all clones, it was considered a polymorphism and excluded from the mutation repertoire.

cDNA synthesis and quantitative real-time RT-PCR. Total RNA was extracted from transfectants using Trizol reagent (Invitrogen, Carlsbad, CA). RNA was treated with DNase I (Invitrogen) to eliminate the genomic DNA. cDNA was synthesized from 1 µg of total RNA with the Superscript preamplification system (Invitrogen) according to the manufacturer's protocol. cDNA product was quantified by real-time reverse transcription-PCR (RT-PCR) with power SYBR green PCR master mix and 7900HT fast real-time PCR system (both from Applied Biosystems) according to the manufacturer's instructions. The specific primers for the *HA3G* gene were 5'-CGCGTGCCACCATGAAGATC-3' (forward) and 5'-TGTGGGTGGATCCATCGAGT-3' (reverse). The specific primers for the human *AID* gene were 5'-TTCACCGCGCCCTACTT-3' (forward) and 5'-GCTGTCTGGAGAGACGAAC-3' (reverse). Target cDNA was normalized to the amount of endogenous mRNA of *ACTB*. The primers used for *ACTB* were 5'-AGGCCAACCGCGAGAAGATG-3' (forward) and 5'-CCAGAGGC GTACAGGGATAG-3' (reverse). PCRs were carried out in triplicate. Data were analyzed by the comparative threshold cycle (*C_T*) method according to protocols of the manufacturer (Applied Biosystems).

In vitro editing of the HTLV-1 genome by HA3G. QT6 quail cells were transfected with an infectious molecular clone of HTLV-1 (pX1MT-M) (30) in the presence or absence of the HA3G expression plasmid. Twenty-four hours post-transfection, HTLV-1-producing QT6 cells were treated with 200 µg/ml of mitomycin C at 37°C for 30 min, washed well, and then cocultivated 1:1 with fresh QT6 cells (13). Seventy-two hours later, the cells were collected, and genomic DNA was extracted. HA3G-mediated DNA hypermutation was detected using three-dimensional (3D) PCR that allows differential amplification of G-to-A hypermutants using a *Taq* DNA polymerase as previously described (25). The first round of PCR was performed with a denaturation temperature of 94°C using the primer pair 5'-CTGCAGATACAAAGTTAAC-3' (forward) and 5'-TGG AGGAAGGAGGGTGGAAAT-3' (reverse). PCR conditions were as follows: an initial step of 5 min at 94°C; 35 cycles, with 1 cycle consisting of 30 s at 94°C, 30 s at 58°C, and 30 s at 72°C; and a final extension step of 10 min at 72°C. Then, 0.5 µl of the first-round reaction product was used as the template for the second round of PCR using the following primers: 5'-CCATGCTTATTATCAGCCC A-3' (forward) and 5'-GTTGGGGTTGTATGAGTGA-3' (reverse). The PCR program was the same as for the first round of reaction except that the denaturation gradient ranged from 94 to 82°C. PCR products were purified and cloned, and up to 20 clones were sequenced.

Nucleotide sequence accession number. The nucleotide sequence of ATL-25, selected as the standard sequence, has been submitted to the DDBJ (DNA Data Bank of Japan) database under accession number AB513134.

RESULTS

HTLV-1 proviral sequence changes in ATL. Previous studies showed that *Tax* expression was frequently lost by genetic changes in the *tax* gene, by DNA methylation of the 5'LTR, and by deletion of the 5'LTR (29). Similarly, although not well studied thus far, the expression of other viral proteins might also be lost through genetic changes in the provirus sequence. To analyze the genetic changes of the HTLV-1 provirus in ATL cells, we determined the entire sequence of the provirus (9,034 bp) in 60 ATL cases. For this study, we chose ATL cases with HTLV-1 proviruses containing both LTRs. As shown in Table 1, we detected genetic changes (nonsense mutations, deletions, and insertions) in the coding regions of viral genes in 28 of 60 ATL cases (46.7%). Deletions in the proviruses were detected in 27 cases (45%), while insertions were found in 10 cases (16.7%) (see Table S1 in the supplemental material). These deletions and insertions resulted in the loss or truncation of protein(s) in 18 cases (Table 1; also see Table S1 in the supplemental material).

