

Circulating CD4⁺ T-Cell Differentiation Subsets Are Redistributed in HTLV-1 –Associated Adult T-Cell Leukemia/Lymphoma

Jéssica Bilar Cavalcanti¹

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Priscila Lima Barros¹

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Davi Ferreira Jofre¹

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Giulia Sales Guimarães¹

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Karine Sobral Marques¹

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Ana Beatriz Rocha¹

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Vinícius de Camargo Callefi

Faculdade de Medicina da Universidade de São Paulo

Lucas Ruffo

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Gabriel Palhares

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Julia Augusto Passos

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Nélio César de Aquino

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Sandro Albuquerque

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Vivian Furlan

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Sofia Cattena

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Lilian de Souza França

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Letícia Montanha de Assis

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Sheila de Oliveira Garcia Mateos

Universidade Municipal de São Caetano do Sul

Carolina Rosadas

Imperial College London

Jorge Simão do Rosário Casseb

Faculdade de Medicina da Universidade de São Paulo

Vanderson Rocha

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Youko Nukui

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Luís Alberto de Pádua Covas Lage

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Juliana Pereira

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Cadiele Oliana Reichert

Faculdade de Medicina da Universidade de São Paulo

Hebert Fabricio Culler¹

hebert.culler@hc.fm.usp.br

Hospital das Clínicas HCFMUSP, Universidade de São Paulo

Research Article

Keywords: Adult T-cell leukemia/lymphoma, CD4-positive T lymphocytes, effector memory T cells, flow cytometry, HTLV-1, T-cell differentiation, T-cell memory subsets, transmission route

Posted Date: August 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10413678/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Virology Journal on October 15th, 2026. See the published version at <https://doi.org/10.1186/s12985-026-03323-x>.

Abstract

Background

Human T-cell lymphotropic virus type 1 (HTLV-1) establishes lifelong infection in CD4 + T lymphocytes and can cause adult T-cell leukemia/lymphoma (ATLL). The distribution of circulating CD4 + T-cell differentiation subsets across asymptomatic infection and ATLL, and its relationship with presumed transmission route, remain incompletely characterized.

Methods

This single-center, cross-sectional study included 117 adults: 80 asymptomatic HTLV-1 carriers, including 51 with presumed vertical/breastfeeding transmission and 29 with presumed sexual transmission; 21 patients with ATLL; and 16 healthy controls. Peripheral blood mononuclear cells were analyzed by multiparameter flow cytometry. Total CD4 + T-cell frequency was measured among viable CD3 + T cells, while naïve, T memory stem cells (TSCM), central memory T cells (TCM), effector memory T cells (TEM), and terminally differentiated effector T cells (TTE) were defined within viable CD3 + CD4+ T cells using CD45RA, CCR7, and CD95. Group comparisons used Kruskal–Wallis tests followed by Dunn’s tests with Holm adjustment. False-discovery-rate correction was applied across six outcomes, and median regression estimated age- and sex-adjusted differences.

Results

Total CD4 + T-cell frequency showed a nominal difference among groups ($p = 0.0478$) that did not remain significant after false-discovery-rate correction or covariate adjustment. Patients with ATLL had lower median naïve-cell frequencies than asymptomatic carriers and healthy controls (2.91% vs. 27.95% and 24.40%), lower TCM frequencies (12.30% vs. 25.35% and 31.85%), and higher TEM frequencies (49.80% vs. 16.15% and 17.80%). All three global comparisons remained significant after false-discovery-rate correction. After adjustment for age and sex, ATLL remained associated with lower naïve-cell frequencies versus asymptomatic carriers and healthy controls (– 25.00 and – 21.29 percentage points), lower TCM frequencies (– 12.80 and – 16.97), and higher TEM frequencies (+ 34.40 and + 31.85), respectively. TSCM showed no significant unadjusted global difference, and TTE did not differ among groups. No immunophenotypic differences were identified between presumed vertical/breastfeeding and sexual transmission groups.

Conclusions

ATLL was associated with contraction of the naïve and TCM compartments and expansion of TEM cells, independently of age and sex. This pattern reflected redistribution across differentiation states rather

than a consistent increase in total CD4 + T cells. Presumed transmission route was not associated with CD4 + T-cell subset distribution in asymptomatic carriers.

BACKGROUND

CD4 + T lymphocytes comprise phenotypically and functionally heterogeneous populations that coordinate adaptive immune responses, support the activation of other immune cells, and contribute to the establishment and maintenance of immunological memory [1, 2]. Following antigenic stimulation, naïve CD4 + T cells undergo clonal expansion and differentiate into effector and memory populations with distinct migratory, proliferative, and functional properties [3, 4]. Although the developmental relationships among these populations remain incompletely resolved, a widely accepted differentiation framework includes naïve T cells, central memory T cells (TCM), effector memory T cells (TEM), and terminally differentiated effector T cells (TTE), the latter corresponding to a CD45RA-re-expressing effector phenotype frequently referred to as TEMRA [4–6].

These subsets can be distinguished by combinations of cell-surface markers associated with differentiation, survival, and tissue-homing capacity [5–7]. TCM cells generally express C-C chemokine receptor type 7 (CCR7), retain substantial proliferative potential, and preferentially migrate to secondary lymphoid tissues [2, 8]. In contrast, TEM cells lack CCR7 expression, circulate between peripheral tissues and blood, and display more immediate effector functions [1, 6, 8]. TTE cells represent a more differentiated compartment characterized by re-expression of CD45RA, limited proliferative capacity, and enhanced effector activity [5, 6, 9]. In addition to these conventional populations, T memory stem cells (TSCM) constitute a rare and relatively undifferentiated memory subset with long-term self-renewal, multipotency, and high proliferative capacity [10–12]. Phenotypically, TSCM cells share CD45RA and CCR7 expression with naïve cells but can be distinguished by the expression of CD95 and other memory-associated markers [10, 11].

The longevity and self-renewal capacity of TSCM cells have generated particular interest in chronic viral infections and malignancies because these cells may contribute to the maintenance of long-lived infected or transformed cellular compartments [11, 13]. In human T-cell lymphotropic virus type 1 (HTLV-1) infection, previous evidence indicates that infected clones may be distributed across the TSCM, TCM, and TEM compartments [13]. TSCM cells have also been proposed to occupy an early position within the cellular hierarchy of adult T-cell leukemia/lymphoma (ATLL), although their precise contribution to viral persistence, clonal evolution, and leukemic transformation remains under investigation [13, 14].

HTLV-1 was the first human retrovirus associated with malignancy and remains an important cause of chronic infection and severe disease worldwide [14–16]. An estimated 5–10 million people are infected, although the true burden is likely underestimated because population-based surveillance and routine screening remain limited in many endemic regions [16, 17]. Transmission occurs primarily through prolonged breastfeeding, sexual contact, exposure to infected blood components, and the sharing of contaminated needles [16, 18]. Brazil carries a substantial burden of infection, with marked geographic

heterogeneity and persistent challenges related to underdiagnosis, limited awareness, insufficient screening, and restricted access to specialized care [19, 20].

