Patient 3 had HLA-A*02:01 and A*11:01. Although peptides of CTL epitopes for both HLA alleles were available, we chose the Tax11-19 peptide for HLA-A2 because HLA-A2 has a higher frequency in Japanese individuals. After each vaccination, the patient developed a low-grade fever and dermatitis (grade 2); however, no other severe adverse events were noted. She achieved a PR with an improvement in the KPS 8 weeks after the initiation of the Tax-DC vaccine therapy. Thereafter, the level of sIL2R returned to normal (Fig 2B). The patient subsequently achieved a CR at 6 months and has remained in this status for more than 19 months after the completion of the Tax-DC vaccine therapy.

Immunological responses after the Tax-DC vaccine therapy

In Patient 1, Tax-specific CD8$^+$ CTLs (HLA-A*24:02/Tax301-309 tetramer$^+$) were detectable prior to vaccination, and their frequency in peripheral CD8$^+$ cells transiently decreased during the Tax-DC vaccine administration, then recovered and maintained a constant level with some fluctuation (Fig 3A). The IFN-γ production from Tax-specific CTL also fluctuated. It is noteworthy that a vigorous proliferative response of Tax-specific CTLs was observed in vitro in the PBMC sample obtained at 20 weeks after the initiation of the Tax-DC vaccine therapy (Fig 3B), in which the proportion of HLA-A*24:02/Tax301-309 tetramer$^+$ cells in CD8$^+$ cells increased up to 22·5% within 2 weeks of the culture. A mild proliferative response of CTLs was also observed at 12 weeks. Samples obtained from the same patient prior to vaccination lacked such strong responses, implying a functional improvement in CTLs after the Tax-DC vaccine therapy.

Similar to that observed in Patient 1, a markedly increased level of spontaneous in vitro proliferative responses of Tax-specific CTLs was observed in the PBMC samples obtained from Patient 2 at 16 weeks after the initiation of the Tax-DC vaccine therapy, although the CTLs of this patient had exhibited a proliferative response prior to vaccination to a lesser degree (Fig 3B). The IFN-γ producing response of the CTL in this patient slightly improved after vaccination and showed some peaks at later time points.

As the size of the lymph nodes in Patient 2 did not improve within the first 8 weeks, a biopsy of the inguinal lymph node was performed at 9 weeks during the study period. The tumour cells isolated from the lymph node were CD4$^+$ CD8$^+$ CCR4$^+$ (Fig 4A) and possessed HTLV-I proviruses (849·5 copies/1000 cells). However, HTLV-I Tax proteins or mRNA expression was not induced in the lymph node cells after a short-term in vitro culture, whereas the viral expression was inducible in the PBMC sample of the same patient before vaccination (Fig 4B,C).

Tax-specific CTLs were below detectable levels prior to vaccination in Patient 3. However, 2 weeks after the initiation of the vaccine therapy with Tax 11–19 peptide-pulsed DCs, CD8$^+$ Tax-specific CTLs became detectable with HLA-A*0201/Tax11-19 tetramers, but not HLA-A*1101/Tax88-96 tetramers (Fig 3A). Although the IFN-γ producing response was barely detectable because of the low CTL frequency, an in vitro proliferative response of Tax-specific CTLs was observed in the PBMC samples obtained from Patient 3 most clearly at 16 weeks of the Tax-DC vaccine therapy, upon stimulation with Tax11-19 peptides, but not Tax 88–96 peptides (Fig 3B).

In all three patients, the level of the proviral load in the peripheral blood mostly remained below 100 copies per 1000 PBMCs at least for 1 year after vaccination, with the exception of sporadic small spikes (Fig 3A).

Discussion

Although various therapeutic trials have been conducted, the prognosis of ATL remains dismal. According to the simplified ATL prognostic index (ATL-PI) (Katsuya et al, 2012), the median survival time is only 4·6, 7·0 and 16·2 months, while the 2-year overall survival rate is 6%, 17% and 37%, for patients in high-, intermediate- and low-risk groups, respectively. According to the ATL-PI, Patients 1 and 2 were classified as intermediate-risk, while Patient 3 was classified as high-risk. Therefore, it is quite unique and surprising that all three patients remained in a favourable condition, without the need for any additional anti-tumour therapy, for at least 24, 14 and 19 months respectively, after only three administrations of the Tax-DC vaccine. In particular, Patients 1 and 3 obtained PR by 8 weeks after the initiation of the Tax-DC vaccine therapy.

Although these results are exciting, we cannot completely rule out the persisting effects of lenalidomide and/or mogamulizumab, which were previously administered in each patient prior to the Tax-DC vaccine therapy. These previous treatments may also have positively contributed to the present results via their immunomodulatory effects. According to recent reports, mogamulizumab has been shown to decrease the level of CCR4$^+$ regulatory T cells (Ishida & Ueda, 2011), and lenalidomide has immunomodulatory effects indirectly enhancing the activity of natural killer and T cells (Wu et al, 2008; De Keersmaecker et al, 2012).

The biopsy specimen of a residual surface lymph node from Patient 2 contained HTLV-I proviruses, although the viral expression was not inducible in the isolated cells even after in vitro culture (Fig 4). In general, induction of Tax expression after short-term culture is observed in approximately 50% of ATL cases (Kurihara et al, 2005). In the other 50% of ATL cases, the ATL cells lack the ability to express Tax, presumably due to the genomic and epigenetic changes in the HTLV-I proviruses (Takeda et al, 2004). Given that the viral expression was inducible in PBMCs of Patient 2 obtained prior to vaccination, the absence of viral induction
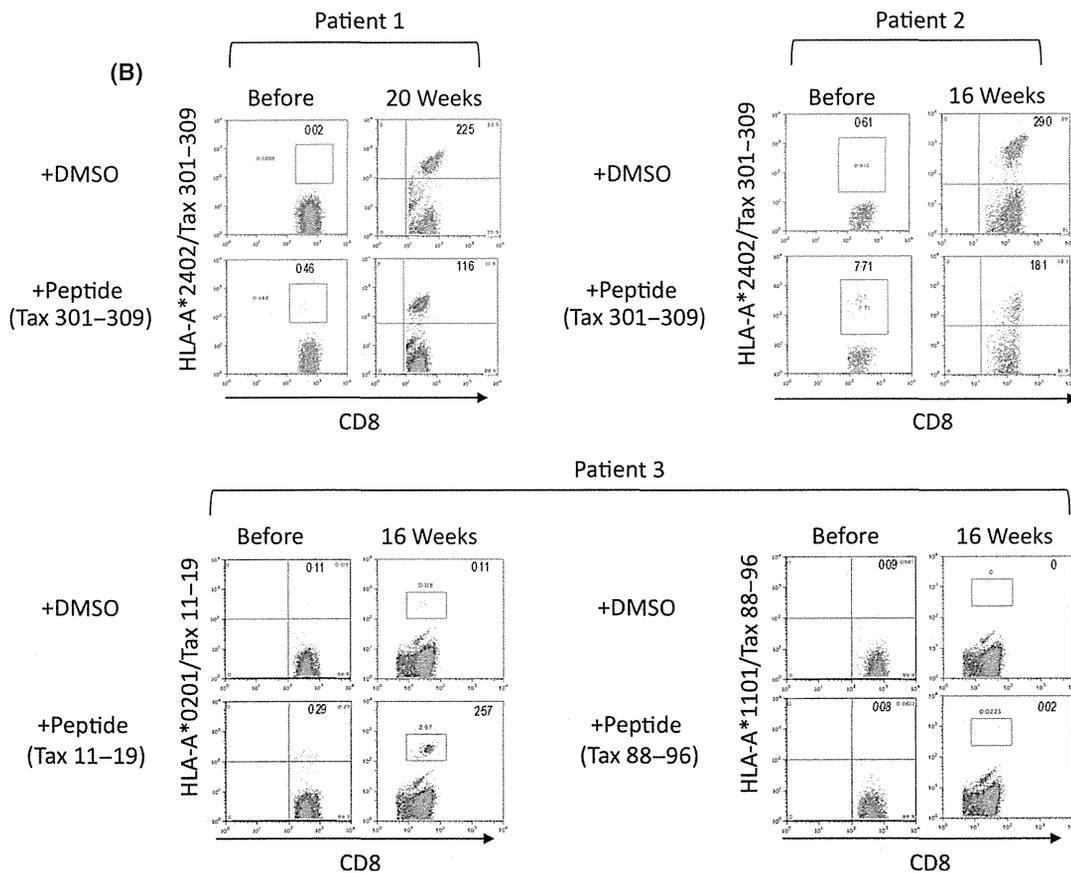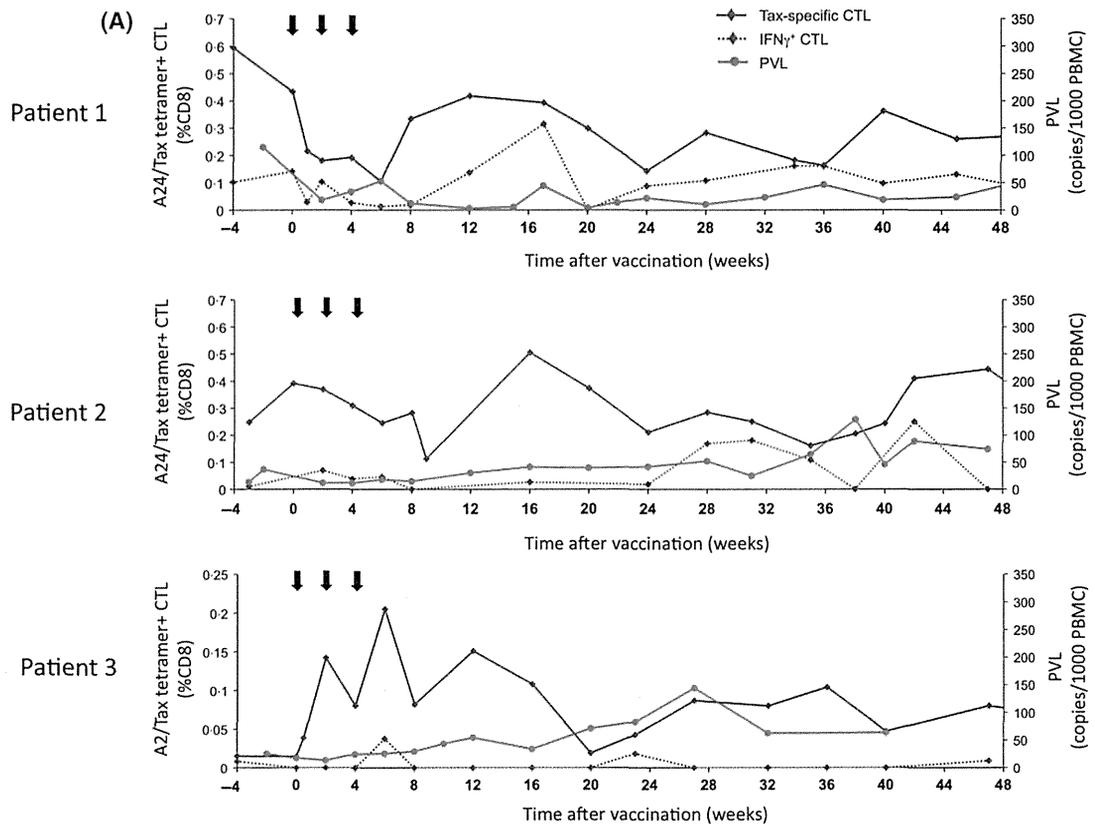
Fig 3. Immunological responses in the three patients after the Tax-DC vaccine therapy. (A) Long-term kinetics of the Tax-specific cytotoxic T cells (CTLs; % CD8+ cells, blue solid line), and γ-interferon (IFN-γ)-producing Tax-specific CTLs (% CD8+ cells, blue broken line), and human T cell leukaemia virus type-I proviral load (HTLV-I PVL) [copies/1000 peripheral blood mononuclear cells (PBMCs), red] in the peripheral blood of the three patients. Each arrow indicates administration of the vaccine. (B) The proliferative ability of the Tax-specific CTLs was evaluated using flow cytometry following incubation of the PBMCs for 13–15 d *in vitro* with cognate Tax peptide (100 nmol/l) or dimethyl sulfoxide (DMSO) in the presence of 10 u/ml of recombinant human IL2. The cells were stained with HLA/Tax tetramer-PE, anti-human CD8-PE-Cy5 mAb and anti-human CD3-FITC mAb. The values represent the percentage of tetramer+ cells/CD3+ CD8+ cells.
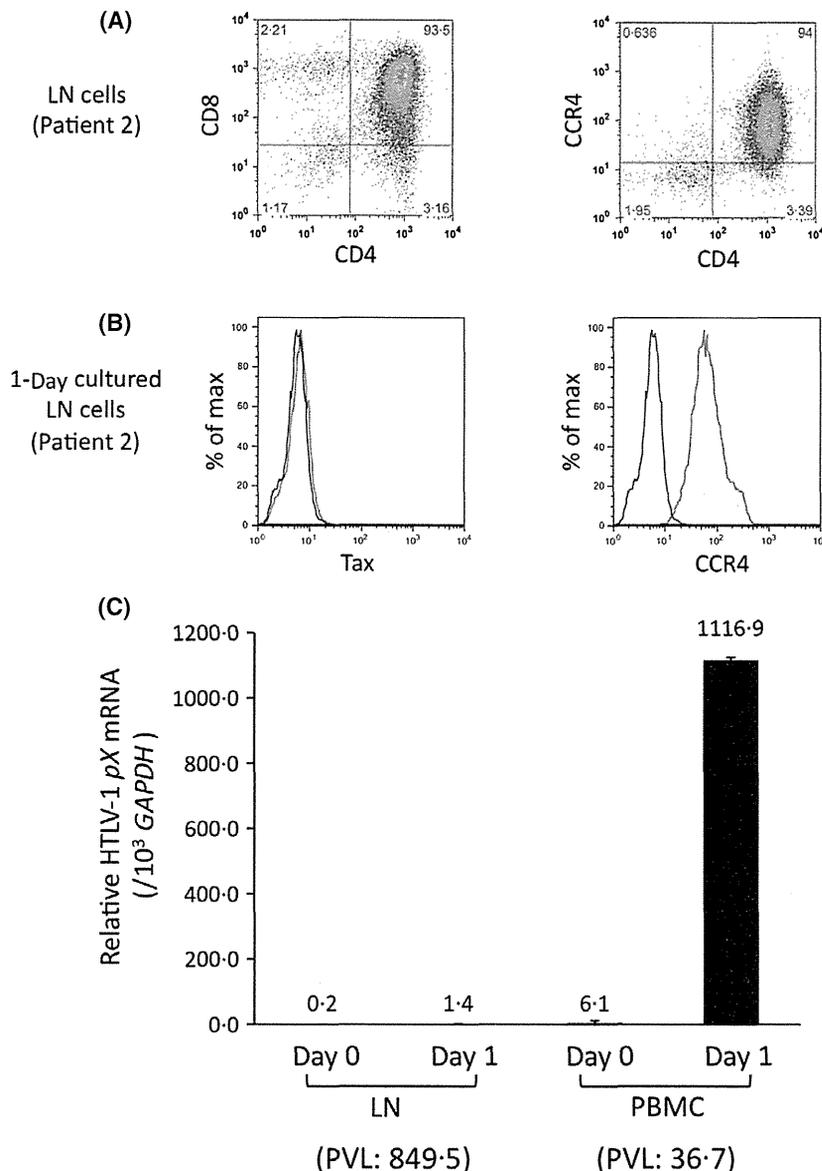
Fig 4. Absence of HTLV-I expression in the lymph node cells obtained from Patient 2. Cells were isolated from a biopsy specimen of the inguinal lymph node (LN) from Patient 2 at 9 weeks after the initiation of the Tax-DC vaccine therapy and subjected for characterization. (A) The cell surface phenotype of the LN cells immediately after isolation was analysed following staining with the indicated mAbs. (B) The intracellular Tax and CCR4 expression levels (red) in the LN cells after a 1-d culture *in vitro* were analysed following fixation of the cells with methanol. The blue histogram indicates the results of control antibody staining. (C) The HTLV-I *pX* mRNA expression levels in the LN cells before and after a 1-d culture *in vitro* were evaluated by quantitative reverse transcription polymerase chain reaction. The viral mRNA expression in peripheral blood mononuclear cells (PBMCs) obtained from the same patient before vaccination was similarly analysed as a positive control. The relative values standardized by *GAPDH* mRNA copy numbers were indicated as the means and standard deviations of duplicate samples. The proviral load (PVL) in the samples (copies/1000 cells) is indicated in parenthesis.