Nonsense mutations, deletions, and insertions in the viral genes were found in all viral genes except the *HBZ* gene (Table 1). Since the *p30* and *p13* genes use an identical coding frame, the same mutations generate nonsense changes in both *p30*

TABLE 1. Abortive genetic changes of HTLV-1 viral genes in 60 ATL cases^a

ATL case	Genetic change ^b in the following viral gene:								
	<i>gag</i>	<i>pol</i>	<i>env</i>	<i>p12</i>	<i>p30</i>	<i>p13</i>	<i>p27</i>	<i>tax</i>	<i>HBZ</i>
ATL-2		DEL							
ATL-3								W248*	
ATL-7			DEL						
ATL-8			DEL				W139*		
ATL-10	DEL	DEL	DEL						
ATL-11		W680*	W427*						
ATL-12	DEL		W387*		W165*	W11*	W139*	W56*	
ATL-13	W59*	W680*	DEL		W165*	W11*		W56*	
					W190*	W36*			
ATL-14							DEL	DEL	
ATL-15		W680*							
ATL-27		DEL							
ATL-29			W431*		W165*	W11*		W56*	
ATL-32		W271*							
ATL-36								DEL	
ATL-37		W485*	W50*					W56*	
		W680*							
ATL-39				W87*					
ATL-40					R184*	R30*			
ATL-41		W485*						W56*	
ATL-42			DEL						
ATL-43		DEL							
ATL-44	DEL	DEL		W87*					
ATL-45	DEL							Y286*	
ATL-46	DEL	DEL							
ATL-47	DEL	DEL							
ATL-48	DEL							IN	
ATL-49	DEL	DEL							
ATL-51					W178*	W24*			
ATL-53	DEL								

^a Only ATL cases with deletions, insertions, or nonsense mutations are shown.

^b Abbreviations: DEL, deletion; IN, insertion; W, tryptophan; R, arginine; Y, tyrosine. A number after an amino acid indicates the position of the nonsense mutation. Stop codons are indicated by an asterisk.

and *p13* (Fig. 1). The frequency of genetic changes in the *tax* gene was 16.7% (10/60) (Table 1), which was higher than in previous studies (7, 41). Of a total of 29 nonsense mutations, 27 preferentially accumulated in tryptophan codons (27/29 [93.1%]) (see Table S2 in the supplemental material). Since the tryptophan codon is TGG, a G-to-A mutation generates either a TGA or a TAG stop codon. In one nonsense mutation,

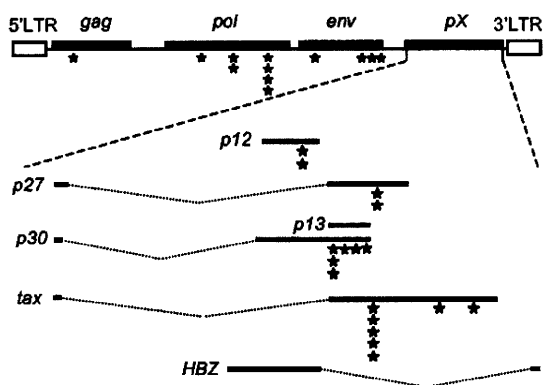


FIG. 1. Schema of the distribution of nonsense mutations in the HTLV-1 provirus. The *p13* gene has the same reading frame as the *p30* gene. Nonsense mutations observed in ATL cases are indicated by black stars.

the codon for arginine (CGA) was converted to the TGA stop codon by a C-to-T mutation. Interestingly, nonsense mutations tended to accumulate in the same cases (Table 1).

It was difficult to discriminate between mutations and polymorphisms in HTLV-1 provirus from ATL patients, since the proviral sequences in nonleukemic cells could not be analyzed in most ATL cases without an internal deletion of provirus. Some mutations might be characteristic of a subgroup of proviruses (polymorphism). Since nonsense mutations at the *tax* gene (nucleotide position 7469) were detected in 5 ATL cases, we tentatively classified base substitutions as polymorphisms when more than 5 identical base substitutions were observed in different cases. On the basis of this criterion, we observed 591 mutations in all ATL cases (see Table S3 in the supplemental material). Among them, the G-to-A mutation was the most frequent, contributing to 28.8% of the total mutations. Other frequently detected mutations were the C-to-T (23.7%), A-to-G (17.1%), and T-to-C (12.9%) mutations. Frequent G-to-A and C-to-T mutations strongly suggested the activity of deamination enzymes in the generation of these mutations.