The virus preferentially infects CD4 + T lymphocytes, in which the provirus integrates into the host genome and persists predominantly through the proliferation of infected cellular clones rather than through continuous production of cell-free viral particles [14, 17]. Viral regulatory proteins, particularly Tax and HTLV-1 basic leucine zipper factor (HBZ), dysregulate pathways involved in cellular proliferation, survival, genomic stability, immune evasion, and resistance to apoptosis [14, 21]. Over time, persistent clonal expansion and the accumulation of genetic and epigenetic alterations may promote malignant transformation and the development of ATLL [14, 22, 23].

Most infected individuals remain asymptomatic throughout life; however, approximately 2–5% develop ATLL after a prolonged latency period [14, 24]. ATLL is a mature T-cell neoplasm characterized by the clonal proliferation of infected lymphocytes and substantial clinical, pathological, and molecular heterogeneity [23, 24]. According to the Shimoyama classification, ATLL is categorized into acute, lymphoma, chronic, and smoldering subtypes, which differ markedly in clinical presentation, laboratory abnormalities, therapeutic approach, and prognosis [24, 25]. Despite advances in understanding the molecular evolution of the disease, the redistribution of circulating CD4 + T-cell differentiation subsets across asymptomatic infection and established malignancy remains incompletely characterized [13, 17, 23].

The route and timing of viral acquisition may influence the duration of infection and the subsequent risk of HTLV-1-associated disease. Infection acquired during infancy, particularly through breastfeeding, has traditionally been associated with a greater lifetime risk of ATLL, whereas sexual acquisition generally occurs later in life [14, 17, 24, 26]. Nevertheless, it remains uncertain whether the presumed transmission route is associated with measurable differences in the distribution of circulating CD4 + T-cell subsets. Comparative immunophenotypic evidence addressing this question remains particularly limited in both HTLV-1 and ATLL populations.

Therefore, this study aimed to characterize and compare circulating naïve, TSCM, TCM, TEM, and TTE CD4 + T-cell subsets among asymptomatic carriers, patients with ATLL, and healthy controls. As a secondary objective, we investigated whether the distribution of these subsets differed between asymptomatic carriers with presumed vertical/breastfeeding transmission and those with presumed sexual transmission.

METHODS

Study design and setting

This analytical, single-center, cross-sectional observational study was conducted at the HTLV-1 Outpatient Clinic of the Division of Hematology, Hemotherapy and Cell Therapy, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP), São Paulo, Brazil.

The source population comprised adults with confirmed HTLV-1 infection who were followed at the institutional outpatient clinic, including asymptomatic carriers and patients with ATLL. Healthy controls were recruited among platelet apheresis donors at Fundação Pró-Sangue Hemocentro de São Paulo, São Paulo, Brazil.

Peripheral blood samples were collected between February 2023 and December 2025. The final analytical cohort included 117 participants: 80 asymptomatic carriers, 21 patients with ATLL, and 16 healthy controls. Among the asymptomatic carriers, 51 were classified as having presumed vertical/breastfeeding transmission and 29 as having presumed sexual transmission.

The study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology guidelines for cross-sectional studies [27]. The study design and analytical workflow are summarized in Fig. 1.

Study population and eligibility criteria

Eligible participants were adults aged 18 years or older who provided written informed consent and had sufficient biological material available for flow-cytometric analysis. HTLV-1 infection was confirmed by two independent positive polymerase chain reaction results obtained from samples collected at different time points before study enrollment. Molecular confirmation was performed at the Laboratory of Medical Investigation in Dermatology and Immunodeficiencies (LIM-56), Instituto de Medicina Tropical, Faculdade de Medicina, Universidade de São Paulo.

Asymptomatic carriers were defined as participants with confirmed infection who had no clinical, morphological, histopathological, or immunophenotypic evidence of ATLL or another HTLV-1-associated disease at enrollment. These participants showed no evidence of an abnormal or monoclonal mature T-cell expansion during the institutional diagnostic assessment.

The clinical and laboratory assessment used to establish asymptomatic-carrier status was performed independently of the experimental flow-cytometry analysis conducted in this study. Accordingly, the frequencies of total CD4 + T cells and differentiation subsets investigated as study outcomes were not used to determine eligibility or assign participants to the asymptomatic-carrier group.

For asymptomatic carriers, additional eligibility required sufficient information to classify the presumed route of infection according to the predefined study algorithm. Exclusion criteria comprised inability to classify the presumed transmission route, concomitant acute or chronic infection, malignant neoplasms other than non-melanoma skin cancer, primary cutaneous aggressive epidermotropic CD8 + cytotoxic T-cell lymphoma, and a history of intravenous drug use involving shared needles.

The final sample size reflected the number of eligible participants with available biological samples who fulfilled the clinical and laboratory criteria during the study period. No participant was excluded from the primary analyses because of missing immunophenotypic measurements.

Classification of presumed HTLV-1 transmission route

The presumed route of HTLV-1 acquisition was classified using a predefined clinical and epidemiological algorithm based on maternal anti-HTLV-1/2 serological status, the serological status of current or previous sexual partners, previous exposure to blood-component transfusion, and history of intravenous drug use involving shared needles.

Because the route of infection was not confirmed by phylogenetic analysis or direct molecular linkage between the presumed source and the participant, all categories were designated as presumed transmission routes.

Participants were classified as having presumed vertical/breastfeeding transmission when all the following criteria were fulfilled:

1. positive maternal anti-HTLV-1/2 serology;
2. negative anti-HTLV-1/2 serology in current and previous sexual partners for whom information was available;
3. no history of transfusion of blood components before November 1993, when anti-HTLV-1/2 screening became mandatory in Brazilian blood banks under Ordinance No. 1376/1993 of the Brazilian Ministry of Health; and
4. no history of intravenous drug use involving shared needles.

Participants were classified as having presumed sexual transmission when all the following criteria were met:

1. negative maternal anti-HTLV-1/2 serology;
2. positive anti-HTLV-1/2 serology in a current or previous sexual partner;
3. no history of transfusion of blood components before November 1993; and
4. no history of intravenous drug use involving shared needles.

No participant included in the final analytical cohort was classified as having presumed transfusion-associated infection. Asymptomatic carriers who did not meet all criteria for either presumed vertical/breastfeeding or presumed sexual transmission were not eligible for the transmission-route analysis.

Diagnosis and clinical classification of ATLL

ATLL diagnosis was established through an integrated clinicopathological assessment in accordance with the 2017 World Health Organization classification [28]. The diagnostic evaluation included clinical presentation, hematological findings, peripheral-blood cytomorphology, flow-cytometric identification of an abnormal or clonal mature T-cell population, molecular assessment of T-cell receptor clonality when

indicated, and histopathological examination of involved tissues in patients with nodal or extranodal disease.

In patients with peripheral-blood involvement, the institutional diagnostic assessment included identification of an abnormal mature T-cell population, typically expressing CD3, CD4, and CD25. T-cell clonality was evaluated, when required, by multiplex polymerase chain reaction amplification of the T-cell receptor gamma gene.

In patients presenting predominantly with lymphadenopathy or extranodal disease, diagnosis was supported by histopathological and immunophenotypic examination of tissue-biopsy specimens compatible with a mature T-cell neoplasm in the context of confirmed HTLV-1 infection.