in the lymph node cells suggests that these tumour cells had escaped from Tax-specific CTLs.

Intriguingly, the Tax-specific CTLs demonstrated a vigorous proliferative response *in vitro* in all three patients at approximately 16–20 weeks after the initiation of the Tax-DC vaccine therapy. In particular, in Patients 1 and 2, the CTLs proliferated spontaneously without stimulation (Fig 3B). Similar phenomena have been reported in patients with HTLV-I-Associated Myelopathy/Tropical Spastic Paraparesis (Jacobson *et al*, 1990; Takamori *et al*, 2011) and occasionally in ATL patients post-HSCT (Harashima *et al*, 2005), interpreted to be the result of a normal CTL response against HTLV-I-infected cells *in vivo*. In the present study, although it is unclear whether the Tax-DC vaccine newly induced CTLs or simply activated pre-existing CTLs, Tax-specific CTLs appear to survey infected cells, at least for several months after the Tax-DC vaccine therapy, in responding to the dynamic activity of HTLV-I-infected cells *in vivo*.

In Patient 3, the Tax-specific CTLs emerged after vaccination and exhibited a clear proliferative response that peaked at 16 weeks. This response was preferentially directed toward the HLA-A2-restricted Tax epitope used for the therapy, not the HLA-A11-restricted epitope, suggesting the contribution of the Tax-DC vaccine therapy to CTL induction.

Although active CTL responses were observed in the first several months in all three patients, the responses diminished thereafter. At later time points (6 months or later) the sIL2R levels gradually increased in Patients 1 and 2 (Fig 2B). This finding suggests the need for a boosting vaccination or additional treatment to decrease the degree of immune suppression in order to maintain long-lasting anti-tumour effects.

In conclusion, the Tax-DC vaccine therapy is a safe and feasible treatment for ATL patients in stable condition. The promising clinical outcomes observed in the present study imply that the Tax-DC vaccine therapy has the potential to be an effective second-line treatment for ATL, although the anti-tumour effects of this vaccine therapy must be confirmed in further clinical trials with an increased number of patients. To our knowledge, this is the first clinical report to show the significance of a therapeutic vaccine targeting viral antigens as a new treatment modality for HTLV-I-induced malignancies. Given that Tax-specific CTL responses are impaired in patients with smouldering types of ATL and also in a small subset of asymptomatic HTLV-I carriers (Takamori *et al*, 2011), the vaccine therapy may be beneficial in these populations as well. The present study thus provides important information in a new era of anti-ATL immune therapies with the potential to be extended for prophylaxis of the disease in the future.

Acknowledgements

Authorship

Y.S. designed the study, prepared the protocol, administered the Tax-DC therapy in patients and analysed the data. A.H. designed the study, prepared the protocol, established the method of Tax-DC preparation and analysed the data. T.I. administered the Tax-DC therapy in the patients. A.S., R.T., A.T., I.C., T.F., O.M. and T.T. participated in the protocol preparation. M.M. performed the provirus analysis. N.W. and A.T. performed the flow cytometric analysis. S.T. and K.A. supervised the institutional cell processing. M.K. proposed the initial idea and concept, designed the study, prepared the protocol and analysed the data. N.U. and J.O. supervised and coordinated the clinical and basic studies. M.K., Y.S., A.H. and J.O. wrote the manuscript. All co-authors approved the final version of the manuscript.

Disclosure

Tokyo Medical and Dental University holds a patent for the Tax epitope for HLA-A*11:01, of which M. Kannagi and R. Tanosaki are included in the inventors. This epitope was not used for a vaccine in the present study. S. Takaishi receives grants and personal fees from the MEDINET Co. Ltd., outside the submitted work.

References

De Keersmaecker, B., Allard, S.D., Lacor, P., Schots, R., Thielemans, K. & Aerts, J.L. (2012) Expansion of polyfunctional HIV-specific T cells upon stimulation with mRNA electroporated dendritic cells in the presence of immunomodulatory drugs. *Journal of Virology*, **86**, 9351–9360.

Fuessel, S., Meye, A., Schmitz, M., Zastrow, S., Linne, C., Richter, K., Lobel, B., Hakenberg, O.W., Hoelig, K., Rieber, E.P. & Wirth, M.P.

(2006) Vaccination of hormone-refractory prostate cancer patients with peptide cocktail-loaded dendritic cells: results of a phase I clinical trial. *The Prostate*, **66**, 811–821.

Gill, P.S., Harrington, W. Jr, Kaplan, M.H., Ribeiro, R.C., Bennett, J.M., Liebman, H.A., Bernstein-Singer, M., Espina, B.M., Cabral, L., Allen, S., Kornblau, M.D., Pike, M.C. & Levine, A.M. (1995) Treatment of adult T-cell leukemia-lymphoma with a combination of interferon alfa and zidovudine. *New England Journal of Medicine*, **332**, 1744–1748.

Hanabuchi, S., Ohashi, T., Koya, Y., Kato, H., Hasegawa, A., Takemura, F., Masuda, T. & Kannagi, M. (2001) Regression of human T-cell leukemia virus type I (HTLV-I)-associated lymphomas in a rat model: peptide-induced T-cell immunity. *Journal of the National Cancer Institute*, **93**, 1775–1783.

Harashima, N., Kurihara, K., Utsunomiya, A., Tanosaki, R., Hanabuchi, S., Masuda, M., Ohashi, T., Fukui, F., Hasegawa, A., Masuda, T., Takaue, Y., Okamura, J. & Kannagi, M. (2004) Graft-versus-Tax response in adult T-cell

leukemia patients after hematopoietic stem cell transplantation. *Cancer Research*, **64**, 391–399.

Harashima, N., Tanosaki, R., Shimizu, Y., Kurihara, K., Masuda, T., Okamura, J. & Kannagi, M. (2005) Identification of two new HLA-A*1101-restricted tax epitopes recognized by cytotoxic T lymphocytes in an adult T-cell leukemia patient after hematopoietic stem cell transplantation. *Journal of Virology*, **79**, 10088–10092.

Hermine, O., Bouscary, D., Gessain, A., Turlure, P., Leblond, V., Franck, N., Buzyn-Veil, A., Rio, B., Macintyre, E., Dreyfus, F. & Bazarbachi, A. (1995) Brief report: treatment of adult T-cell leukemia-lymphoma with zidovudine and interferon alfa. *New England Journal of Medicine*, **332**, 1749–1751.

Hinuma, Y., Nagata, K., Hanaoka, M., Nakai, M., Matsumoto, T., Kinoshita, K.I., Shirakawa, S. & Miyoshi, I. (1981) Adult T-cell leukemia: antigen in an ATL cell line and detection of antibodies to the antigen in human sera. *Proceedings of the National Academy of Sciences of the United States of America*, **78**, 6476–6480.

Hishizawa, M., Imada, K., Kitawaki, T., Ueda, M., Kadowaki, N. & Uchiyama, T. (2004) Depletion and impaired interferon-alpha-producing capacity of blood plasmacytoid dendritic cells in human T-cell leukaemia virus type I-infected individuals. *British Journal of Haematology*, **125**, 568–575.

Hishizawa, M., Kanda, J., Utsunomiya, A., Taniguchi, S., Eto, T., Moriuchi, Y., Tanosaki, R., Kawano, F., Miyazaki, Y., Masuda, M., Nagafuji, K., Hara, M., Takanashi, M., Kai, S., Atsuta, Y., Suzuki, R., Kawase, T., Matsuo, K., Nagamura-Inoue, T., Kato, S., Sakamaki, H., Morishima, Y., Okamura, J., Ichinohe, T. & Uchiyama, T. (2010) Transplantation of allogeneic hematopoietic stem cells for adult T-cell leukemia: a nationwide retrospective study. *Blood*, **116**, 1369–1376.

Ishida, T. & Ueda, R. (2011) Immunopathogenesis of lymphoma: focus on CCR4. *Cancer Science*, **102**, 44–50.