Mutations in leukemic and nonleukemic cells in ATL cases. In ATL patients, nonleukemic HTLV-1-infected cells coexist with leukemic cells. To analyze the proviral sequences of nonleukemic cells, we amplified proviral sequences using primers within regions deleted in leukemic cells to avoid amplification of provirus from leukemic cells. Nonsense mutations that were

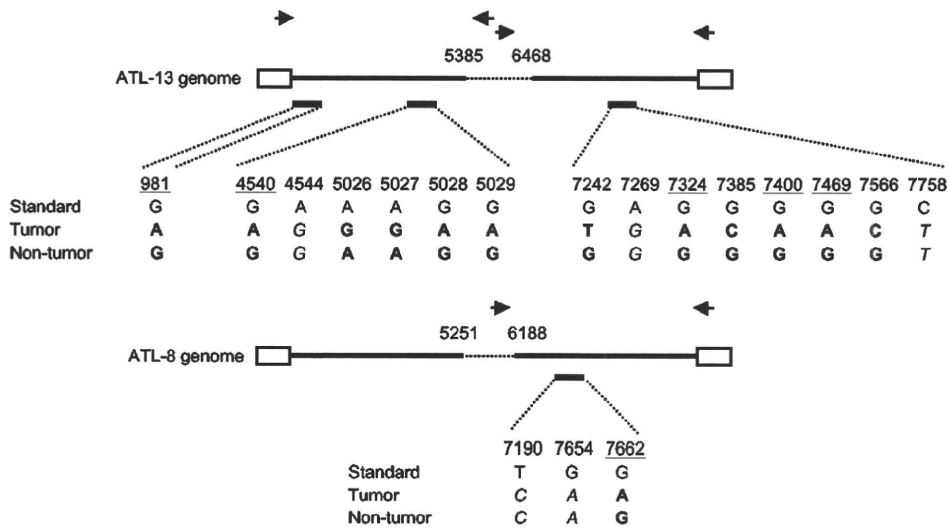


FIG. 2. Mutations in non-tumor-cell-derived proviral sequence in ATL patients. HTLV-1 genomes with internal deleted regions from two samples from two ATL patients are displayed schematically. As indicated by horizontal arrows, one of the PCR primers was set in the deleted region in order to specifically amplify a proviral region that originated from nontumor cells (see Materials and Methods). Nested PCR products, represented by short black lines, were subjected to direct sequencing. Nucleotide bases mutated in tumor cells according to the criteria mentioned in Materials and Methods are displayed. A base substitution is considered a polymorphism if the same base was found also in the nontumor cells derived from the same patient sample. Positions at which a G-to-A mutation will result in a nonsense mutation are underlined.

observed in leukemic cells were not detected in nonleukemic cells (Fig. 2), indicating that nonsense mutations generated in the provirus of leukemic cells were selected for. We compared the entire sequence of the HTLV-1 provirus in both leukemic cells and nonleukemic cells from five ATL cases. As shown in Table 2, only 37% of base substitutions were considered real mutations, since the nucleotide differed between leukemic cells and nonleukemic cells. However, 71% of G-to-A transitions were actually mutations, while 73% of other base substitutions were polymorphisms (Table 2), indicating that G-to-A mutations are predominant in ATL cells.