ATLL cases were classified according to the Shimoyama criteria as acute, lymphoma, chronic, or smoldering. For secondary clinical and laboratory comparisons, acute and lymphoma subtypes were grouped as aggressive ATLL, whereas chronic and smoldering subtypes were grouped as indolent ATLL [25].

The experimental characterization of circulating CD4 + T-cell differentiation subsets performed in this study was not used as an isolated criterion to establish ATLL diagnosis or assign clinical subtype.

Healthy control group

Healthy controls were recruited among individuals undergoing their first or second platelet apheresis donation at Fundação Pró-Sangue Hemocentro de São Paulo.

All controls fulfilled the clinical, hematological, and serological eligibility requirements for blood donation established by applicable regulations and institutional procedures. Routine donor screening included negative serological testing for HTLV-1/2.

Restriction to individuals with limited prior exposure to platelet apheresis was intended to minimize the potential influence of repeated procedures on circulating lymphocyte populations, particularly CD4 + T-cell frequencies [29].

Controls had no known history of HTLV-1 infection, hematological malignancy, active infection, or any clinical condition that would preclude blood donation.

Clinical and laboratory data collection

Demographic, clinical, and laboratory data were extracted from institutional electronic medical records by authorized members of the research team and entered into a study-specific Research Electronic Data Capture database hosted at HCFMUSP [30]. Each participant was assigned a unique coded identifier, and direct personal identifiers were not included in the analytical dataset.

Variables collected included age, sex, presumed HTLV-1 transmission route, ATLL subtype, absolute lymphocyte count, lactate dehydrogenase, albumin, total calcium, uric acid, beta-2 microglobulin, alkaline phosphatase, and urea. For participants with serum albumin concentrations below 3.5 g/dL, total calcium was corrected for albumin according to the institutional laboratory procedure.

Laboratory measurements corresponded to the results obtained closest to the date of peripheral blood collection for immunophenotypic analysis. When more than one result was available, the measurement nearest to the collection date was selected.

Missing laboratory data were not imputed. Participants without an available result for a specific variable were excluded only from the corresponding analysis and remained eligible for all other analyses for which data were available.

Peripheral blood collection and mononuclear cell isolation

For asymptomatic carriers and patients with ATLL, up to 50 mL of peripheral blood was collected into tubes containing ethylenediaminetetraacetic acid as anticoagulant. For healthy controls, leukoreduction filters were obtained immediately after platelet apheresis, and retained leukocytes were recovered using the standardized institutional procedure before mononuclear cell isolation.

Peripheral blood mononuclear cells were isolated by density-gradient centrifugation using Ficoll-Paque PLUS, according to the manufacturer's instructions. After centrifugation, the mononuclear cell layer was collected and washed to remove residual platelets, plasma proteins, and Ficoll.

Cell viability was assessed by Trypan Blue exclusion using 0.4% Trypan Blue solution, a Neubauer hemocytometer, and optical microscopy. Following isolation, lymphocyte concentration was measured using a Sysmex XN-550 automated hematology analyzer in CBC + DIFF mode.

Flow-cytometry staining panel

For each sample, approximately 15×10^6 viable lymphocytes were used for flow-cytometric staining. Cells were incubated with a multiparameter monoclonal-antibody panel comprising anti-CD45RA conjugated to allophycocyanin, clone HI100; anti-CCR7 conjugated to ECD, clone 2-L1-A; anti-CD95 conjugated to BB515, clone DX2; anti-CD4 conjugated to PE-Cy7, clone SK3; and anti-CD3 conjugated to BB700, clone SK7. All monoclonal antibodies were obtained from BD Biosciences. Fixable Viability Stain 780, detected in the APC-H7 channel, was used to exclude non-viable cells and was also obtained from BD Biosciences.

All antibodies were titrated before study initiation and used at their experimentally determined optimal concentrations. Cells were incubated with the antibody cocktail for 20 min at room temperature,

protected from light. After staining, cells were washed and resuspended in phosphate-buffered saline containing 2 mmol/L ethylenediaminetetraacetic acid and 0.5% human serum albumin.

Samples were acquired using a BD FACSAria Fusion flow cytometer and analyzed with FlowJo software version 10.10.0.

Fluorescence controls

Unstained and fluorescence-minus-one controls were used to define the boundaries between negative and positive populations and to assess potential fluorescence spillover affecting the principal differentiation markers [31].

Separate fluorescence-minus-one controls were prepared for CD45RA, CCR7, and CD95 by staining cells with the complete antibody panel except for the antibody corresponding to the marker under evaluation.

The fluorescence-minus-one controls were used to support consistent placement of the gates separating CD45RA-positive and CD45RA-negative, CCR7-positive and CCR7-negative, and CD95-positive and CD95-negative populations. Representative fluorescence-minus-one controls are shown in Fig. 2.

Flow-cytometry gating strategy

A standardized sequential gating strategy was applied to all samples. Cellular aggregates and doublets were first excluded using forward-scatter height versus forward-scatter area and side-scatter area versus side-scatter height parameters. The lymphocyte population was then identified according to forward- and side-scatter characteristics.

Non-viable cells were excluded using Fixable Viability Stain 780. CD3 + T lymphocytes were identified within the viable lymphocyte population, and CD4 + T lymphocytes were subsequently identified within viable CD3 + T cells. Total CD4 + T-cell frequency was expressed as the percentage of viable CD3 + T cells.

CD4 + T-cell differentiation subsets were identified within the viable CD3 + CD4+ population according to the expression of CD45RA, CCR7, and CD95 [6, 9, 10, 13, 32]. The operational phenotypic definitions are summarized in Table 1. Subset frequencies were expressed as percentages of viable CD3 + CD4+ T cells.

Table 1
Operational phenotypic definitions of circulating CD4 + T-cell subsets

CD4 + T-cell subset	Operational phenotype
Naïve	CD45RA+CCR7 + CD95–
T memory stem cells (TSCM)	CD45RA+CCR7 + CD95+
Central memory T cells (TCM)	CD45RA–CCR7 + CD95+
Effector memory T cells (TEM)	CD45RA–CCR7 – CD95+
Terminally differentiated effector T cells (TTE)	CD45RA+CCR7 – CD95+

RESULTS

Study population and data completeness

A total of 117 participants were included in the final analytical cohort: 80 asymptomatic carriers, 21 patients with ATLL, and 16 healthy controls. Among asymptomatic carriers, 51/80 (63.8%) were classified as having presumed vertical/breastfeeding transmission and 29/80 (36.3%) as having presumed sexual transmission.

All participants had complete measurements for the six prespecified immunophenotypic outcomes: total CD4 + T-cell frequency and the frequencies of naïve, TSCM, TCM, TEM, and TTE cells. Therefore, all 117 participants were included in the primary unadjusted comparisons and in the age- and sex-adjusted median regression models.

Clinical and biochemical variables were analyzed using an available-case approach. Depending on the variable, laboratory measurements were available for 94–98 participants in the overall cohort. Missingness in variables restricted to patients with ATLL, including clinical subtype and disease behavior, was structural and did not affect the primary immunophenotypic analyses. No immunophenotypic value was excluded because of an extreme observation, and no missing value was imputed.