Ishida, T., Joh, T., Uike, N., Yamamoto, K., Utsunomiya, A., Yoshida, S., Saburi, Y., Miyamoto, T., Takemoto, S., Suzushima, H., Tsukasaki, K., Nosaka, K., Fujiwara, H., Ishitsuka, K., Inagaki, H., Ogura, M., Akinaga, S., Tomonaga, M., Tobinai, K. & Ueda, R. (2012) Defucosylated anti-CCR4 monoclonal antibody (KW-0761) for relapsed adult T-cell leukemia-lymphoma: a multicenter phase II study. *Journal of Clinical Oncology*, **30**, 837–842.

Ishida, T., Hishizawa, M., Kato, K., Tanosaki, R., Fukuda, T., Takatsuka, Y., Eto, T., Miyazaki, Y., Hidaka, M., Uike, N., Miyamoto, T., Tsudo, M., Sakamaki, H., Morishima, Y., Suzuki, R. & Utsunomiya, A. (2013) Impact of graft-versus-host disease on allogeneic hematopoietic cell transplantation for adult T cell leukemia-lymphoma focusing on preconditioning regimens: nationwide retrospective study. *Biology of Blood and Marrow Transplantation*, **19**, 1731–1739.

Jacobson, S., Shida, H., McFarlin, D.E., Fauci, A.S. & Koenig, S. (1990) Circulating CD8+ cytotoxic

T lymphocytes specific for HTLV-I pX in patients with HTLV-I associated neurological disease. *Nature*, **348**, 245–248.

Jones, K.S., Petrow-Sadowski, C., Huang, Y.K., Bertolette, D.C. & Ruscetti, F.W. (2008) Cell-free HTLV-1 infects dendritic cells leading to transmission and transformation of CD4(+) T cells. *Nature Medicine*, **14**, 429–436.

Kannagi, M., Shida, H., Igarashi, H., Kuruma, K., Murai, H., Aono, Y., Maruyama, I., Osame, M., Hattori, T., Inoko, H. & Harada, S. (1992) Target epitope in the Tax protein of human T-cell leukemia virus type I recognized by class I major histocompatibility complex-restricted cytotoxic T cells. *Journal of Virology*, **66**, 2928–2933.

Katsuya, H., Yamanaka, T., Ishitsuka, K., Utsunomiya, A., Sasaki, H., Hanada, S., Eto, T., Moriuchi, Y., Saburi, Y., Miyahara, M., Sueoka, E., Uike, N., Yoshida, S., Yamashita, K., Tsukasaki, K., Suzushima, H., Ohno, Y., Matsuoka, H., Jo, T., Suzumiya, J. & Tamura, K. (2012) Prognostic index for acute- and lymphoma-type adult T-cell leukemia/lymphoma. *Journal of Clinical Oncology*, **30**, 1635–1640.

Kurihara, K., Harashima, N., Hanabuchi, S., Masuda, M., Utsunomiya, A., Tanosaki, R., Tomonaga, M., Ohashi, T., Hasegawa, A., Masuda, T., Okamura, J., Tanaka, Y. & Kannagi, M. (2005) Potential immunogenicity of adult T cell leukemia cells in vivo. *International Journal of Cancer*, **114**, 257–267.

Lee, B., Tanaka, Y. & Tozawa, H. (1989) Monoclonal antibody defining tax protein of human T-cell leukemia virus type-I. *The Tohoku Journal of Experimental Medicine*, **157**, 1–11.

Linette, G.P., Zhang, D., Hodi, F.S., Jonasch, E.P., Longerich, S., Stowell, C.P., Webb, I.J., Daley, H., Soiffer, R.J., Cheung, A.M., Eapen, S.G., Fee, S.V., Rubin, K.M., Sober, A.J. & Haluska, F.G. (2005) Immunization using autologous dendritic cells pulsed with the melanoma-associated antigen gp100-derived G280-9V peptide elicits CD8+ immunity. *Clinical Cancer Research*, **11**, 7692–7699.

Makino, M., Wakamatsu, S., Shimokubo, S., Arima, N. & Baba, M. (2000) Production of functionally deficient dendritic cells from HTLV-I-infected monocytes: implications for the dendritic cell defect in adult T cell leukemia. *Virology*, **274**, 140–148.

Nagayama, H., Sato, K., Morishita, M., Uchimaru, K., Oyaizu, N., Inazawa, T., Yamasaki, T., Enomoto, M., Nakaoka, T., Nakamura, T., Maekawa, T., Yamamoto, A., Shimada, S., Saida, T., Kawakami, Y., Asano, S., Tani, K., Takahashi, T.A. & Yamashita, N. (2003) Results of a phase I clinical study using autologous tumour lysate-pulsed monocyte-derived mature dendritic cell vaccinations for stage IV malignant melanoma patients combined with low dose interleukin-2. *Melanoma Research*, **13**, 521–530.

Ohashi, T., Hanabuchi, S., Kato, H., Tateno, H., Takemura, F., Tsukahara, T., Koya, Y., Hasegawa, A., Masuda, T. & Kannagi, M. (2000) Prevention of adult T-cell leukemia-like

lymphoproliferative disease in rats by adoptively transferred T cells from a donor immunized with human T-cell leukemia virus type 1 Tax-coding DNA vaccine. *Journal of Virology*, **74**, 9610–9616.

Okamura, J., Utsunomiya, A., Tanosaki, R., Uike, N., Sonoda, S., Kannagi, M., Tomonaga, M., Harada, M., Kimura, N., Masuda, M., Kawano, F., Yufu, Y., Hattori, H., Kikuchi, H. & Saburi, Y. (2005) Allogeneic stem-cell transplantation with reduced conditioning intensity as a novel immunotherapy and antiviral therapy for adult T-cell leukemia/lymphoma. *Blood*, **105**, 4143–4145.

Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D. & Gallo, R.C. (1980) Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America*, **77**, 7415–7419.

Rende, F., Cavallari, I., Corradin, A., Silic-Benussi, M., Toulza, F., Toffolo, G.M., Tanaka, Y., Jacobson, S., Taylor, G.P., D'Agostino, D.M., Bangham, C.R. & Ciminale, V. (2011) Kinetics and intracellular compartmentalization of HTLV-1 gene expression: nuclear retention of HBZ mRNAs. *Blood*, **117**, 4855–4859.

Takamori, A., Hasegawa, A., Utsunomiya, A., Maeda, Y., Yamano, Y., Masuda, M., Shimizu, Y., Tamai, Y., Sasada, A., Zeng, N., Choi, I., Uike, N., Okamura, J., Watanabe, T., Masuda, T. & Kannagi, M. (2011) Functional impairment of Tax-specific but not cytomegalovirus-specific CD8+ T lymphocytes in a minor population of asymptomatic human T-cell leukemia virus type 1-carriers. *Retrovirology*, **8**, 100.

Takeda, S., Maeda, M., Morikawa, S., Taniguchi, Y., Yasunaga, J., Nosaka, K., Tanaka, Y. & Matsuoka, M. (2004) Genetic and epigenetic inactivation of tax gene in adult T-cell leukemia cells. *International Journal of Cancer*, **109**, 559–567.

Tanosaki, R., Uike, N., Utsunomiya, A., Saburi, Y., Masuda, M., Tomonaga, M., Eto, T., Hidaka, M., Harada, M., Choi, I., Yamanaka, T., Kannagi, M., Matsuoka, M. & Okamura, J. (2008) Allogeneic hematopoietic stem cell transplantation using reduced-intensity conditioning for adult T cell leukemia/lymphoma: impact of antithymocyte globulin on clinical outcome. *Biology of Blood and Marrow Transplantation*, **14**, 702–708.

Thomas-Kaskel, A.K., Zeiser, R., Jochim, R., Robbel, C., Schultze-Seemann, W., Waller, C.F. & Veelken, H. (2006) Vaccination of advanced prostate cancer patients with PSCA and PSA peptide-loaded dendritic cells induces DTH responses that correlate with superior overall survival. *International Journal of Cancer*, **119**, 2428–2434.

Tsukasaki, K., Hermine, O., Bazarbachi, A., Ratner, L., Ramos, J.C., Harrington, W. Jr, O'Mahony, D., Janik, J.E., Bittencourt, A.L., Taylor, G.P., Yamaguchi, K., Utsunomiya, A., Tobinai, K. & Watanabe, T. (2009) Definition, prognostic

factors, treatment, and response criteria of adult T-cell leukemia-lymphoma: a proposal from an international consensus meeting. *Journal of Clinical Oncology*, **27**, 453–459.

Tsukasaki, K., Tobinai, K., Hotta, T. & Shimoyama, M. (2012) Lymphoma study group of JCOG. *Japanese Journal of Clinical Oncology*, **42**, 85–95.

Uchiyama, T., Yodoi, J., Sagawa, K., Takatsuki, K. & Uchino, H. (1977) Adult T-cell leukemia: clinical and hematologic features of 16 cases. *Blood*, **50**, 481–492.

Ueda, Y., Itoh, T., Nukaya, I., Kawashima, I., Okugawa, K., Yano, Y., Yamamoto, Y., Naitoh, K., Shimizu, K., Imura, K., Fuji, N., Fujiwara, H., Ochiai, T., Itoi, H., Sonoyama, T., Hagiwara, A.,

Takesako, K. & Yamagishi, H. (2004) Dendritic cell-based immunotherapy of cancer with carcinoembryonic antigen-derived, HLA-A24-restricted CTL epitope: clinical outcomes of 18 patients with metastatic gastrointestinal or lung adenocarcinomas. *International Journal of Oncology*, **24**, 909–917.

Utsunomiya, A., Miyazaki, Y., Takatsuka, Y., Hanada, S., Uozumi, K., Yashiki, S., Tara, M., Kawano, F., Saburi, Y., Kikuchi, H., Hara, M., Sao, H., Morishima, Y., Kodera, Y., Sonoda, S. & Tomonaga, M. (2001) Improved outcome of adult T cell leukemia/lymphoma with allogeneic hematopoietic stem cell transplantation. *Bone Marrow Transplantation*, **27**, 15–20.

Wierecky, J., Muller, M.R., Wirths, S., Halder-Oehler, E., Dorfel, D., Schmidt, S.M., Hantschel, M., Brugger, W., Schroder, S., Horger, M.S., Kanz, L. & Brossart, P. (2006) Immunologic and clinical responses after vaccinations with peptide-pulsed dendritic cells in metastatic renal cancer patients. *Cancer Research*, **66**, 5910–5918.

Wu, L., Adams, M., Carter, T., Chen, R., Muller, G., Stirling, D., Schafer, P. & Bartlett, J.B. (2008) lenalidomide enhances natural killer cell and monocyte-mediated antibody-dependent cellular cytotoxicity of rituximab-treated CD20+ tumor cells. *Clinical Cancer Research*, **14**, 4650–4657.

VIROLOGY JOURNAL
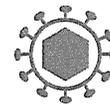
RESEARCH **Open Access**

HTLV-1 proviral integration sites differ between asymptomatic carriers and patients with HAM/TSP

Heather A Niederer[1], Daniel J Laydon[1], Anat Melamed[1], Marjet Elemans[1], Becca Asquith[1], Masao Matsuoka[2] and Charles RM Bangham[1*]

Abstract

Background: HTLV-1 causes proliferation of clonal populations of infected T cells in vivo, each clone defined by a unique proviral integration site in the host genome. The proviral load is strongly correlated with odds of the inflammatory disease HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP). There is evidence that asymptomatic HTLV-1 carriers (ACs) have a more effective CD8 + T cell response, including a higher frequency of HLA class I alleles able to present peptides from a regulatory protein of HTLV-1, HBZ. We have previously shown that specific features of the host genome flanking the proviral integration site favour clone survival and spontaneous expression of the viral transactivator protein Tax in naturally infected PBMCs ex vivo. However, the previous studies were not designed or powered to detect differences in integration site characteristics between ACs and HAM/TSP patients. Here, we tested the hypothesis that the genomic environment of the provirus differs systematically between ACs and HAM/TSP patients, and between individuals with strong or weak HBZ presentation.

Methods: We used our recently described high-throughput protocol to map and quantify integration sites in 95 HAM/TSP patients and 68 ACs from Kagoshima, Japan, and 75 ACs from Kumamoto, Japan. Individuals with 2 or more HLA class I alleles predicted to bind HBZ peptides were classified 'strong' HBZ binders; the remainder were classified 'weak binders'.

Results: The abundance of HTLV-1-infected T cell clones in vivo was correlated with proviral integration in genes and in areas with epigenetic marks associated with active regulatory elements. In clones of equivalent abundance, integration sites in genes and active regions were significantly more frequent in ACs than patients with HAM/TSP, irrespective of HBZ binding and proviral load. Integration sites in genes were also more frequent in strong HBZ binders than weak HBZ binders.