Association of G-to-A mutations with human APOBEC3G. A high frequency of G-to-A mutations suggests the role of deaminase(s) in generating these mutations. Deamination enzymes are known to have individually preferred target sequences for deamination (44). Sequences surrounding the nonsense mutations generated by the G-to-A mutation in HTLV-1 proviruses showed a predominance of the 5'-GG dinucleotide context (target underlined) (Fig. 3A). Among trinucleotides

containing GG dinucleotides, CGG, TGG, and GGG were preferred (Fig. 3A); these are consistent with the target sequences of human APOBEC3G (hA3G) (48). The AGG sequence was also targeted in HTLV-1, unlike HIV-1. When we checked tetrameric sequences that contain a central GG dinucleotide, CGGG and TGGG were the preferred targets in the HTLV-1 provirus, similar to those reported for HIV-1 (Fig. 3B). As in HIV-1, a C at the +2 position (NGGC [N indicates T or C]) was not favored in HTLV-1. These findings suggest that the observed G-to-A mutations were generated by hA3G. Another deaminase, human APOBEC3F (hA3F), which is largely coexpressed with hA3G, is also reported to target single-stranded minus-strand DNA (46). Unlike hA3G, which favors the 5'-GG dinucleotide, the consensus target sequence of hA3F is 5'-GA (21). G-to-A mutations at GA sites contributed 13.4% of all G-to-A mutations (Fig. 3A), suggesting that hA3F might also play a role in these mutations in the HTLV-1 provirus.

Since hA3G targets the minus strand during reverse tran-

TABLE 2. Concordance rate of mutations between leukemic cells and nonleukemic cells in five ATL cases

ATL case or parameter	No. of mutations ^a												
	Total ^b	G-to-A	C-to-T	A-to-G	T-to-C	A-to-C	A-to-T	C-to-G	C-to-A	T-to-G	G-to-C	G-to-T	T-to-A
ATL-8	10 (4)	2 (1)	3 (1)	2 (0)	1 (0)			1 (1)			1 (1)		
ATL-12	25 (8)	8 (6)	7 (2)	5 (0)	3 (0)			2 (0)					
ATL-13	36 (20)	10 (8)	8 (3)	7 (2)	2 (1)	2 (0)			1 (0)	1 (0)	3 (3)	1 (0)	1 (0)
ATL-14	12 (2)	1 (1)	7 (0)	1 (0)	2 (0)				1 (1)				
ATL-45	16 (3)	3 (1)	3 (0)	4 (0)	2 (1)		3 (0)		1 (1)				
Total	99 (37)	24 (17)	28 (6)	19 (2)	10 (2)	2 (0)	3 (0)	3 (1)	3 (2)	1 (0)	4 (4)	1 (0)	1 (0)
Concordance rate (%)	37	71	21	11	20	0	0	33	67	0	100	0	0

^a Each entry [A (B)] shows the number of each type of mutation in ATL cases judged by our criteria (A) and the number of confirmed mutations (B). The concordance rate of mutations was calculated by dividing B by A.

^b Number of mutations of all types in each ATL case.