Among the 21 patients with ATLL, eight had the smoldering subtype (38.1%), seven had acute disease (33.3%), four had the lymphoma subtype (19.0%), and two had chronic disease (9.5%). Acute and lymphoma cases were classified as aggressive ATLL (n = 11), whereas chronic and smoldering cases were classified as indolent ATLL (n = 10).

At the time of peripheral blood collection, 17 patients (81.0%) were evaluated at diagnosis before treatment initiation. The remaining four samples were obtained after one cycle of CHOP, after six cycles of mini-CHOP, after initiation of zidovudine, and during late relapse, respectively.

Demographic and clinical characteristics

Healthy controls had a median age of 42.5 years (interquartile range [IQR], 36.5–59.5), compared with 50.0 years (IQR, 46.0–56.0) among asymptomatic carriers with presumed sexual transmission, 57.0 years (IQR, 44.0–66.5) among carriers with presumed vertical/breastfeeding transmission, and 57.0 years (IQR, 43.0–68.0) among patients with ATLL. Although healthy controls were descriptively younger than the HTLV-1–infected groups, the global difference in age did not reach statistical significance (Kruskal–Wallis $p = 0.063$).

Women represented 68.8% of healthy controls, 89.7% of carriers with presumed sexual transmission, 82.4% of carriers with presumed vertical/breastfeeding transmission, and 71.4% of patients with ATLL. Sex distribution did not differ significantly among the four study groups (Fisher's exact $p = 0.232$).

Among the 80 asymptomatic carriers, 68 were women (85.0%) and 12 were men (15.0%). The presumed vertical/breastfeeding transmission group comprised 42 women and nine men, whereas the presumed sexual transmission group comprised 26 women and three men. Clinical and laboratory characteristics are described in Table S1.

Among the 21 patients with ATLL, eight had the smoldering subtype (38.1%), seven had acute ATLL (33.3%), four had the lymphoma subtype (19.0%), and two had chronic ATLL (9.5%). Acute and lymphoma cases were grouped as aggressive ATLL ($n = 11$), whereas chronic and smoldering cases were grouped as indolent ATLL ($n = 10$).

The aggressive ATLL group comprised seven women (63.6%) and four men (36.4%), whereas the indolent group comprised eight women (80.0%) and two men (20.0%). Median age was 55.0 years (IQR, 43.0–63.5) among patients with aggressive ATLL and 62.0 years (IQR, 53.2–71.0) among those with indolent ATLL, with no statistically significant difference between groups ($p = 0.359$).

At the time of peripheral-blood collection, 17 patients with ATLL (81.0%) were evaluated at diagnosis before treatment initiation. The remaining four samples were obtained after one cycle of CHOP, after six cycles of mini-CHOP, after initiation of zidovudine, and during late relapse, respectively. The clinical and laboratory characteristics of patients with aggressive and indolent ATLL are presented in Table 2.

Table 2
Demographic and laboratory characteristics of patients with aggressive and indolent ATLL

Characteristic	Aggressive ATLL (n = 11)	Indolent ATLL (n = 10)	Unadjusted <i>p</i> value
Demographic characteristics			
Age, years	55.0 (43.0–63.5)	62.0 (53.2–71.0)	0.359
Sex, n (%)			
Female	7 (63.6)	8 (80.0)	—
Male	4 (36.4)	2 (20.0)	—
Laboratory characteristics			
Absolute lymphocyte count, $\times 10^3/\mu\text{L}$	3.5 (1.3–9.7)	3.0 (1.5–4.2)	0.698
Lactate dehydrogenase, U/L	613.0 (488.5–1518.0)	204.5 (170.8–214.3)	< 0.001
Corrected total calcium, mg/dL	9.9 (9.2–10.9)	9.4 (9.2–9.5)	0.161
Albumin, g/dL	3.20 (2.52–3.88)	4.40 (4.30–4.60)	0.002
Uric acid, mg/dL	5.9 (5.4–8.9)	4.8 (3.9–5.4)	0.052
Beta-2 microglobulin, mg/L	2.8 (2.2–5.4)	1.8 (1.7–2.5)	0.037
Alkaline phosphatase, U/L	146.0 (112.5–300.0)	76.0 (67.3–82.0)	0.003
Urea, mg/dL	37.0 (24.5–51.5)	29.5 (27.2–33.8)	0.324

Asymptomatic HTLV-1 carriers versus healthy controls

No immunophenotypic outcome differed independently between asymptomatic HTLV-1 carriers and healthy controls.

The adjusted median difference in total CD4 + T-cell frequency was – 2.67 percentage points (95% CI, – 10.06 to 4.71; *p* = 0.479; FDR-adjusted *p* = 0.651).

For the differentiation subsets, the adjusted median differences were + 3.71 percentage points for naïve cells (95% CI, – 8.20 to 15.61; *p* = 0.543; FDR-adjusted *p* = 0.651), + 0.38 percentage points for TSCM (95% CI, – 0.25 to 1.01; *p* = 0.237; FDR-adjusted *p* = 0.651), – 4.17 percentage points for TCM (95% CI, – 15.40 to 7.07; *p* = 0.469; FDR-adjusted *p* = 0.651), – 2.55 percentage points for TEM (95% CI, – 10.75 to 5.64; *p* = 0.543; FDR-adjusted *p* = 0.651), and – 0.01 percentage points for TTE (95% CI, – 0.52 to 0.49; *p* = 0.954; FDR-adjusted *p* = 0.954).

All confidence intervals included zero, and none of these comparisons remained statistically significant after FDR correction.

Integrated interpretation of the adjusted models

The adjusted analyses confirmed that the principal ATLL-associated immunophenotypic pattern was not explained by differences in age or sex. The most consistent alterations were pronounced contraction of the naïve compartment, lower representation of TCM cells, and marked expansion of TEM cells in patients with ATLL relative to both asymptomatic carriers and healthy controls.

The absence of adjusted differences in total CD4 + T-cell frequency indicates that the principal distinction among the clinical groups involved redistribution across differentiation states rather than a uniform increase in the overall circulating CD4 + compartment. Similarly, the absence of increased TTE frequencies suggests that the ATLL-associated profile was concentrated predominantly in the TEM compartment rather than reflecting generalized accumulation across all later stages of differentiation.

The small adjusted reduction in TSCM frequency observed only between patients with ATLL and asymptomatic carriers should be interpreted cautiously because of its limited absolute magnitude and the absence of a significant unadjusted global difference. Overall, the convergence between the unadjusted and adjusted analyses supports a robust ATLL-associated shift from naïve and central memory compartments toward an effector memory–predominant circulating CD4 + T-cell profile.

Presumed HTLV-1 transmission route was not associated with the circulating immunophenotypic profile

Analyses restricted to the 80 asymptomatic HTLV-1 carriers showed no evidence that presumed HTLV-1 transmission route was independently associated with total CD4 + T-cell frequency or with the distribution of the five CD4 + T-cell differentiation subsets. Of these participants, 51 were classified as having presumed vertical/breastfeeding transmission and 29 as having presumed sexual transmission.