Conclusion: Clonal abundance is correlated with integration in a transcriptionally active genomic region, and these regions may promote cell proliferation. A clone that reaches a given abundance in vivo is more likely to be integrated in a transcriptionally active region in individuals with a more effective anti-HTLV-1 immune response, such those who can present HBZ peptides or those who remain asymptomatic.

Keywords: HTLV-1, Human T cell lymphotropic virus-type 1, HBZ, HTLV-1 basic leucine zipper factor, HAM/TSP, HTLV-1-associated myelopathy/tropical spastic paraparesis, Integration site, CD8+ T cell

* Correspondence: c.bangham@imperial.ac.uk
[1]Department of Immunology, Wright-Fleming Institute, Imperial College London, London W2 1PG, UK
Full list of author information is available at the end of the article

Background

Human T cell lymphotropic virus-type 1 (HTLV-1) is estimated to infect over 10 million people [1], and is endemic in sub-Saharan Africa, the south islands of Japan, the Caribbean and parts of South America. HTLV-1 is primarily found in CD4$^+$ T cells, where predominantly only a single copy of the virus integrates into the genome [2]. The virus is almost 100% cell-associated, and the viral burden is defined as the fraction of PBMCs that carry the integrated provirus, termed the proviral load (PVL). Infected cells proliferate in vivo, producing clonal populations of cells, each defined by its unique proviral integration site. The viral regulatory proteins Tax and HTLV-1 basic leucine zipper factor (HBZ) are known to drive proliferation of the infected cells [3-5].

More than 90% of HTLV-1-infected individuals remain lifelong asymptomatic carriers (AC), but 1–6% develop an aggressive malignancy known as adult T cell leukaemia/lymphoma (ATLL). A further 0.25 to 4% develop a chronic inflammatory disease of the central nervous system, HAM/TSP, characterised by a slowly progressive spastic paraparesis with pain and neurogenic bladder disturbance [6]. Risk factors for HAM/TSP include female gender and high PVL [7].

There is strong evidence that the CD8$^+$ T cell response is important in limiting PVL, and reducing the risk of HAM/TSP [8], although innate immunity also plays a role in the host response to HTLV-1 [9]. Certain HLA class I alleles are associated with a reduction in PVL and prevalence of HAM/TSP, in particular HLA-A*02 and Cw*08 in a population from Southern Japan [10]. Tax is the dominant CD8$^+$ T cell target antigen of HTLV-1 [11,12]: Tax escape mutations in the HLA-A2-restricted epitope Tax 11-19 are more frequent in individuals with the HLA-A2 allele [13], and Tax expression is frequently silenced in the expanded clone in ATLL by mutations in *tax* or methylation or deletion of the 5'LTR [14-17]. The rate of lysis of Tax$^+$CD4$^+$ cells by CD8$^+$ cells has been inversely correlated with PVL [18], although Tax mRNA is virtually undetectable directly ex vivo. Individuals who remain asymptomatic were shown to have a lower PVL than those with HAM/TSP at a given lysis rate [18], and had a greater CD8$^+$ T-cell lytic efficiency as measured by proportion of Tax-specific CTL which degranulate when exposed to their cognate epitope ex vivo [19].

Unlike Tax, HBZ expression is uniformly maintained in HTLV-1-infected T cells, including ATLL cells [4], and this expression correlates with PVL in both ACs and patients with HAM/TSP [20]. On average, HBZ peptides bind to HLA class I alleles with lower affinity than Tax peptides, and the frequency of HBZ-specific CD8+ T cells [21] is correspondingly lower. HBZ expression may be maintained because it can drive expansion of an infected clone without presenting a strong target to the CD8$^+$ T cell response. The frequency of HLA class I alleles that are predicted to strongly bind HBZ peptides is greater in ACs than patients with HAM/TSP, and is inversely correlated with PVL in each group [21]. These observations suggest that a CD8$^+$ T-cell response to the HBZ protein is protective against HTLV-1-associated inflammatory disease.

The equilibrium abundance in vivo of a particular HTLV-1-infected T-cell clone is the result of the interplay between the proliferation of the clone and counter-selection by the host response, chiefly the CD8$^+$ T cell response. Both factors are governed by the program of proviral expression by the clone. Since the proviral sequence is very stable [22], the chief unique attribute of each HTLV-1-infected T-cell clone is the genomic position of the integrated provirus – the proviral integration site. Specific features of the genomic environment of the HTLV-1 proviral integration site are associated with the frequency of spontaneous reactivation of Tax expression ex vivo [23]. Integration in the same transcriptional orientation as a flanking host gene is associated with suppression of Tax expression: same-sense orientation is more frequent in high-abundance clones, and more frequent in vivo than during in vitro infection, suggesting that this orientation confers a selective advantage by allowing escape from the Tax-specific CD8$^+$ T cell response [23]. There are no published data on the influence of the integration site environment on HBZ expression.

Since the genomic integration site influences proviral expression, we reasoned that the selection pressure exerted by a protective immune response will alter the abundance of clones which have integration site genomic environments with certain characteristics. Our previous reports were neither designed nor powered to examine the relationship between the integration site environment and either disease status or host immunogenetics. In this study, we investigated the differences in integration site environment between Japanese individuals who remained AC and those who developed HAM/TSP, and between those that differed in their capacity to present HBZ peptides on protective HLA class I alleles. We report that integration sites in genes and active regions are significantly more frequent in ACs than in patients with HAM/TSP, even after accounting for clone abundance and PVL. Integration sites in genes are also more frequent in strong HBZ binders.

Results

Possession of HBZ-binding HLA class I alleles was associated with reduced HTLV-1 proviral load, but did not affect oligoclonality of the HTLV-1 infected cell population or total number of clones by load

We have previously shown that in both ACs and patients with HAM/TSP in Kagoshima prefecture (Southern Japan),

Table 1 Total patients analysed and integration sites identified

| | Kagoshima | | | | | | Kumamoto | | |
| | HAM/TSP | | | AC | | | AC | | |
Number of patients in:	Strong HBZ	Weak HBZ	*Total*	Strong HBZ	Weak HBZ	*Total*	Strong HBZ	Weak HBZ	*Total*
Whole cohort	81	148	***229***	92	110	***202***	34	64	***98***
All LM-PCR	42	62	***104***	41	59	***100***	34	64	***98***
LM-PCR QC pass	39	56	***95***	30	38	***68***	24	51	***75***
Number of UIS* with abundance:									
<1 per 10,000 PBMC	7954	12341	***20295***	1451	5485	***6936***	4572	9744	***14316***
1-10 per 10,000 PBMC	10051	12429	***22480***	2178	2515	***4693***	1034	1853	***2887***
>10 per 10,000 PBMC	219	364	***583***	102	67	***169***	42	40	***82***
All			***43358***			***11798***			***17285***

*UIS identified from LM-PCR QC pass samples only.
AC: Asymptomatic carrier, UIS: Unique integration site, QC pass: samples that passed LM-PCR quality controls (QC).

the number of HLA class I A and B alleles predicted to bind HBZ epitopes is inversely correlated with PVL [21]. Here, we extended this analysis with a second southern Japanese AC cohort, from nearby Kumamoto prefecture (Table 1). The predicted ability of individual HLA class I A and B alleles to bind HBZ epitopes was determined using the rank of the top HBZ-binding peptide in the peptide binding prediction software Metaserver, as previously described [21] (Table 2). We then used linear regression to analyse the relationship between the number of HBZ-binding alleles and PVL. In the Kumamoto cohort, as in the two Kagoshima cohorts, there was an inverse correlation between log PVL and the number of HBZ-binding alleles (linear regression: slope = -0.12, Table 3). This correlation was significant in a regression analysis of the combined Kagoshima cohorts (p = 0.02) as previously reported, and all three cohorts combined (p = 0.006), although it did not reach significance in the smaller Kumamoto cohort alone.

We used our recently developed high-throughput protocol [24] to map and quantify the abundance of unique proviral integration sites in the HAM/TSP cohort and the two AC cohorts (Table 1). Sufficient DNA was not available from all subjects, but the median PVL of the HAM/TSP samples that passed LM-PCR quality controls (QC) was not significantly different from that of the full HAM/TSP cohort (Additional file 1: Figure S1A). In the AC cohorts the median PVL of successfully analysed samples exceeded that of all ACs, because a proportion of ACs have a PVL too low to permit accurate quantification [7] and LM-PCR. Since this reduces the difference in median PVL between the analysed AC and HAM/TSP cohorts, our results represent a conservative estimate of the difference between ACs and patients with HAM/TSP.

Since strong HBZ peptide binding is associated with efficient control of PVL, we tested whether this also altered the frequency distribution of the infected cell population. Individuals were designated strong HBZ binders if they

Table 2 HLA class I alleles divided by HBZ epitope binding status

Strong HBZ binding alleles	Weak HBZ binding alleles
A0201, A0206, A0207, A0210	A0101
A2601, A2602, A2603	A0203
A3001	A0301, A0302
A3303	A1101
	A2402
	A3101
	A3201
B3701	B0702
B4001, B4002, B4006	B1301, B1302
B41	B1501, B1507, B1511, B1518
B4403	B2704
	B3501, B3520, B3532/B3568
	B3802
	B3901
	B4003, B4009/B4027/B40105
	B4601
	B4801
	B5101
	B5201
	B5401
	B5502, B5504
	B5601, B5605
	B5801
	B5901, B5902
	B6701

Table 3 Linear regression of HBZ binding alleles to PVL

Outcome	Predictor	Cohort*	n	Slope	p value	Controlling for:
Log PVL	# HBZ binding alleles	Kagoshima HAM, AC	HAM: 221 AC: 200	-0.12	0.021	Age, sex, disease status
Log PVL	# HBZ binding alleles	Kumamoto AC	AC: 98	-0.12	0.28	
Log PVL	# HBZ binding alleles	Kagoshima HAM, AC & Kumamoto AC	HAM: 229 AC: 300	-0.13	0.0058	Disease status

*Only individuals who had all the relevant information were included in the regression. Age, sex and disease status were included as factors where possible as they may vary with PVL.
Age and sex information was not available for the Kumamoto AC cohort.
PVL: proviral load, AC: Asymptomatic carrier, #: number of.

carried two or more predicted HBZ binding HLA class I alleles. The total number of HTLV-1 proviruses mapped in strong HBZ binders was similar to that in weak HBZ binders in each cohort (Additional file 1: Figure S1B), allowing unbiased estimates of clone structure. We estimated the total number of integration site clones in the blood using our novel in silico approach [25]. The estimated number of clones correlated with PVL in all cohorts (Figure 1A). There was no significant difference in estimated total number of clones in the blood, after stratifying by PVL, between AC and HAM/TSP cohorts, or between strong HBZ binders and weak HBZ binders (Figure 1A). The Oligoclonality Index (OCI) quantifies the diversity in clone abundance in an infected T cell population: an OCI approaching 1 indicates an essentially monoclonal population, whereas an OCI of 0 indicates that all clones have the same abundance. Consistent with our previous observations [24], OCI did not correlate with PVL and did not differ between ACs and HAM/TSP cohorts (Figure 1B); there was also no difference in OCI between strong and weak binders of HBZ in any cohort (Figure 1C).


Figure 1 Host disease status and HBZ-binding capacity do not alter total unique HTLV-1-infected clones or oligoclonality. HTLV-1 infected Japanese asymptomatic carriers (AC; triangles) from Kagoshima and Kumamoto prefectures and HAM/TSP patients (circles) from Kagoshima were stratified on the basis of predicted HBZ peptide binding affinity of host HLA class I alleles (strong binders, sHBZ, filled symbols; weak binders, wHBZ, open symbols). Each symbol represents one individual. Genomic DNA samples were processed by sonication-based LM-PCR and samples passing data quality checks were analysed further, as previously described ([23,24]). **(A)** The DivE method was used to estimate the total number of clones in the blood in each subject; the total clone number was positively correlated (Spearman correlation) with proviral load in each cohort both HAM (blue, p = 0.0004) and AC (black, p = 0.0001)). There was no significant difference in slope of the linear regression line between AC and HAM/TSP cohorts (likelihood ratio test, p = 0.30) or between sHBZ (continuous line) and wHBZ (dashed line) in the combined AC and HAM/TSP cohorts (likelihood ratio test, p = 0.06) indicating that there is no systematic difference in total clone number by disease status or HBZ binding status. **(B)** The oligoclonality index (OCI), which quantifies the diversity in observed clone abundance in each sample, was not significantly correlated with proviral load in ACs or HAM/TSP (Spearman correlation) and there was no significant difference in slope of the linear regression line between AC and HAM/TSP cohorts. **(C)** The median OCI did not differ significantly by HBZ binding status within individual cohorts (Mann-Whitney U test). Bars denote median +/- interquartile range.