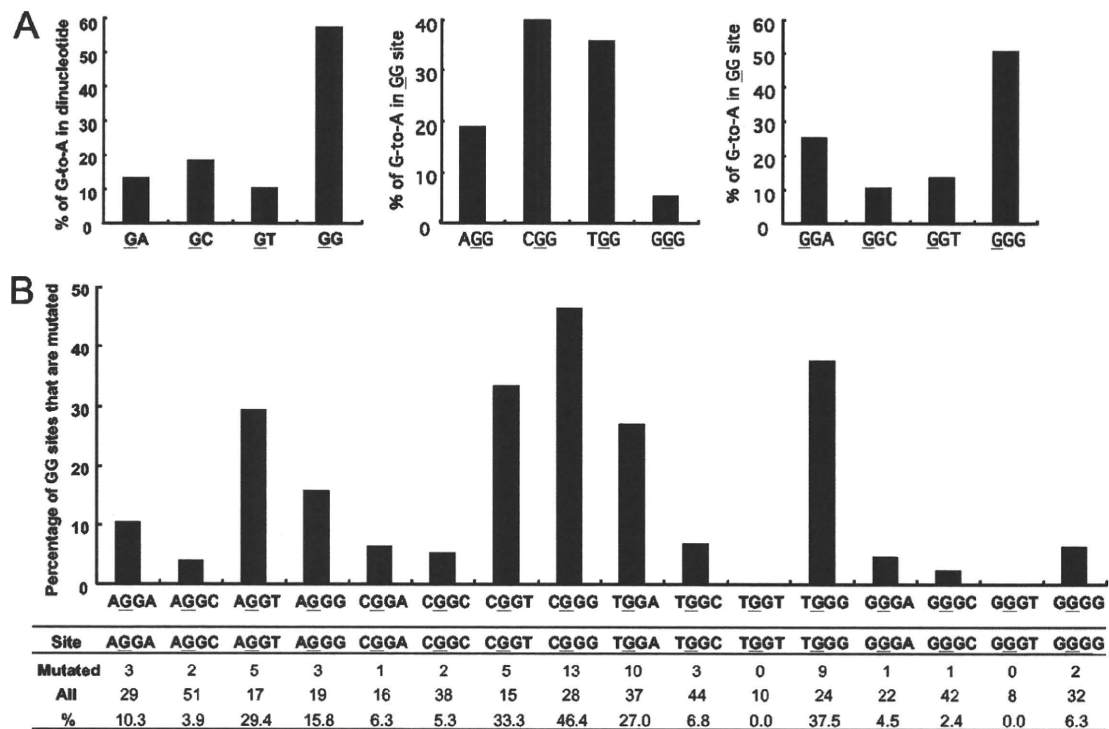


FIG. 3. Target sequence preference of G-to-A mutations. (A) Sequence context of G-to-A mutations. In all G-to-A mutations, the percentage of mutations at the indicated dinucleotide sequence context is shown (left panel). In the GG context, the influence of 5' (middle panel) and 3' (right panel) nucleotides neighboring the GG dinucleotide on G-to-A mutations is displayed. (B) Frequency of G-to-A mutations on sites of tetranucleotides containing a central GG. HTLV-1 genome sequence from ATL-25 was used as a standard sequence with mutated bases corrected to calculate the total number of GG sites that exist in the provirus genome (see Materials and Methods). The percentage of GG sites that underwent G-to-A mutation at least once in 60 ATL cases was plotted against each sequence context. Mutated bases are underlined. CGGG and TGGG were the preferred targets for G-to-A mutation in the HTLV-1 provirus. A C at position +2 (NGGC) was not favored.

scription, nonsense mutations generated by G-to-A mutations are not found in the *HBZ* gene. A G-to-A mutation in the proviral sense strand caused a nonsense mutation in the *p12* gene (TGG to TGA), whereas the same mutation generated a missense mutation in the *HBZ* gene (CCA to TCA) (Fig. 1). G-to-A mutations occurred most frequently at GG sites (target underlined) with a T or C at their 5' terminus (Fig. 3A). We therefore counted the number of TGG or CGG sites within individual HTLV-1 regions (Table 3). As NGGC (N indicates

T or C) was not favored by hA3G (Fig. 3B), it was not included. Among HTLV-1 regions, the plus-strand sequence of the *HBZ* gene had the fewest TGG and CGG sites, even when the number of sites was normalized by the size of the region. This indicates that mutations in the *HBZ* gene-coding region caused by hA3G are also rare.

Mutations in the HTLV-1 proviral sequences from carriers. Frequent detection of G-to-A mutations in HTLV-1 proviral sequences in ATL cases suggests that these nonsense mutations potentially occur in both HTLV-1 carriers and ATL patients. To explore this possibility, we next studied the proviral sequences of the *pol*, *env*, *p12*, *p13/p30*, *p27*, *tax*, and *HBZ* genes in asymptomatic HTLV-1 carriers. The proviral sequences in asymptomatic carriers were amplified using a high-fidelity DNA polymerase with proofreading activity, and after subcloning, the sequences were determined for at least 20 subclones. As seen in ATL cases, the mutation frequency differed drastically among individual carriers (representative data are shown in Tables S4 and S5 in the supplemental material), and the distribution of base substitutions among clones varied within individual carriers. G-to-A mutations accounted for 81% and 72% of all mutations in the *pol* and *tax* genes, respectively (Table 4 and 5, respectively). A high frequency of G-to-A mutations was also observed in other viral genes (data not shown). Analyses of the sequence context of G-to-A mutations showed a predominance of GG sequences, which is