In the unadjusted four-group analyses no statistically significant pairwise difference was detected between carriers with presumed sexual transmission and those with presumed vertical/breastfeeding transmission for any of the five differentiation subsets. Median naïve-cell frequencies were 26.30% (IQR, 13.60–34.80) and 31.60% (IQR, 17.05–39.70), respectively, and the pairwise comparison was not significant after Holm correction.

Median TSCM frequencies were 1.14% (IQR, 0.53–2.60) among carriers with presumed sexual transmission and 1.60% (IQR, 0.45–3.48) among those with presumed vertical/breastfeeding transmission. Median TCM frequencies were 22.40% (IQR, 17.20–34.20) and 25.60% (IQR, 15.40–34.05), respectively. Neither comparison was statistically significant after Holm adjustment.

Similarly, median TEM frequencies were 17.10% (IQR, 14.00–23.50) in the presumed sexual-transmission group and 14.80% (IQR, 12.10–23.20) in the presumed vertical/breastfeeding-transmission group.

Median TTE frequencies were 0.91% (IQR, 0.21–1.29) and 0.44% (IQR, 0.21–1.14), respectively. No significant pairwise difference was detected for either subset.

The absence of route-associated differences was maintained after adjustment for age and sex. Using presumed sexual transmission as the reference category, presumed vertical/breastfeeding transmission was associated with an adjusted median difference of – 3.26 percentage points in total CD4 + T-cell frequency. The corresponding adjusted median differences were + 2.00 percentage points for naïve cells, + 0.17 percentage points for TSCM, + 1.29 percentage points for TCM, – 2.30 percentage points for TEM, and – 0.36 percentage points for TTE.

The 95% confidence intervals for all six adjusted estimates included zero. None of the raw associations remained statistically significant after correction for multiple testing, and the FDR-adjusted p value was 0.771 for each outcome.

Accordingly, the absence of an association between presumed transmission route and the circulating CD4 + T-cell profile was consistent across the unadjusted pairwise comparisons, age- and sex-adjusted median regression models, and effect-size analyses. Within the limitations of the predefined classification algorithm, asymptomatic carriers with presumed sexual and presumed vertical/breastfeeding acquisition showed comparable total CD4 + T-cell frequencies and similar distributions of naïve, TSCM, TCM, TEM, and TTE cells.

Aggressive ATLL was associated with an adverse biochemical profile and pronounced contraction of naïve CD4 + T cells

The 21 patients with ATLL were classified according to clinical behavior. Eleven patients had aggressive ATLL, comprising the acute and lymphoma subtypes, whereas ten had indolent ATLL, comprising the chronic and smoldering subtypes.

Median age was 55.0 years (IQR, 43.0–63.5) in the aggressive group and 62.0 years (IQR, 53.2–71.0) in the indolent group ($p = 0.359$). Women represented 7/11 patients (63.6%) with aggressive ATLL and 8/10 patients (80.0%) with indolent ATLL, whereas men represented 4/11 (36.4%) and 2/10 (20.0%), respectively. Thus, no marked demographic imbalance was identified between the two groups.

Absolute lymphocyte counts showed considerable dispersion, particularly in aggressive disease. Median lymphocyte counts were $3.5 \times 10^3/\mu\text{L}$ (IQR, 1.3–9.7) in aggressive ATLL and $3.0 \times 10^3/\mu\text{L}$ (IQR, 1.5–4.2) in indolent ATLL. This difference was not statistically significant in the unadjusted analysis ($p = 0.698$) and remained non-significant after correction for multiple comparisons within the clinical and laboratory family.

Patients with aggressive ATLL had substantially higher lactate dehydrogenase concentrations than those with indolent disease. Median lactate dehydrogenase was 613.0 U/L (IQR, 488.5–1518.0) in aggressive ATLL and 204.5 U/L (IQR, 170.8–214.3) in indolent ATLL. The unadjusted difference was highly significant ($p < 0.001$) and remained significant after Holm correction (adjusted $p = 0.001$).

Serum albumin was lower in patients with aggressive ATLL. Median albumin concentrations were 3.20 g/dL (IQR, 2.52–3.88) in the aggressive group and 4.40 g/dL (IQR, 4.30–4.60) in the indolent group. This difference was significant before ($p = 0.002$) and after Holm correction (adjusted $p = 0.019$).

Alkaline phosphatase was also higher in aggressive ATLL, with median concentrations of 146.0 U/L (IQR, 112.5–300.0), compared with 76.0 U/L (IQR, 67.3–82.0) in indolent ATLL. The difference was statistically significant in the unadjusted analysis ($p = 0.003$) and remained significant after Holm correction (adjusted $p = 0.024$).

Beta-2 microglobulin concentrations were nominally higher in aggressive ATLL. The median was 2.8 mg/L (IQR, 2.2–5.4) in the aggressive group and 1.8 mg/L (IQR, 1.7–2.5) in the indolent group ($p = 0.037$). However, this association did not remain statistically significant after correction for multiple comparisons within the clinical and laboratory family.

Median uric acid concentration was 5.9 mg/dL (IQR, 5.4–8.9) in aggressive ATLL and 4.8 mg/dL (IQR, 3.9–5.4) in indolent ATLL. This difference approached, but did not reach, the prespecified threshold for statistical significance in the unadjusted analysis ($p = 0.052$) and remained non-significant after multiplicity correction.

Albumin-corrected total calcium did not differ significantly between the groups. Median corrected calcium was 9.9 mg/dL (IQR, 9.2–10.9) in aggressive ATLL and 9.4 mg/dL (IQR, 9.2–9.5) in indolent ATLL ($p = 0.161$). Similarly, median urea concentrations were 37.0 mg/dL (IQR, 24.5–51.5) and 29.5 mg/dL (IQR, 27.2–33.8), respectively ($p = 0.324$). Neither variable remained significant after Holm correction. These clinical and laboratory findings are summarized in Table 2.

The exploratory immunophenotypic comparison showed the most pronounced difference in the least differentiated CD4 + T-cell compartment. Patients with aggressive ATLL had a median naïve-cell frequency of 0.79% (IQR, 0.24–2.08), compared with 22.00% (IQR, 6.11–26.12) among patients with indolent ATLL. This difference was significant in the unadjusted comparison ($p = 0.008$) and remained at the threshold of statistical significance after Holm correction across the six immunophenotypic outcomes (adjusted $p = 0.0496$).

In contrast, total CD4 + T-cell frequency and the TSCM, TCM, TEM, and TTE compartments did not differ significantly between aggressive and indolent ATLL after correction for multiple testing. Therefore, within the relatively small ATLL subgroup, aggressive disease was characterized predominantly by biochemical evidence of greater disease activity and tissue involvement, together with a marked contraction of the circulating naïve CD4 + T-cell compartment, Fig. 6A–D.

Correlation analyses connected CD4 + differentiation states with markers of disease activity

Correlation analyses were performed to characterize the relationships among circulating CD4 + T-cell differentiation subsets and between immunophenotypic and laboratory variables. Spearman correlation coefficients were calculated in the overall cohort and subsequently within the ATLL subgroup. Because a large number of pairwise associations were examined, statistical significance was evaluated using false-discovery-rate-adjusted p values.