Integration in transcriptionally active areas was positively selected in asymptomatic carriers and was associated with high clonal abundance

Detected integration sites were binned according to their absolute abundance in 10,000 PBMCs (Additional file 1: Figure S1C). In this study, we compared the frequency of transcriptionally active sites between ACs and HAM/TSP patients, after stratifying by host HBZ binding strength and absolute clonal abundance. There was no significant difference in the genomic environment of integration sites between the AC cohorts from Kumamoto and Kagoshima (Additional file 1: Figure S2B, Additional file 1: Figure S3A), which enabled combination of the two cohorts for further statistical analysis. We previously

reported that the proportion of integration sites located within a gene and in active genomic regions increased in clones which had a higher absolute abundance; conversely that high abundance clones had a reduced frequency of sites modified with inhibitory epigenetic marks [24]. Here, we also observe significant correlations between clonal abundance and integration in a gene and frequency of epigenetic marks associated with active genomic regions (Figure 2, Additional file 1: Figure S3).

The proportion of integration sites within genes was significantly higher in ACs than in HAM/TSP patients (Figure 2A). The proportion of integration sites within genes did not correlate with PVL (Additional file 1: Figure S2A), suggesting that the difference between ACs and patients
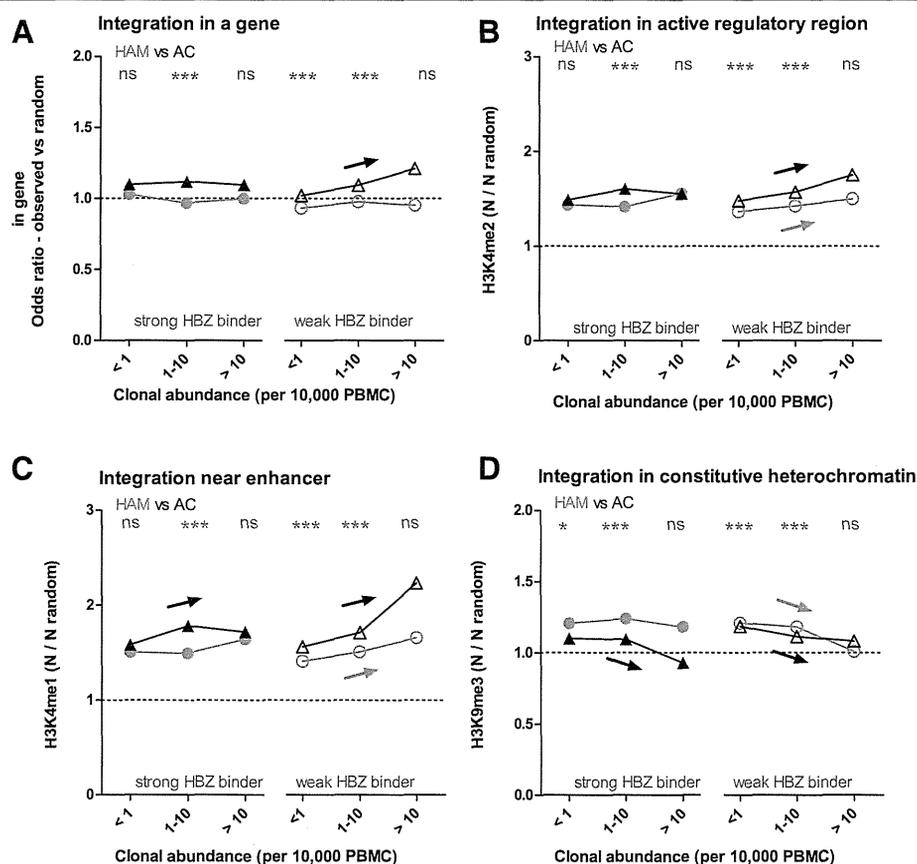
Figure 2 Transcriptionally active integration sites associate with AC status and clone abundance in weak HBZ binders. HTLV-1 unique integration sites from Japanese asymptomatic carriers (AC, black triangles) from Kagoshima and Kumamoto were compared to those from HAM/TSP patients (blue circles, from Kagoshima). Integration sites were stratified on the basis of predicted HBZ peptide binding affinity of host HLA class I alleles (strong binders, sHBZ, filled symbols; weak binders, wHBZ, open symbols) and binned by absolute abundance. Data is expressed versus an in silico generated random integration site dataset. **(A)** AC individuals (Kagoshima and Kumamoto cohorts combined) had a greater proportion of integration sites in genes than HAM/TSP patients (chi-squared test). Percentage of clones with integration sites in genes was correlated with clone abundance only in individuals with HLA class I alleles which could not bind HBZ (arrow, significant chi-squared test for trend). **(B)** Asymptomatic carriers had a higher frequency than HAM/TSP of H3K4me2 marks, enriched in transcriptionally active areas, within 10 Kb of integration sites (Mann-Whitney U test), and **(C)** a higher frequency of H3K4me1 marks, associated with enhancers. Mean epigenetic mark frequency near integration sites in a bin is divided by frequency near random sites. **(D)** In contrast, AC had a lower frequency of H3K9me3 marks, associated with constitutively heterochromatic DNA. Spearman correlation shows a significant (arrow) association between epigenetic mark frequency and log absolute abundance. Statistical comparisons AC vs HAM by Mann-Whitney U test after correction for multiple testing: * $0.05 > p > 0.01$, ** $0.01 > p > 0.001$, *** $p < 0.001$.

with HAM/TSP could not be explained simply by a lower PVL in ACs. There was no significant difference between ACs and patients with HAM/TSP in the proportion of integration sites in the same transcriptional orientation as the nearest host gene (Additional file 1: Figure S2C) or nearby transcriptional start site (Additional file 1: Figure S2D).

Compared to patients with HAM/TSP, integration sites from ACs had higher counts of H3K4me2 and H3K27ac marks within 10 Kb (indicating active regulatory regions and open chromatin; Figure 2B, Additional file 1: Figure S3B) and higher counts of H3K4me1 (associated with enhancers; Figure 2C). Conversely, integration sites had higher counts of H3K9me3 (associated with constitutively repressed chromatin) in patients with HAM/TSP than ACs (Figure 2D). There was no difference in the frequency of H3K27me3 marks, which are associated with facultative heterochomatin (Additional file 1: Figure S3C). Within genes, integration sites marked by H4K20me1 (enriched at 5′ ends of actively expressed gene bodies) were more frequent in ACs (Additional file 1: Figure S3D), but the percentage of integration sites marked by H3K36me3 (enriched at 3′ ends of actively expressed gene bodies) did not consistently correlate with either clone abundance or disease status (Additional file 1: Figure S3E). The increased frequency of active sites and sites in genes in AC individuals was also significant when the characteristics of all integration sites within a patient were averaged, and the cohorts analysed at the patient level (Additional file 1: Figure S4).

Since transcriptional activity of the genome is correlated with gene density, we used a multivariable logistic regression model to simultaneously test the independence of association of integration in a gene, active region and heterochomatin with disease status. This association was tested both at the level of individual integration sites, and using a single value per patient representing the averaged integration sites from that patient. The results showed that integration within a host gene and within active and inhibitory genomic regions were each independently associated with disease status, after controlling for clone abundance and host PVL and predicted HBZ binding affinity (Table 4 and Additional file 1: Figure S3F).

Controlling for HBZ affinity using counts of HBZ binding alleles per patient, rather than a strong/weak designation, gave a very similar significance and odds ratio.

We next compared the frequency of transcriptionally active sites between individuals with strong or weak HBZ binding potential. Integration sites from HBZ strong binders were more likely to be in a gene than those from weak binders, even when host disease status and clone absolute abundance were taken into account (Table 5); there was no significant association with frequency of H3K4me2 or H3K9me3 epigenetic marks. There was a similar odds ratio between strong and weak HBZ binders in the percentage of integration sites in a gene per patient, although at this level of individual patients the association did not reach statistical significance.

Discussion

At all levels of clone abundance, ACs had a significantly higher frequency than patients with HAM/TSP of integration sites within host genes and in genomic regions marked by activating epigenetic modifications. This enrichment was associated with disease status per se and was independent of variation in proviral load. The odds ratio of this enrichment in each case was modest (~1.1), however the finding that both integration in a gene and integration in an active genomic region were independently associated with AC status strongly suggests a consistent underlying biological mechanism. The question arises: what are the forces that favour selective survival of these integration sites (i.e. in active transcribed regions) in ACs?

We previously reported that Tax-silenced proviruses were more likely to lie in the same orientation as a flanking host gene or nearby upstream transcriptional start site; we concluded that this effect may be attributed to transcriptional interference [23]. The integration site locations that inhibit Tax expression were also associated with increased clone abundance. Consistent with these two findings, more abundant clones were less likely to express Tax [23]. These observations raised the possibility that a provirus integrated in the same orientation as a host gene might enjoy a selective advantage in individuals with an

Table 4 Proviral integration within genes and active genomic regions are independently associated with disease status

Data	n	Outcome		In gene*	# H3K4me2	# H3K9me3
Integration sites	AC: 29083 HAM/TSP: 43358	AC vs HAM	OR	1.08	1.13	0.88
			p value	4.2×10^{-5}	4.8×10^{-7}	2.1×10^{-6}
Patients	AC: 143 HAM/TSP: 95	AC vs HAM	OR	1.12	1.14	1.0
			p value	0.010	0.011	1.0

Additional factors controlled in multiple logistic regression model: Host HBZ binding status and proviral load, and log absolute clone abundance (per integration site) or log average absolute clone abundance (per patient).
*Percent of integration sites in gene per patient in patient-level analysis.
#: Number of specified epigenetic marks within 10 Kb of integration site. Averaged per patient in patient-level analysis.
OR: Odds ratio, AC: Asymptomatic carrier.

Table 5 Proviral integration within genes and active genomic regions is associated with disease status

Data	n	Outcome		In gene*
Integration sites	sHBZ: 27603 wHBZ: 45439	sHBZ vs wHBZ	OR	1.04
			p value	0.007
Patients	sHBZ: 93 wHBZ: 146	sHBZ vs wHBZ	OR	1.03
			p value	0.2

Additional factors controlled in multiple logistic regression model: Disease status and log absolute clone abundance (per integration site) or log average absolute clone abundance (per patient).
*Percent of integration sites in gene per patient in patient-level analysis.
OR: Odds ratio, sHBZ: strong HBZ binder, wHBZ: weak HBZ binder.

effective immune response because it is less exposed to the strong anti-Tax CD8$^+$ T cell response. However, the integration site environments associated with Tax silencing were not those associated with an increased frequency in ACs compared to HAM/TSP in this study.

A second possibility is that a more efficient cellular and innate immune response in ACs [8] means that a clone needs a greater proliferative capacity to reach a given absolute abundance. Interestingly, we also observe an increase in the percentage of integration sites in a gene amongst integration sites from individuals with strong predicted HBZ binding capacity compared those less likely to bind HBZ epitopes, particularly in the less abundant clones. Increasing numbers of HLA class I alleles able to present HBZ are associated with decreasing PVL, suggesting that there is greater control of infected cells in these patients [21]. Integration sites located in genes and active genomic regions are associated with increased clone abundance ([24], and this study): we postulate that these environments support virus-driven cell proliferation allowing clones to survive under stronger immune control.

Tax is known to drive proliferation of the infected cell; could the clones which selectively survive in ACs express more Tax? We previously observed, in a very small sample of integration sites (n = 40, [26]), that Tax-expressing cells had a higher frequency of integration sites in genes. However, with the advent of high-throughput sequencing, we have recently shown that clones that spontaneously express Tax ex vivo have a minor increase in the frequency of integration in a gene and in regions with activating epigenetic marks ([23] and AM, unpublished observations). Since we have previously observed that more abundant clones are less likely to express Tax [23], Tax expression is unlikely to completely account for the success of these integration site clones. HBZ is also known to drive cell proliferation [4], and integration in a gene or active region may also promote increased expression of HBZ. To definitively determine, however, whether transcriptionally active regions promote increased HBZ expression will require high-throughput sorting and integration site analysis of HBZ expressing clones directly ex vivo: this is currently precluded by the difficulty in sorting cells based on

detection of HBZ protein in naturally-infected PBMCs. The role of the integration site in driving proliferation via either Tax or HBZ expression is also complicated by the effect of Tax and HBZ on the expression or function of each other.