TABLE 3. Distribution of hA3G target sequence in the plus strand of the HTLV-1 genome

Gene or genome region	Size (bp)	No. of sites ^a			
		CGGX	CGGX/kb	TGGX	TGGX/kb
5'LTR	755	8	10.60	3	3.97
<i>gag</i>	1,290	12	9.3	10	7.75
<i>pol</i>	2,691	16	5.95	25	9.29
<i>env</i>	1,467	8	5.45	17	11.59
<i>p27</i>	570	3	5.26	10	17.54
<i>p30</i>	726	2	2.75	7	9.64
<i>tax</i>	1,062	5	4.71	11	10.36
<i>HBZ</i>	621	1	1.61	1	1.61
3'LTR	755	7	9.27	3	3.97

^a Number of sites (e.g., CGGX sites or TGGX sites, where X indicates A, T, or G) alone or the number of CGGX or TGGX sites normalized by the size of the region (e.g., number of CGGX sites in 1 kb [CGGX/kb]).

TABLE 4. Mutations in the *pol* region of 10 HTLV-1 carriers

HTLV-1 carrier (no. of clones) ^a	No. of the indicated single-nucleotide substitution in HTLV-1 carrier:								No. of sequences with the following sequence context for the G-to-A mutation ^b :				
	G-to-A	C-to-T	A-to-G	T-to-C	A-to-C	C-to-G	G-to-C	G-to-T	T-to-A	<u>GA</u>	<u>GC</u>	<u>GT</u>	<u>GG</u>
C72 (25)	5	1					1						5
C79 (20)	8		1		1			1		1	1		6
C63 (28)	1							1				1	
C40 (26)	13		1										13
C32 (22)		2											
C62 (22)	1	2											1
C87 (22)	14									1	1		13
C5 (23)	19	1						1	2				19
C29 (25)	9					1		1			1		8
C82 (27)	7			1							1		6
Total	77	6	2	1	1	1	1	4	2	2	4	1	71

^a The number of clones studied from each carrier is shown in parentheses.
^b Mutated bases are underlined.

consistent with the finding that hA3G deaminates the viral genome during reverse transcription (48). In order to compare the frequencies of nonsense mutations in carriers versus ATL cases, the occurrence of nonsense mutations was further analyzed. Consistent with the findings for ATL cases, nonsense mutations were detected in the *pol*, *env*, *p30/p13*, *p27*, and *tax* genes, but not in the *p12* and *HBZ* genes (Table 6). As in ATL cases, nonsense mutations in carriers were most frequently observed in the *tax* and *pol* genes. This result suggests that there is no bias for mutations in specific viral genes in ATL cases compared with carriers. These results suggest that nonsense mutations in provirus are not generated during oncogenesis but are present in the carrier state.

In addition to G-to-A mutations, C-to-T mutations were detected in the *pol* and *tax* genes in carrier cells as in ATL cells (see Table S2 in the supplemental material); they were also detected in other viral regions (data not shown). Comparative analyses of leukemic cell and nonleukemic cells showed that 79% of C-to-T base substitutions were polymorphisms (Table 2). However, after G-to-T mutations, C-to-T mutations were

still the most frequent mutations in ATL patients and HTLV-1 carriers.

Expression of the *hA3G* gene and activation-induced deaminase gene. In order to analyze the correlation between G-to-A mutation and *hA3G* expression, we first studied the mRNA level of *hA3G* using a real-time PCR assay. The *hA3G* gene was expressed in normal T cells, and the level of expression of *hA3G* moderately decreased in ATL cells. In ATL samples, there was no obvious correlation between the level of expression of the *hA3G* gene and the number of G-to-A mutations in the proviral genome.