Correlation structure in the overall cohort

Among the immunophenotypic variables, naïve-cell frequency was positively correlated with TSCM frequency (Spearman's $\rho = 0.484$; FDR-adjusted $p < 0.001$). Conversely, naïve-cell frequency was inversely correlated with TEM frequency ($\rho = -0.474$; FDR-adjusted $p < 0.001$). These associations were consistent with coordinated variation between less differentiated and effector-memory compartments.

TSCM frequency was positively correlated with TTE frequency ($\rho = 0.394$; FDR-adjusted $p < 0.001$) and inversely correlated with TEM frequency ($\rho = -0.313$; FDR-adjusted $p = 0.006$). A weaker positive association was observed between TEM and TTE frequencies ($\rho = 0.276$; FDR-adjusted $p = 0.021$). Together, these findings showed that the evaluated differentiation subsets were not statistically independent and that variation in one compartment was accompanied by changes in other regions of the operational CD4 + T-cell differentiation continuum.

Among the laboratory variables, beta-2 microglobulin was positively correlated with urea ($\rho = 0.544$; FDR-adjusted $p < 0.001$), uric acid ($\rho = 0.412$; FDR-adjusted $p < 0.001$), and lactate dehydrogenase (LDH; $\rho = 0.381$; FDR-adjusted $p = 0.002$).

LDH was also associated with the immunophenotypic profile. Higher LDH concentrations were correlated with lower naïve-cell frequencies ($\rho = -0.303$; FDR-adjusted $p = 0.021$) and higher TEM frequencies ($\rho = 0.283$; FDR-adjusted $p = 0.035$). Thus, increasing biochemical evidence of disease activity was associated with reduced representation of the naïve compartment and greater representation of effector memory cells in the overall cohort.

Correlation structure within the ATLL subgroup

Several associations were stronger when the analysis was restricted to the 21 patients with ATLL. TEM and TSCM frequencies were strongly inversely correlated ($\rho = -0.745$; FDR-adjusted $p = 0.011$), indicating that patients with greater representation of the TEM compartment tended to have lower circulating TSCM frequencies.

LDH showed a strong inverse correlation with serum albumin ($\rho = -0.709$; FDR-adjusted $p = 0.020$) and with naïve-cell frequency ($\rho = -0.666$; FDR-adjusted $p = 0.026$). In contrast, LDH was positively correlated with alkaline phosphatase ($\rho = 0.687$; FDR-adjusted $p = 0.020$).

An additional inverse association was observed between TCM and TTE frequencies within the ATLL subgroup ($\rho = -0.631$; FDR-adjusted $p = 0.045$).

The convergence between these correlation analyses and the aggressive-versus-indolent comparison was noteworthy. Clinically aggressive ATLL was characterized by higher LDH, lower albumin, and markedly lower naïve-cell frequency. In the continuous correlation analysis restricted to patients with ATLL, increasing LDH was likewise associated with decreasing albumin and decreasing naïve-cell frequency.

Nevertheless, these subgroup associations should be interpreted as exploratory because they were estimated from only 21 patients with ATLL and involved multiple pairwise correlations. They should not be considered independently validated prognostic relationships. The correlation matrices for the overall cohort and the ATLL subgroup are presented in Fig. 6E and Fig. 6F, respectively.

DISCUSSION

In this cross-sectional study, we identified a reproducible redistribution of circulating CD4 + T-cell differentiation subsets associated with established ATLL. The principal pattern consisted of marked contraction of the naïve compartment, reduced representation of TCM cells, and pronounced expansion of TEM cells. These differences remained evident after correction for multiple testing and adjustment for age and sex, were supported by large or moderate effect sizes, and were concordant with the correlation and multivariate analyses. In contrast, the overall frequency of circulating CD4 + T cells showed only a nominal and non-robust difference, while TSCM and TTE frequencies were comparatively preserved. To our knowledge, this study provides one of the most comprehensive evaluations of the circulating CD4 + differentiation continuum across healthy controls, asymptomatic HTLV-1 infection, and clinically established ATLL, while also examining whether the same profile varies according to the presumed route of viral acquisition.

The nominally higher total CD4 + T-cell frequency observed in patients with ATLL is biologically plausible because most leukemic clones display a mature CD4 + phenotype [14, 24, 34]. HTLV-1 persists predominantly through mitotic expansion of infected cells, and ATLL emerges after the selection and expansion of clones that have accumulated viral, genetic, epigenetic, and host-cell alterations affecting proliferation, survival, immune regulation, and genomic stability [14, 23, 34–36]. Nevertheless, the total CD4 + result should not be regarded as a defining finding of this study. The initial global comparison was only nominally significant, the effect size was small, and the association did not remain significant after false-discovery-rate correction or adjustment for age and sex. Moreover, the outcome represented the relative frequency of CD4 + cells among viable CD3 + T lymphocytes rather than an absolute CD4 + count or direct measurement of the malignant clone. The observed values may therefore reflect varying contributions from leukemic cells, non-malignant infected cells, and residual uninfected CD4 + T lymphocytes. The wide distribution within the ATLL group is also compatible with the substantial heterogeneity of leukemic and lymphomatous presentations, clinical subtypes, peripheral-blood involvement, and treatment exposure [24, 34].

In contrast, contraction of the naïve CD4 + compartment was one of the most robust findings. Patients with ATLL had substantially lower naïve-cell frequencies than both asymptomatic carriers and healthy controls, with large global and pairwise effect sizes. The median frequency in ATLL was approximately one tenth of that observed in the other groups, and the association remained significant in the adjusted models. This finding may reflect proportional replacement of the less differentiated compartment by expanded antigen-experienced or neoplastic populations. Chronic HTLV-1 infection is accompanied by persistent immune stimulation, proliferation of infected clones, and continuing interactions between viral products and host-cell signaling pathways [14, 21, 35, 37, 38], all of which may reshape the circulating CD4 + T-cell pool over time.

The reduction in naïve cells may also reflect impairment of immune-compartment maintenance in clinically established disease. Ageing is associated with thymic involution and progressive loss of naïve T cells [39, 40], and the healthy controls in this study were descriptively younger than the infected groups. However, the naïve-cell difference remained pronounced after adjustment for age and sex, making demographic variation an insufficient explanation for the observed contraction. Nevertheless, residual confounding related to immunosenescence, comorbidities, inflammatory burden, or unmeasured treatment effects cannot be entirely excluded. Age- and sex-matched validation cohorts and direct assessment of thymic output, including recent thymic emigrant markers or T-cell receptor excision circles, could help distinguish reduced production from proportional displacement by expanded differentiated populations [41].

The cross-sectional design also requires caution when interpreting the relationship between naïve-cell loss and disease development. The present findings do not demonstrate that naïve cells differentiated directly into the expanded TEM population, that contraction of the naïve compartment preceded transformation, or that naïve-cell depletion contributes causally to leukemogenesis. Rather, they establish a strong association between established ATLL and lower representation of circulating phenotypically naïve cells. Longitudinal studies beginning in asymptomatic carriers will be required to determine whether progressive naïve-cell contraction accompanies clonal evolution or instead emerges because of clinically apparent malignancy.