A recent study in primary infection with BLV, a related retrovirus [27], has shown that early in infection, integration is favoured in transcriptionally active areas but is strongly selected against by the host immune response. Yet in subsequent chronic infection, abundant clones have a higher frequency of integration sites in transcriptionally active areas. Similarly, in HTLV-1, the effectiveness of the initial host response against expressed viral proteins is likely to define PVL set point, selecting against highly active clones. Clones in heterochomatin may also represent a dead end for the virus, because it may never be re-expressed. During lifelong chronic infection, however, surviving clones with integration sites in 'intermediate' transcriptionally active areas may have a proliferative advantage, although other factors (including clone TCR specificity or immune escape by Tax silencing or timing of viral expression [16,23,28]) will also contribute to the relative success of a clone. These transcriptionally active clones are more common in ACs than HAM/TSP, plausibly because they compete against the effective host response in ACs, or less likely, because they are selectively lost (by an unknown mechanism) in HAM/TSP.

There are other differences between HAM/TSP patients and AC individuals, in addition to the effectiveness of the T-cell immune response, which could alter the selection of proviral integration site during chronic infection. For example, in HAM/TSP, proliferation of HTLV-1-infected T cells may be maintained by IL-2 [29] and IL-15 [30], which may reduce the advantage conferred by integration sites that increase expression of proliferation-inducing factors such as HBZ.

Our results reflect systematic differences in the characteristics of HTLV-1 integration sites that persist in vivo between HAM/TSP patients and ACs. We propose that these differences are not themselves causative of the disparate clinical outcomes, but rather they reflect an underlying difference between patients with HAM/TSP and ACs in the efficiency of host-mediated

control of HTLV-1 replication, for which there is extensive evidence [8].

Two previous studies compared the integration site environment between ACs and patients with HAM/TSP and found no differences, or a borderline significant (p = 0.049) difference [24,26]. The difference between these reports and the present study may be attributable to the differences in sample size (N = 238 in present study, cp. N = 40 and 24 respectively in [24,26]), the quantitative nature and greater sensitivity of the present high-throughput method, or in the ethnicity of the study population (uniformly southern Japanese vs. predominantly Caribbean). The incidence of HAM/TSP is much lower in the Japanese population (studied here) than in individuals of Caribbean origin in the previous studies (0.25 vs 3% [31,32]). In the current study, we excluded differences in LM-PCR efficiency and mean clone abundance as causes of the observed differences between patients with HAM/TSP and ACs.

Conclusions

The expression levels of Tax and HBZ influence both the rate of proliferation and the rate of CTL killing of each infected T-cell clone, and thus its equilibrium abundance in vivo. The balance between the relative strength of these opposing effects differ between the two genes: *tax* encodes the dominant target of the CTL response and is often silenced in abundant clones despite its capacity to support proliferation, but HBZ expression is typically maintained and is a poorer target for the CTL response. We have previously reported features of the integration site that influence Tax expression but there are no studies directly linking integration site and HBZ expression. We show that integration sites within genes and active genomic regions are more frequent in AC clones compared to equivalently-sized clones from HAM/TSP patients, and in individuals who have HLA class I allelesable to effectively present HBZ peptides to CTL. We postulate that integration in a transcriptionally active area may elevate HBZ and/or Tax expression and increase the equilibrium abundance of a clone. The increased frequency of integration sites observed in transcriptionally active genomic regions in ACs is consistent with the conclusion that greater proliferation is required to reach a given clonal abundance under the selection pressure exerted by an effective anti-HTLV-1 immune response.

Methods
Subjects
Kagoshima cohort. The study population has been previously reported [7,10]. Subjects consisted of 229 patients with HAM/TSP and 202 HTLV-1-infected asymptomatic carriers randomly selected from blood donors; all were of Japanese ethnic origin and residing in Kagoshima, Kyushu, Japan [7,10].

Kumamoto cohort. The study population consisted of 98 HTLV-1-infected asymptomatic carriers from blood donors in Kumamoto Prefecture, Kyushu, Japan.

Research was carried out in compliance with the Helsinki Declaration. The study was approved by the Faculty of Medical and Pharmaceutical Sciences Ethics Review Board, University Hospital, Kumamoto, Japan (Ethics 149). All patients gave written, informed consent for the study and for publication of anonymized results.

HTLV-1 proviral load measurements
HTLV-1 DNA was amplified by quantitative PCR in a ABI7900HT FAST real time PCR system using FastSYBR Green (Applied Biosystems) reagents with the Tax-specific primers SK43 and SK44. A control region in B-actin was also amplified using ActF and ActR primers. The rat cell line TARL-2, which contains one integrated copy of the HTLV-1 provirus, was used to generate a standard curve. The sample copy number was interpolated from the standard curve and PVL was expressed as number of infected cells per 100 PBMCs. Proviral load data for the Kagoshima HAM/TSP and AC cohorts, with TARL-2 as pX region control, were as previously described ([7]). SK43: 5′CGGATACCCAGTCTACGTGT, SK44: 5′GAGCCGA TAACGCGTCCATCG, ActF: 5′TCACCCACACTGTGCC CATCTATGA, ActR: 5′ CATCGGAACCGCTACTTGCC GATAG.

HLA class I alleles
HLA class I typing of the Kagoshima cohort was reported previously ([10]). HLA typing of the Kumamoto cohort was done by Luminex reverse SSOP at the Hammersmith Hospital, London, UK to 2 digit resolution with 'strings' of possible 4 digit resolution alleles. For each individual's string of possible alleles, the most frequent 4-digit allele in the Japanese population (represented by a study of 1018 Japanese individuals [33]) was identified as the most likely allele. If there were multiple possible alleles with a population frequency >3%, all these frequent alleles were retained as possibilities for the individual. If no allele subtype in the string was represented in the population study, all alleles in the string were retained.

Epitope binding prediction
We used the Metaserver algorithm (detailed in [21]) to predict HLA class I epitopes. Metaserver combines predicted TAP transport, proteasomal cleavage and HLA–peptide binding from NetCTL and NetMHC to predict peptide binding to HLA Class I A and B alleles present in the Kagoshima and Kumamoto cohorts. For each HLA class I A or B allele, all HTLV-1 peptides were ranked by binding score. The rank of the top HBZ peptide was

recorded for each HLA class I allele. The rank of top binding peptide for a protein is a more robust method for comparisons between alleles than affinity [21]. HLA class alleles where the rank of the top HBZ peptide was < 20 were included as predicted binders of HBZ. A0201 was also included as a binder (HBZ top rank = 29) as it had been experimentally shown to present HBZ peptides [34]. For alleles not included in Metaserver, we assessed the rank of the binding score using NetMHC only. Alleles where the rank of the top HBZ peptide was < 20 were added to the predicted binders of HBZ. Approximately 30% of individuals have a detectable CTL response to HBZ [35]. Individuals with two or more HBZ-binding alleles constituted 39% of the combined cohorts, and were designated 'strong HBZ binders' representing individuals who were most likely to be able to mount an HBZ-specific response.

High-throughput linker-mediated PCR identification of proviral integration sites

Integration site analysis was carried out on samples from patients with HAM/TSP and ACs from the Kagoshima cohort where sufficient genomic DNA was available (HAM/TSP, n = 104; AC, n = 100). To maximise statistical power and to test for reproducibility, we also analysed DNA of ACs from a neighbouring prefecture in Japan, Kumamoto (n = 98).

HTLV-1 integration sites were mapped and their abundance quantified as previously described ([23,24]). Genomic DNA from peripheral blood mononuclear cells was randomly sheared by sonication and ligated to a partially double-stranded DNA adaptor that incorporated a 6 nt barcode, a reverse sequencing primer site and the P7 sequence for paired-end sequencing. Two rounds of nested PCR were performed between the HTLV-1 LTR and the adaptor, adding a paired end P5 sequence in the LTR primer. The resulting amplicons were combined into libraries of up to 42 samples and sequence data were acquired on an Illumina GAII or HiSeq platform with paired-end 50 bp reads and a 6 bp index (barcode) read. Paired reads were aligned to a human genome reference (build 18, excluding haplotype and "random" chromosomes) using ELAND. A random set of integration sites was derived from ∼ 190,000 50 bp human genome sequences generated using Galaxy, and aligned to the same human genome reference to control for any bias due to alignment limitations.

Unique integration sites (defined by Read1) were quantified on the basis of number of distinct shear sites identified (determined from paired Read2) and calibrated to provide a count of number of sequenced sister cells per clone. The absolute abundance (*Aabs*) of a T cell clonal population, defined by a single HTLV-1 integration site, was calculated as follows:

$$Aabs(i) = \frac{Si}{\sum_{i=1}^{T} Si} * 100 * PVL$$

where *Si* is the number of sister cells of the *i*th clone, T is the total number of observed clones in the patient, and *PVL* is the number of infected cells per 100 PBMC in the patient. Clones were assigned to absolute abundance ranges of <1 per 10^4 PBMC, 1-10 per 10^4 PBMC, >10 per 10^4 PBMC. LM-PCR integration site data sets for each patient were subjected to successive quality control filters as previously described ([23,24]); LM-PCR was designated 'successful' and the data included in further analysis only when the sample contained a minimum of 50 sister cells, a minimum of 15 clones, an average of more than 15 sequence reads per sister cell.

Oligoclonality index

To measure the diversity of clone abundance in the infected cell population from each individual, we used the oligoclonality index [24] which is based on the Gini Index. This index measures the non-uniformity of the distribution of clone abundance: a value of 0 indicates that all clones have the same abundance and 1 is an upper bound where the proviral load effectively consists of a single clone.

Estimating total number of clones in the blood

Estimation of total numbers of clones in the blood (observed and unobserved) was carried out using a computational diversity estimation approach (DivE), as described previously [25]. Briefly, many mathematical models were fitted to species-accumulation data, and to successively smaller nested subsamples thereof. Novel criteria were used to score models in how consistently they can reproduce existing observations from incomplete data. The estimates from the best performing models were aggregated (using the geometric mean) to estimate the number of clones in the circulation.

Genetic and epigenetic analysis of integration site

Analysis of the genomic region surrounding the integration site was carried out as previously described ([23,24]). Specifically, the following attributes were identified for each integration site: location within/outside a transcriptional unit and orientation of the provirus versus that gene, proximity to a CpG island, and counts of selected histone marks within a 10 Kb window around the integration site. Locations of transcriptional units were retrieved from the NCBI (ftp.ncbi.nih.gov/gene/), CpG island data from UCSC tables [36], and histone marks from ChIP-seq experiments on primary CD4+ T cells (detailed in [24]).

integration site positions were compared to the locations of specific relevant annotations using the R package hiAnnotator (http://malnirav.github.com/hiAnnotator), kindly provided by N. Malani and F. Bushman (University of Pennsylvania, USA). An integration site was designated as being in an area enriched in a particular histone mark if there were more counts of the mark within a window of 10 Kb around the integration site than around 90% of random sites.

Statistical analysis

Statistical tests were performed using R version 2.15 (http://www.R-project.org/). A multivariable linear regression model was used to compare log PVL with the number of HBZ-binding HLA alleles per individual, taking into account disease status, as well as age and sex in the Kagoshima cohort-only analysis as these have been suggested to vary with PVL. Differences in OCI, PVL and total number of sisters between cohorts were analysed by non-parametric Mann-Whitney U tests as distributions were non-normal. Spearman correlation was used to determine the correlation in each cohort between log PVL and either log estimated number of clones in the blood or OCI. A likelihood ratio test was used to compare a null multivariable linear regression model associating log clone number with log PVL, disease status and HBZ binding capacity to an alternative model which added an interaction term between PVL and HBZ binding capacity. This allowed a test of whether the association between total clones in the blood and PVL differed with HBZ binding capacity. A second likelihood ratio test compared the same null model with a different alternative model including an interaction term between log PVL and disease status. This method was repeated for the association of OCI with log PVL.