Although only 21% of C-to-T base substitutions were mutations (Table 2), C-to-T mutations were the second most frequent mutation (see Table S2 in the supplemental material). Activation-induced deaminase (AID) was found to be a key factor for the switch recombination of immunoglobulin (Ig), and it is also implicated in hypermutations of Ig genes (15). AID deaminates cytosine to uracil. Aberrant expression of AID driven by NF-κB activation is found in gastric cancers and hepatomas, and AID expression has been invoked to explain

TABLE 5. Mutations in the *tax* region of 10 HTLV-1 carriers

HTLV-1 carrier (no. of clones) ^a	No. of the following single-nucleotide substitution in HTLV-1 carrier:								No. of sequences with the following sequence context for the G-to-A mutation ^b :				
	G-to-A	C-to-T	A-to-G	A-to-C	C-to-A	G-to-C	G-to-T	T-to-A	<u>GA</u>	<u>GC</u>	<u>GT</u>	<u>GG</u>	
C72 (24)	3								1				2
C79 (24)	2	4		2	1	1					2		
C63 (21)	19												19
C40 (21)	2												2
C32 (25)	4	1											4
C62 (27)	5	1								1			4
C87 (24)	1	1							1				
C5 (26)	2												2
C29 (25)	3		3										3
C82 (22)	5						1	1					5
Total	46	7	3	2	1	1	1	1	2	1	2		41

^a The number of clones studied from each carrier is shown in parentheses.
^b Mutated bases are underlined.

TABLE 6. Nonsense mutations in 10 HTLV-1 carriers

HTLV-1 carrier	No. of nonsense mutations ^a in the indicated gene and position in HTLV-1 carrier:														
	<i>pol</i>		<i>env</i>						<i>p12</i>	<i>p30/p13</i> at 7361	<i>p27^b</i>		<i>tax</i>		<i>HBZ</i>
	3955	4540	5465	6158	6342	6362	6482	6494			7652	7660	7469	8045	
C72	1/25	0/25	0/23	0/23	0/23	0/23	0/23	0/23	0/31	0/31	0/31	0/31	1/24	0/24	0/31
C79	2/20	2/20	0/20	0/20	0/20	1/20	0/20	0/20	0/24	0/24	0/24	0/24	1/24	0/24	0/24
C63	0/28	0/28	0/20	0/20	0/20	0/20	0/20	0/20	0/24	2/24	0/24	0/24	9/21	0/21	0/24
C40	0/26	13/26	0/22	0/22	0/22	0/22	0/22	0/22	0/30	0/30	0/30	0/30	2/21	0/21	0/30
C32	0/22	0/22	0/22	0/22	0/22	0/22	0/22	0/22	0/32	0/32	0/32	1/32	0/25	0/25	0/32
C62	1/22	0/22	0/25	0/25	2/25	0/25	0/25	0/25	0/28	1/28	0/28	0/28	1/27	0/27	0/28
C87	1/22	3/22	0/20	0/20	0/20	0/20	0/20	0/20	0/21	0/21	1/21	0/21	0/24	0/24	0/21
C5	0/22	19/22	0/20	0/20	0/20	1/20	0/20	0/20	0/24	0/24	0/24	1/24	1/26	0/26	0/24
C29	0/25	3/25	0/20	1/20	0/20	0/20	0/20	1/20	0/24	2/24	5/24	0/24	1/25	0/25	0/24
C82	3/27	2/27	1/20	0/20	0/20	0/20	1/20	0/20	0/24	1/24	0/24	0/24	1/22	1/22	0/24

^a Each entry *A/B* in the table shows the number of clones harboring nonsense mutations at a given nucleotide position from a given carrier (*A*) along with the number of clones analyzed in that region from that carrier (*B*). The boldface values indicate that a nonsense mutation was found.

^b Sequences of exon 2 were determined.

induced mutations observed in cancer cells (28). Therefore, we studied *AID* expression in 7 ATL cases with different proviral mutation frequencies. As shown in Fig. 4, the *AID* gene transcription increased in ATL-15. In this case, we found six C-to-T mutations in the provirus, and the sequence context showed more *AID* preferred sequences in C-to-T mutated sites compared to other cases without *AID* activation (data not shown). Thus, *AID* might also play a role in the deamination of the HTLV-1 genome in at least some ATL cases. Since *AID* is thought to target the double-stranded DNA of provirus, it likely generates C-to-T mutations in the plus strand of the provirus. However, C-to-T base substitutions contained many polymorphisms, as shown in Table 2. Therefore, correlation between C-to-T mutations and *AID* expression remains to be studied.