The reciprocal pattern of lower TCM and higher TEM frequencies provided the clearest evidence of altered differentiation architecture in ATLL. TCM cells retain CCR7 expression, lymphoid-homing capacity, and substantial proliferative potential, whereas TEM cells lack CCR7 and constitute a more differentiated circulating compartment capable of rapid effector responses [6, 8, 42]. Patients with ATLL had lower TCM frequencies than both healthy controls and asymptomatic carriers and markedly higher TEM frequencies than every comparison group. TEM showed the largest global effect size among all evaluated subsets and the largest pairwise effects involving ATLL, indicating that this was not merely a statistically detectable difference but the dominant feature of the circulating phenotype.

The combination of TCM contraction and TEM expansion is compatible with an effector-memory-predominant profile. Similar changes toward mature TEM phenotypes have been described in clinically

manifest HTLV-1-associated inflammatory disease, including untreated HTLV-1-associated myelopathy [37]. Although inflammatory neurological disease and ATLL are biologically distinct, the observations suggest that clinically manifest HTLV-1-associated conditions may be accompanied by remodeling of the circulating memory compartment. In ATLL, however, TEM predominance may result from several non-mutually exclusive mechanisms, including the differentiation phenotype of the dominant leukemic clone, selective survival of differentiated infected clones [14, 23], persistent immune activation of non-malignant T cells, altered trafficking between blood and tissues, and treatment-related changes in the residual immune compartment [37].

Chronological genomic and single-cell studies indicate that ATLL evolution involves the emergence, selection, and replacement of genetically and transcriptionally distinct subclones, with alterations affecting T-cell receptor signaling, STAT3, NOTCH, immune regulation, and other pathways arising before and during clinically apparent disease [22, 23, 34, 43]. The circulating phenotype detected here may therefore represent an accessible population-level correlate of a considerably more complex clonal and transcriptional architecture. It should not be interpreted as evidence that TEM cells are the exclusive cell of origin of ATLL or that leukemogenesis follows a simple linear TCM-to-TEM developmental transition.

The absence of a corresponding expansion of TTE cells is an important component of this interpretation. TTE frequencies were low across all clinical groups, and both the three-group and four-group comparisons showed essentially null global effect sizes. TTE frequency also remained non-significant in the age- and sex-adjusted models. Therefore, the ATLL-associated profile did not reflect uniform accumulation across all late differentiation stages. Instead, the alteration appeared concentrated in the TEM compartment, without a comparable increase in the operationally defined CD45RA+CCR7 – CD95+ population.

This distinction argues against describing the findings simply as generalized terminal differentiation. The term “effector-memory-predominant” more accurately reflects the observed profile. At the same time, the absence of a quantitative difference in TTE cells should not be interpreted as proof of functional equivalence. TTE cells were rare and showed considerable interindividual variation. Additional markers of activation, senescence, exhaustion, cytotoxic capacity, and proliferative history would be required to determine whether phenotypically similar TTE frequencies conceal functional or molecular differences among the groups.

TSCM cells also remained comparatively stable in the unadjusted analyses. These cells constituted a small fraction of circulating CD4 + T cells, and no significant global difference was detected across the three or four clinical groups. The adjusted comparison between ATLL and asymptomatic carriers identified a small reduction of approximately 0.65 percentage points, but the association was close to the multiplicity-adjusted significance threshold, was limited in absolute magnitude, and was not reproduced in the comparison between ATLL and healthy controls. It should therefore be interpreted as a secondary and comparatively weak finding.

The absence of a large quantitative difference in total circulating TSCM cells does not exclude an important biological role. TSCM cells are long-lived, self-renewing, and multipotent [11], characteristics that make them plausible reservoirs for persistent infection and clonal evolution. Nagai et al. demonstrated that HTLV-1-infected TSCM cells may occupy an early position in a functional ATLL cellular hierarchy and reproduce the phenotypic diversity of the original leukemic population [13]. The present study measured the total phenotypically defined TSCM compartment but did not determine which cells carried integrated HTLV-1, their proviral burden, integration sites, mutational profile, clonogenic capacity, or relationship to the dominant ATLL clone. A small infected TSCM fraction could therefore have major biological relevance without producing a detectable increase in total TSCM frequency.

This distinction between abundance and biological function is essential. Stable total TSCM frequency should not be interpreted as evidence that TSCM cells are uninvolved in ATLL pathogenesis. Instead, the findings indicate that frequency alone is insufficient to define their contribution. Cell sorting followed by subset-specific proviral-load measurement, integration-site analysis, T-cell receptor clonality assessment, and genomic or transcriptomic characterization would be required to determine whether phenotypically similar TSCM compartments differ substantially in infection burden and clonal composition.

The age- and sex-adjusted models strengthened the central interpretation of the study. Relative to asymptomatic carriers, patients with ATLL had adjusted median frequencies approximately 25 percentage points lower for naïve cells, 13 percentage points lower for TCM cells, and 34 percentage points higher for TEM cells. Similar differences were observed relative to healthy controls, with reductions of approximately 21 percentage points in naïve cells and 17 percentage points in TCM cells and an increase of approximately 32 percentage points in TEM cells. All these associations remained significant after correction for multiple testing. In contrast, neither total CD4 + nor TTE frequency differed independently between ATLL and the reference groups.

These adjusted results demonstrate that the principal ATLL-associated pattern cannot be explained solely by the older age distribution or sex composition of the clinical groups. They also reinforce the distinction between a modest, non-robust change in total CD4 + frequency and a large redistribution within the differentiation compartments. The convergence of the unadjusted comparisons, effect sizes, and adjusted models supports the conclusion that the composition of the circulating CD4 + compartment is more informative than its overall proportional abundance in characterizing ATLL at the group level.

No comparable redistribution was identified between asymptomatic carriers classified as having presumed vertical/breastfeeding transmission and those classified as having presumed sexual transmission. The two groups did not differ significantly in total CD4+, naïve, TSCM, TCM, TEM, or TTE frequencies in either the unadjusted or adjusted analyses. This finding does not contradict epidemiological evidence that infection acquired early in life [24], particularly through prolonged

breastfeeding [26], is associated with a greater lifetime risk of ATLL. Route and timing of acquisition may influence duration of infection, cumulative clonal expansion, and opportunity for acquisition of oncogenic alterations without necessarily producing a stable route-specific distribution of circulating CD4 + subsets decades later [14].

The result should nevertheless be interpreted within the limits of the exposure-classification strategy. Transmission route was inferred from maternal and partner serology, transfusion history, and reported exposure rather than confirmed through paired viral sequencing or phylogenetic linkage. Although the predefined criteria increased classification consistency, historical information may have been incomplete, and more than one plausible exposure could have occurred during a participant's lifetime. Molecular comparison of viral sequences from participants and probable sources would reduce misclassification and help determine whether the absence of immunophenotypic differences persists when transmission route is more directly established [18].

The comparison between aggressive and indolent ATLL provided clinically coherent secondary findings. After correction for multiple testing, aggressive ATLL was associated with higher lactate dehydrogenase and alkaline phosphatase concentrations and lower serum albumin. These abnormalities are consistent with greater disease activity [44], tissue involvement, and systemic burden in the acute and lymphoma subtypes [24]. Beta-2 microglobulin was nominally higher in aggressive ATLL but did not remain significant after multiplicity adjustment and should therefore not be presented as an independently supported difference in this cohort.