In the genomic environment analysis, integration sites were grouped by absolute abundance range, and by the prefecture, disease status and HBZ binding status of the patient. For integration site data statistical analysis, the two AC cohorts were combined into a single AC group as they showed very similar integration site characteristics. Within subsets, the proportion of integration sites located within genes was plotted as odds ratio compared to random integration sites. Chi-squared tests were used to compare the total numbers of UIS lying inside or outside genes at each absolute abundance bin level. A chi-squared test for trend was used to measure the significance of a trend within a cohort across bins of increasing abundance. We plotted the mean number of a specified epigenetic mark within a 10 Kb window around integration sites within a group (N) divided by the mean number of that mark within 10 Kb of in silico random sites (N random). Mann-Whitney U tests were used to compare the numbers of specified epigenetic marks near integration sites

from ACs versus HAM/TSP patients. Spearman correlation was used to test the association between log absolute abundance of an integration site clone and the frequency of a specified epigenetic mark within 10 Kb of the integration site.

Correction for multiple comparisons was made using a Bonferroni-Holm correction to control the family-wise error rate for each set of tests (Mann-Whitney, Spearman) carried out across all analysed epigenetic marks and integration site subsets. Two epigenetic marks (H3K9ac, H3K4ac) were analysed but not reported as their results were very similar to reported ones; the p values from their analyses were included in the calculation of the Bonferroni-Holm correction. Correction for multiple comparisons was also made across subsets in the analyses (Chi-square, Chi square test for trend) of integration sites in genes.

A multivariable logistic regression model was used to test whether disease status was independently associated with integration site in a gene, active genomic region and inhibitory genomic region at the integration site level (all integration sites used in analysis) or patient level (characteristics of integration sites averaged per patient). Host HBZ binding status and PVL, as well as log absolute clone abundance were controlled as additional factors in the model as these may vary with disease status and integration site environment. A multivariable logistic regression model was also used to test the association of integration in the same orientation as a gene or TSS with disease status (controlling for HBZ binding status, PVL and clone abundance). A further multivariable logistic regression tested the association of host HBZ binding status with integration in a gene (controlling for disease status and clone abundance) at the patient and integration site level.

Additional file

Abbreviations
AC: Asymptomatic carrier; ATLL: adult T cell leukaemia/lymphoma; CTL: Cytotoxic lymphocyte; HAM/TSP: HTLV-1-associated myelopathy/tropical spastic paraparesis; HBZ: HTLV-1 basic leucine zipper factor; HTLV-1: Human T cell lymphotropic virus-type 1; OCI: Oligoclonality index; LM-PCR: Linker mediated polymerase chain reaction; PBMC: Peripheral blood mononuclear cell; PVL: Proviral load.

Competing interests
The authors declare that they have no competing interests.

Author details

[1]Department of Immunology, Wright-Fleming Institute, Imperial College London, London W2 1PG, UK. [2]Institute for Viral Research, Kyoto University, Kyoto 606-8507, Japan.

References

1. Gessain A, Cassar O: **Epidemiological Aspects and World Distribution of HTLV-1 Infection.** *Front Microbiol* 2012, 3:388.
2. Cook LB, Rowan AG, Melamed A, Taylor GP, Bangham CRM: **HTLV-1-infected T cells contain a single integrated provirus in natural infection.** *Blood* 2012, 120:3488–3490.
3. Arnold J, Zimmerman B, Li M, Lairmore MD, Green PL: **Human T-cell leukemia virus type-1 antisense-encoded gene, Hbz, promotes T-lymphocyte proliferation.** *Blood* 2008, 112:3788–3797.
4. Satou Y, Yasunaga J-i, Yoshida M, Matsuoka M: **HTLV-I basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells.** *Proc Natl Acad Sci* 2006, 103:720–725.
5. Boxus M, Willems L: **Mechanisms of HTLV-1 persistence and transformation.** *Br J Cancer* 2009, 101:1497–1501.
6. Yamano Y, Sato T: **Clinical pathophysiology of human T-lymphotropic virus-type 1-associated myelopathy/tropical spastic paraparesis.** *Front Microbiol* 2012, 3:389.
7. Nagai M, Usuku K, Matsumoto W, Kodama D, Takenouchi N: **Analysis of HTLV-I proviral load in 202 HAM/TSP patients and 243 asymptomatic HTLV-I carriers: high proviral load strongly predisposes to HAM/TSP.** *J Neurovirol* 1998, 4:586–593.
8. Bangham CRM: **CTL quality and the control of human retroviral infections.** *Eur J Immunol* 2009, 39:1700–1712.
9. Journo C, Mahieux R: **HTLV-1 and innate immunity.** *Viruses* 2011, 3:1374–1394.
10. Jeffery KJ, Usuku K, Hall SE, Matsumoto W, Taylor GP, Procter J, Bunce M, Ogg GS, Welsh KI, Weber JN, Lloyd AL, Nowak MA, Nagai M, Kodama D, Izumo S, Osame M, Bangham CR: **HLA alleles determine human T-lymphotropic virus-I (HTLV-I) proviral load and the risk of HTLV-I-associated myelopathy.** *Proc Natl Acad Sci U S A* 1999, 96:3848–3853.
11. Goon PKC, Biancardi A, Fast N, Igakura T, Hanon E, Mosley AJ, Asquith B, Gould KG, Marshall S, Taylor GP, Bangham CRM: **Human T cell lymphotropic virus (HTLV) type-1-specific CD8+ T cells: frequency and immunodominance hierarchy.** *J Infect Diseases* 2004, 189:2294–2298.
12. Kannagi M, Harada S, Maruyama I, Inoko H, Igarashi H, Kuwashima G, Sato S, Morita M, Kidokoro M, Sugimoto M: **Predominant recognition of human T cell leukemia virus type I (HTLV-I) pX gene products by human CD8+ cytotoxic T cells directed against HTLV-I-infected cells.** *Int Immunol* 1991, 3:761–767.
13. Niewiesk S, Daenke S, Parker CE, Taylor G, Weber J, Nightingale S, Bangham CR: **Naturally occurring variants of human T-cell leukemia virus type I Tax protein impair its recognition by cytotoxic T lymphocytes and the transactivation function of Tax.** *J Virol* 1995, 69:2649–2653.
14. Koiwa T, Hamano-usami A, Ishida T, Okayama A, Yamaguchi K, Kamihira S, Koiwa T, Hamano-usami A, Ishida T, Okayama A, Yamaguchi K, Kamihira S, Watanabe T: **5′-Long Terminal Repeat-Selective CpG Methylation of Latent Human T-Cell Leukemia Virus Type 1 Provirus In Vitro and In Vivo.** *J Virol* 2002, 76:9389–9397.
15. Tamiya S, Matsuoka M, Etoh K, Watanabe T, Kamihira S, Yamaguchi K, Takatsuki K: **Two types of defective human T-lymphotropic virus type I provirus in adult T-cell leukemia.** *Blood* 1996, 88:3065–3073.
16. Cook LB, Melamed A, Niederer H, Valganon M, Laydon D, Foroni L, Taylor GP, Matsuoka M, Bangham CRM: **The role of HTLV-1 clonality, proviral structure and genomic integration site in adult T cell leukemia/lymphoma.** *Blood* 2014, doi:10.1182/blood-2014-02-553602.
17. Furukawa Y, Kubota R, Tara M, Izumo S, Osame M: **Existence of escape mutant in HTLV-I tax during the development of adult T-cell leukemia.** *Blood* 2001, 97:987–993.
18. Asquith B, Mosley AJ, Barfield A, Marshall SEF, Heaps A, Goon P, Hanon E, Tanaka Y, Taylor GP, Bangham CRM: **A functional CD8+ cell assay reveals individual variation in CD8+ cell antiviral efficacy and explains differences in human T-lymphotropic virus type 1 proviral load.** *The Journal of General Virology* 2005, 86:1515–1523.
19. Sabouri AH, Usuku K, Hayashi D, Izumo S, Ohara Y, Osame M, Saito M: **Impaired function of human T-lymphotropic virus type 1 (HTLV-1)-specific CD8+ T cells in HTLV-1-associated neurologic disease.** *Blood* 2008, 112:2411–2420.
20. Saito M, Matsuzaki T, Satou Y, Yasunaga J-I, Saito K, Arimura K, Matsuoka M, Ohara Y: **In vivo expression of the HBZ gene of HTLV-1 correlates with proviral load, inflammatory markers and disease severity in HTLV-1 associated myelopathy/tropical spastic paraparesis (HAM/TSP).** *Retrovirology* 2009, 6:19.
21. Macnamara A, Rowan A, Hilburn S, Kadolsky U, Fujiwara H, Suemori K, Yasukawa M, Taylor G, Bangham CRM, Asquith B: **HLA class I binding of HBZ determines outcome in HTLV-1 infection.** *PLoS Pathog* 2010, 6:e1001117.
22. Van Dooren S, Pybus OG, Salemi M, Liu H-F, Goubau P, Remondegui C, Talarmin A, Gotuzzo E, Alcantara LCJ, Galvão-Castro B, Vandamme A-M: **The low evolutionary rate of human T-cell lymphotropic virus type-1 confirmed by analysis of vertical transmission chains.** *Mol Biol Evol* 2004, 21:603–611.
23. Melamed A, Laydon DJ, Gillet NA, Tanaka Y, Taylor GP, Bangham CRM: **Genome-wide determinants of proviral targeting, clonal abundance and expression in natural HTLV-1 infection.** *PLoS Pathog* 2013, 9:e1003271.
24. Gillet NA, Malani N, Melamed A, Gormley N, Carter R, Bentley D, Berry C, Bushman FD, Taylor GP, Bangham CRM: **The host genomic environment of the provirus determines the abundance of HTLV-1-infected T cell clones.** *Blood* 2011, 117:3113–3122.
25. Laydon DJ, Melamed A, Sim A, Gillet NA, Sim K, Darko S, Kroll JS, Douek DC, Price DA, Bangham CRM, Asquith B: **Quantification of HTLV-1 clonality and TCR diversity.** *PLoS Comput Biol* 2014, 10:e1003646.
26. Meekings KN, Leipzig J, Bushman FD, Taylor GP, Bangham CRM: **HTLV-1 integration into transcriptionally active genomic regions is associated with proviral expression and with HAM/TSP.** *PLoS Pathog* 2008, 4:e1000027.
27. Gillet NA, Gutiérrez G, Rodriguez SM, de Brogniez A, Renotte N, Alvarez I, Trono K, Willems L: **Massive depletion of bovine leukemia virus proviral clones located in genomic transcriptionally active sites during primary infection.** *PLoS Pathog* 2013, 9:e1003687.
28. Gillet NA, Cook L, Laydon DJ, Hlela C, Verdonck K, Alvarez C, Gotuzzo E, Clark D, Farre L, Bittencourt A, Asquith B, Taylor GP, Bangham CRM: **Strongyloidiasis and infective dermatitis alter human T lymphotropic virus-1 clonality in vivo.** *PLoS Pathog* 2013, 9:e1003263.
29. Tendler CL, Greenberg SJ, Blattner W, Manns A, Murphy E, Fleisher T, Hanchard B, Morgan O, Burton JD, Nelson DL: **Transactivation of interleukin 2 and its receptor induces immune activation in human T-cell lymphotropic virus type I-associated myelopathy: pathogenic implications and a rationale for immunotherapy.** *Proc Natl Acad Sci U S A* 1990, 87:5218–5222.
30. Azimi N, Mariner J, Jacobson S, Waldmann TA: **How does interleukin 15 contribute to the pathogenesis of HTLV type 1-associated myelopathy/tropical spastic paraparesis?** *AIDS Res Hum Retroviruses* 2000, 16:1717–1722.
31. Kaplan JEOM, Kubota H, Igata A, Nishitani H, Maeda Y, Khabbaz RF, Janssen RS: **The risk of development of HTLV-I-associated myelopathy/tropical**

spastic paraparesis among persons infected with HTLV-I. *J Acquir Immune Defic Syndr* 1990, **3**:1096–1101.

32. Tosswill JHC, Taylor GP, Tedder RS, Mortimer PP, Voss LD, Mulligan J: HTLV-I/II associated disease in England and Wales, 1993–7: retrospective review of serology requests. *Br Med J* 2000, **320**:611–612.

33. Itoh Y, Mizuki N, Shimada T, Azuma F, Itakura M, Kashiwase K, Kikkawa E, Kulski JK, Satake M, Inoko H: High-throughput DNA typing of HLA-A, -B, -C, and -DRB1 loci by a PCR-SSOP-Luminex method in the Japanese population. *Immunogenetics* 2005, **57**:717–729.