In vitro editing of the HTLV-1 genome by hA3G. Previous reports have suggested that HTLV-1 is relatively resistant to the antiviral effect of hA3G (4), consistent with our finding that the frequency of G-to-A mutations throughout the HTLV-1 proviral genome was low. In order to provide direct evidence that G-to-A mutations observed in the proviruses are the result of hA3G-mediated genome editing, we studied the editing effect of hA3G on the HTLV-1 genome *in vitro*. HTLV-1 viruses were generated in the presence of exogenous hA3G to allow packaging into budding viral particles, a step that is required for exerting editing activity of the cytidine deaminase. After a round of virus infection, the HTLV-1 DNA was analyzed using a highly sensitive PCR-based protocol, referred to as 3D PCR, that is capable of amplifying G-to-A hypermutated genomes as described in Materials and Methods. Differential amplification of a *tax* region from cells infected with viral particles produced in the presence of hA3G was achieved when the denaturation PCR temperature was lowered to between 84 and 86°C (data not shown). All sequences of differentially amplified PCR clones exhibited extensive G-to-A hypermutation, and the number of G-to-A transitions ranged from 8 (13%) to 29 (48%) per clone (Fig. 5). In ATL cases and in the same region, G-to-A mutations at four nucleotide positions resulted in premature stop codons in the *p30* and *tax* genes (Fig. 1). The occurrence of such mutations was observed *in vitro* when viruses were produced together with hA3G. G-to-A

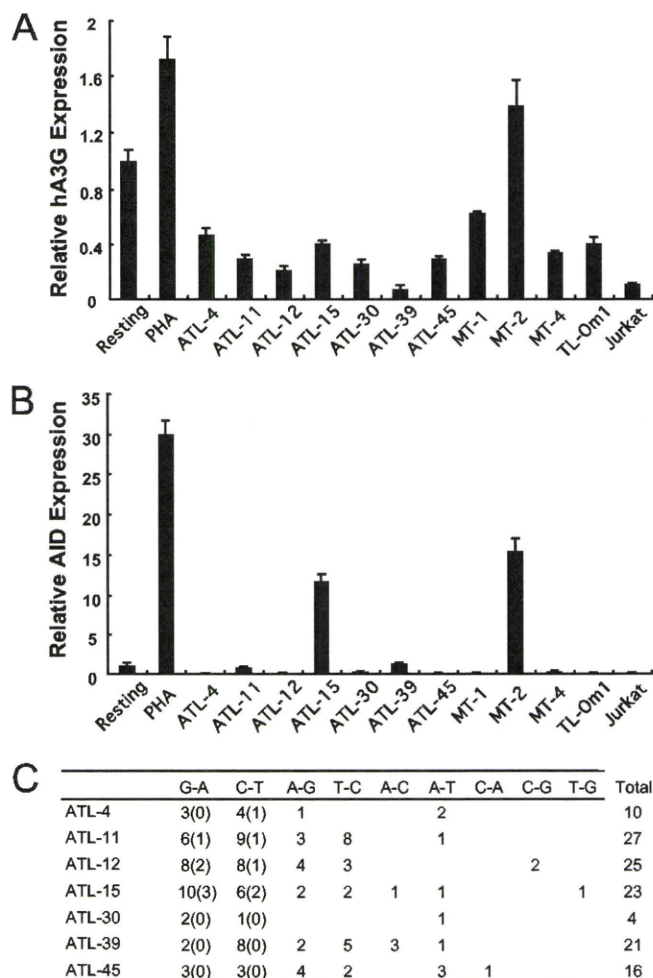


FIG. 4. Expression of the *hA3G* and *AID* genes in ATL cases and in HTLV-1-associated cell lines. Expression of the *hA3G* gene (A) and *AID* gene (B) was investigated by real-time PCR. Normal resting T cells isolated from 3 healthy blood donors were used as a control. *ACTN* was used as an internal control. Relative quantification was performed using a comparative C_T method (Applied Biosystems). Expression of *hA3G* and *AID* in resting T cells was artificially set at 1. (C) Number of nucleotide mutations in the 7 ATL cases. Among mutated sites, the number of *AID*-preferred sites is shown in parentheses.