34. Suemori K, Fujiwara H, Ochi T, Ogawa T, Matsuoka M, Matsumoto T, Mesnard J-M, Yasukawa M: HBZ is an immunogenic protein, but not a target antigen for human T-cell leukemia virus type 1-specific cytotoxic T lymphocytes. *J Gen Virol* 2009, **90**:1806–1811.

35. Hilburn S, Rowan A, Demontis M-A, MacNamara A, Asquith B, Bangham CRM, Taylor GP: In vivo expression of human T-lymphotropic virus type 1 basic leucine-zipper protein generates specific CD8+ and CD4+ T-lymphocyte responses that correlate with clinical outcome. *J Infect Dis* 2011, **203**:529–536.

36. Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, Kent WJ: The UCSC Table Browser data retrieval tool. *Nucleic Acids Res* 2004, **32**:D493–D496.

A Critical Role for IL-17RB Signaling in HTLV-1 Tax-Induced NF-κB Activation and T-Cell Transformation

CrossMark
click for update

Alfonso Lavorgna[1], Masao Matsuoka[2], Edward William Harhaj[1]*

1 Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins School of Medicine, Baltimore, Maryland, United States of America, 2 Laboratory of Virus Control, Institute for Virus Research, Kyoto University, Kyoto, Japan

Abstract

Human T-cell leukemia virus type 1 (HTLV-1) infection is linked to the development of adult T-cell leukemia (ATL) and the neuroinflammatory disease HTLV-1 associated myelopathy/tropical spastic paraparesis (HAM/TSP). The HTLV-1 Tax protein functions as a potent viral oncogene that constitutively activates the NF-κB transcription factor to transform T cells; however, the underlying mechanisms remain obscure. Here, using next-generation RNA sequencing we identified the IL-25 receptor subunit IL-17RB as an aberrantly overexpressed gene in HTLV-1 immortalized T cells. Tax induced the expression of IL-17RB in an IκB kinase (IKK) and NF-κB-dependent manner. Remarkably, Tax activation of the canonical NF-κB pathway in T cells was critically dependent on IL-17RB expression. IL-17RB and IL-25 were required for HTLV-1-induced immortalization of primary T cells, and the constitutive NF-κB activation and survival of HTLV-1 transformed T cells. IL-9 was identified as an important downstream target gene of the IL-17RB pathway that drives the proliferation of HTLV-1 transformed cells. Furthermore, IL-17RB was overexpressed in leukemic cells from a subset of ATL patients and also regulated NF-κB activation in some, but not all, Tax-negative ATL cell lines. Together, our results support a model whereby Tax instigates an IL-17RB-NF-κB feed-forward autocrine loop that is obligatory for HTLV-1 leukemogenesis.

Introduction

The retrovirus human T-cell leukemia virus type 1 (HTLV-1) infects between 10–20 million people worldwide [1]. HTLV-1 is the etiological agent of the neuroinflammatory disease HTLV-1-associated myelopathy (HAM/TSP) and adult T-cell leukemia (ATL), a CD4+CD25+ T-cell malignancy [2,3]. ATL develops in about 5% of HTLV-1-infected individuals after a long latent period spanning 40–60 years [4]. The HTLV-1 genome encodes the Tax protein that exerts pleiotropic roles and is an essential regulator of viral replication and oncogenic cell transformation [5]. Tax modulates the activation of several key signaling pathways and cell cycle proteins to enhance T-cell proliferation and survival. One of the key cellular targets important for transformation by Tax is the NF-κB transcription factor [6].

NF-κB is composed of heterodimeric DNA binding proteins consisting of RelA, c-Rel, RelB, p50 and p52 [7]. In the canonical NF-κB pathway, NF-κB heterodimers are sequestered in the cytoplasm by ankyrin-repeat containing inhibitory proteins including IκBα [8]. A wide variety of stimuli including stress signals, proinflammatory cytokines or virus infection activate the IKK kinase complex consisting of the catalytic subunits IKKα and IKKβ and the regulatory subunit IKKγ (also known as NEMO) [9]. IKKβ phosphorylates IκB proteins to trigger their ubiquitin-dependent degradation thus allowing NF-κB to enter the nucleus

and activate target genes [10]. In the noncanonical NF-κB pathway, tumor necrosis factor receptor (TNFR) superfamily members including BAFF, lymphotoxin-β and CD40 promote proteasomal processing of the p100 (NF-κB2) precursor protein to yield p52, which forms transcriptionally active heterodimers with RelB. The NF-κB inducing kinase (NIK) is a key regulator of this pathway by activating IKKα homodimers which in turn phosphorylate p100 leading to its processing. Tax constitutively activates both the canonical and noncanonical NF-κB pathways, in part by interacting directly with NEMO and IKK [11–14]. There is evidence that Tax may require upstream signaling molecules such as the kinase TAK1 to activate canonical NF-κB signaling [15]. Although the proximal signaling components of TNFR and interleukin-1 receptor (IL-1R) are dispensable for Tax to activate NF-κB [16], whether Tax has usurped a distinct NF-κB pathway is unknown.

Tax activation of the canonical and noncanonical NF-κB pathways fosters the aberrant expression of anti-apoptotic and pro-proliferative genes that leads to oncogenesis. Tax mutants defective in NF-κB activation expressed in an infectious HTLV-1 molecular clone are impaired in the immortalization of primary T cells [17]. NF-κB is also required for the survival of HTLV-1 transformed cell lines and patient-derived ATL cells [18]. Therefore, HTLV-1 transformed cell lines and primary ATL leukemic specimens exhibit a strict "addiction" to NF-κB for

Author Summary

The retrovirus HTLV-1 is the causative agent of an aggressive lymphoproliferative disorder known as adult T-cell leukemia (ATL). The HTLV-1 Tax regulatory protein constitutively activates the host NF-κB transcription factor to promote T-cell proliferation, survival and cell transformation. However, it remains unknown precisely how Tax persistently activates NF-κB in T cells. In this study, we used next-generation sequencing to identify genes that were differentially expressed upon HTLV-1 infection and immortalization of primary T cells. We found that IL-17RB, the receptor for the IL-25 cytokine, was highly induced in HTLV-1 transformed T cells and was required for NF-κB activation, cell proliferation and survival. Tax induced the expression of IL-17RB and established a positive feedback loop together with IL-25 that triggered persistent NF-κB activation and the upregulation of IL-9 and other genes critical for T-cell proliferation and survival. IL-17RB was also overexpressed in a subset of acute ATL patient specimens and therefore may potentially be targeted by monoclonal antibodies as a novel ATL therapy.

survival and proliferation, thus establishing the NF-κB pathway as an attractive target for novel ATL therapeutics. However, since Tax expressing cells are vigorously targeted by cytotoxic T cells and other arms of the host immune response, the majority of ATL tumors exhibit downregulated or lost Tax expression by mutations within Tax or deletion or methylation of the 5′ viral long terminal repeat region (LTR) [19]. Thus, Tax likely plays more important roles in the early events of transformation via persistent NF-κB activation, inactivation of p53 and other tumor suppressors and induction of genomic instability and aneuploidy [5]. However, canonical and noncanonical NF-κB signaling remains constitutive in ATL despite the loss of Tax.

The interleukin 17 (IL-17) cytokine family consists of six members (IL-17A-F) that play essential roles in host immunity and inflammatory diseases. IL-17A is the signature IL-17 cytokine and binds to an IL-17RA/IL-17RC receptor complex to orchestrate the host response against bacterial and fungal infections [20]. IL-17A controls the expression of cytokines and chemokines that enhance neutrophil recruitment. Dysregulation of this pathway has been implicated in numerous autoimmune and metabolic diseases and cancer [21]. IL-17E (also known as IL-25) is essential for host defense against parasites by regulating expression of T helper 2 (Th2) cytokines including IL-4, IL-5 and IL-13 that promote eosinophil recruitment [22]. IL-25 has also been linked to allergic airway inflammation and asthma [23]. IL-25 is produced by diverse cell types such as epithelial cells, T cells, eosinophils, mast cells and basophils [24,25]. IL-25 binds to a heterodimeric receptor composed of IL-17RA and IL-17RB, of which IL-17RB is the specific receptor subunit for IL-25 [26]. IL-17RB is highly expressed in kidney, liver and other peripheral organs as well as memory and effector T lymphocytes [27]. IL-17RB expression can be regulated by IL-4 and TGF-β, however the precise transcriptional regulatory control of IL-17RB is unknown. Upon binding to IL-25, IL-17RB recruits the Act1 (also known as CIKS) adaptor molecule via homotypic SEFIR (similar expression to fibroblast growth factor genes and IL-17R) domain interactions [28,29]. Act1 activates the ubiquitin ligase TRAF6 and the kinase TAK1 that in turn triggers NF-κB and MAP kinase activation to induce type 2 cytokines IL-4, IL-5 and IL-13 as well as IL-9 [30,31]. IL-17B also serves as a ligand for IL-17RB, albeit with a lower affinity for IL-17RB compared to IL-25 [32]. In addition to IL-17RB

regulation of host defense and allergic airway disease, this pathway can be oncogenic if dysregulated. The *IL-17RB* locus is a common site of retroviral integration in murine myeloid leukemias, resulting in the upregulation of IL-17RB expression [33]. IL-17RB is also overexpressed in a subset of breast tumors and is associated with poor prognosis [34]. In breast cancer, IL-17RB engagement by IL-17B triggers TRAF6 recruitment to IL-17RB, NF-κB activation and induction of the *bcl-2* gene to inhibit apoptosis [34].

Although considerable progress has been made in our understanding of HTLV-1 oncogenesis, the precise mechanisms underlying HTLV-1-induced transformation remain unclear. Previous microarray studies have identified several anti-apoptotic, cell cycle and growth regulatory genes dysregulated by HTLV-1 [35–37]. However, due to experimental limitations of these studies and the advent of next-generation sequencing, RNA sequencing (RNA-Seq) has emerged as a powerful tool to evaluate gene expression, differential splicing, noncoding RNAs, RNA editing and gene fusions [38]. In this study, we used RNA-Seq to delineate the transcriptome of primary T lymphocytes immortalized by HTLV-1. This work led to the identification of IL-17RB as an aberrantly overexpressed gene in HTLV-1 transformed cells that was induced by the HTLV-1 Tax protein. Surprisingly, the IL-17RB pathway was required for constitutive NF-κB activation by Tax and in HTLV-1 transformed cell lines. Furthermore, IL-17RB was overexpressed in leukemic cells from acute ATL patients and was essential for NF-κB activation in a subset of Tax-negative ATL cell lines.

Results

Next-generation sequencing identifies the transcriptomes of primary T cells infected and immortalized by HTLV-1

To gain insight into the mechanisms of HTLV-1-induced T-cell immortalization, we used a well-established co-culture model [35,39] whereby primary human CD4+ T cells were purified by immunomagnetic beads from normal donor peripheral blood mononuclear cells (PBMCs) and co-cultured with lethally irradiated HTLV-1 transformed MT-2 cells (to provide a source of HTLV-1). Primary T cells were consistently immortalized in the presence of MT-2 cells between 6–8 weeks. Control T cells cultured in the absence of MT-2 did not proliferate after 4 weeks and were no longer viable at that time. The co-culture assay was performed with T cells from 4 independent blood donors. Of the 4 co-cultures, all produced immortalized T cell clones, however clone #1 ceased proliferation unexpectedly and was excluded from further studies. The immortalized T cell clones (T-MT-2) #2-4 remained dependent on recombinant IL-2 for proliferation and expressed CD3, CD4 and CD25 cell surface markers (Figure 1A).

Total RNA was harvested from T-MT-2 clone #2 (week 12 after co-culture) for RNA-Seq analysis as well as parental primary T cells (week 0), and T cells after 1 week of co-culture. A pure population of viable cells was obtained from the co-culture after removal of dead cells using magnetic labeling and separation. MT-2 RNA was also included as a control for RNA-Seq to confirm that the immortalized T cells expressed a unique genetic signature and were not simply MT-2 contaminants. RNA-Seq and bioinformatics analysis were performed with a total number of reads of 65 million (week 0), 73 million (week 1), 44 million (week 12) and 52 million (MT-2). At 1 week after co-culture, the most abundant induced coding RNAs in the T cells were interferon-stimulated genes (ISGs) such as ISG15, IFI27, OAS1 and MX1