The immunophenotypic comparison showed that aggressive ATLL was additionally associated with marked contraction of the naïve compartment. Patients with aggressive disease had a median naïve-cell frequency below 1%, compared with approximately 22% in indolent ATLL. Although this association remained at the threshold of significance after correction across the immunophenotypic outcomes, the analysis involved only 21 patients and should be considered exploratory. It nevertheless suggests that depletion of the least differentiated circulating compartment may be particularly pronounced in biologically active disease. The absence of significant differences in the other subsets may reflect true similarity, substantial heterogeneity, or limited statistical power after clinical stratification.

The correlation analyses provided further internal support for the principal biological interpretation. Across the complete cohort, naïve-cell frequency was inversely correlated with TEM frequency, indicating coordinated variation between less differentiated and effector-memory compartments. Higher lactate dehydrogenase concentrations were associated with lower naïve-cell and higher TEM frequencies. Within the ATLL subgroup, lactate dehydrogenase was strongly inversely correlated with naïve-cell frequency and serum albumin and positively correlated with alkaline phosphatase. TEM and TSCM frequencies were also strongly inversely related within ATLL.

The convergence between categorical and continuous analyses is noteworthy. Aggressive ATLL was characterized by higher lactate dehydrogenase, lower albumin, and lower naïve-cell frequency, while the subgroup correlation analysis showed that increasing lactate dehydrogenase was associated with

decreasing albumin and naïve-cell representation. These relationships connect the circulating differentiation profile with markers of disease activity. However, because the ATLL subgroup included only 21 patients and multiple correlations were examined, these findings should not be interpreted as independently validated prognostic associations.

The multivariate analyses were also concordant with the univariate results. The first two principal components accounted for more than half of the total immunophenotypic variance, and the principal axis was largely defined by higher TEM and lower naïve-cell frequencies. Healthy controls and asymptomatic carriers showed substantial overlap, whereas patients with ATLL were more frequently displaced toward the region associated with the effector-memory-predominant profile. Clinical group explained approximately 14% of the overall six-dimensional immunophenotypic variation before and after adjustment for age and sex.

Nevertheless, ATLL observations did not form a single compact or completely separated cluster. Patients with ATLL showed broader dispersion, and multivariate dispersion differed significantly among the groups. The PERMANOVA result therefore likely reflected both differences in average group profiles and differences in within-group variability [45]. This is biologically plausible given the heterogeneity of ATLL subtypes, peripheral-blood involvement, clonal composition, disease burden, and treatment exposure [24, 34]. The findings support an ATLL-associated profile at the group level but do not imply that every patient shares an identical phenotype.

The exploratory nonlinear projections, clustering, and classification analyses further emphasized this point. Nonlinear projections showed partial enrichment of ATLL observations in specific regions but substantial overlap remained. Unsupervised clustering did not reproduce the predefined clinical categories and identified only a small subgroup with an extreme phenotype. Supervised models showed comparatively strong discrimination of ATLL but performed poorly in distinguishing healthy controls from asymptomatic carriers. Given the small and imbalanced cohort and the absence of external validation, these analyses should not be interpreted as evidence of diagnostic accuracy. Instead, they reinforce the biological proximity of many asymptomatic carriers and healthy controls and the heterogeneity of the ATLL phenotype.

Several biological mechanisms may underlie the observed redistribution. Expansion of a malignant or premalignant clone with an effector-memory phenotype could directly increase the TEM fraction [14]. Persistent expression of viral regulatory proteins, including Tax and HBZ, may promote proliferation, survival, immune evasion, and altered differentiation of infected cells [46–48]. Chronic antigenic stimulation and inflammatory signaling could also drive differentiation of non-malignant cells or modify the trafficking and persistence of circulating memory populations [37, 49]. In addition, selective tissue migration may reduce the circulating representation of some subsets while enriching others in peripheral blood [2].

Because the current panel did not distinguish infected from uninfected cells or malignant from non-malignant populations, these mechanisms cannot be separated. The observed phenotype likely

represents a composite of the leukemic clone, other infected clones, residual uninfected immunity, and host responses to the disease [50]. This limitation also means that relative subset frequencies cannot establish the cellular origin of ATLL or the developmental sequence leading to transformation. The results are best interpreted as a circulating immunophenotypic signature associated with established disease rather than as direct evidence of lineage progression.

This study has several strengths. It included three clinically meaningful groups spanning uninfected controls, asymptomatic infection, and established ATLL. The asymptomatic cohort was additionally classified using stringent predefined criteria for presumed transmission route, allowing evaluation of a question rarely examined immunophenotypically. The analytical strategy went beyond significance testing by incorporating correction for multiplicity, global and pairwise effect sizes, bootstrap-based confidence intervals, age- and sex-adjusted median regression, correlation analysis, PCA, PERMANOVA, dispersion testing, and exploratory classification.

The laboratory procedures also strengthened interpretability. The study used viability discrimination, sequential gating, standardized operational definitions, and fluorescence-minus-one controls for CD45RA, CCR7, and CD95. Inclusion of TSCM alongside naïve, TCM, TEM, and TTE compartments enabled a broader evaluation of the circulating differentiation continuum than analyses limited to conventional naïve and memory categories [11]. Complete immunophenotypic data were available for all 117 participants, avoiding loss of observations in the primary analyses.

Several limitations must nevertheless be acknowledged. First, the cross-sectional design prevents determination of temporality and causality. It remains unknown whether the identified profile precedes ATLL, develops during subclinical clonal evolution, or results from established malignancy. Second, the number of patients with ATLL was limited, particularly after subdivision into clinical forms, and the group remained biologically heterogeneous [34]. The study was not powered to compare all four Shimoyama subtypes separately or establish prognostic relationships.

Third, four ATLL samples were collected after treatment initiation or during relapse. Chemotherapy, corticosteroids, zidovudine, and other treatments may modify circulating lymphocyte composition. Although most ATLL samples were obtained before treatment, the small cohort precluded adequately powered treatment-restricted sensitivity analyses. Fourth, healthy controls were descriptively younger and differed in sample source. Samples from infected participants were obtained from peripheral blood collected in EDTA tubes, whereas control cells were recovered from leukoreduction filters after platelet apheresis. Differences in collection, transit time, cellular recovery, and processing may have introduced pre-analytical variation [51].

Fifth, the analyses were based primarily on relative frequencies rather than absolute subset counts. A reduction in one proportional compartment may result partly from expansion of another particularly in samples containing large leukemic populations [52]. Absolute subset quantification would help distinguish true numerical depletion from compositional redistribution. Sixth, the study did not directly

quantify proviral load, clone abundance, integration sites, T-cell receptor repertoire, Tax or HBZ expression, or the proportion of infected cells within each subset.

These molecular measurements are particularly important because individuals with similar total subset frequencies may differ substantially in infection burden and clonality [50]. Previous studies have shown that HTLV-1-infected-cell clonality may inform risk of ATLL development and progression [35, 53–55] and that the proportion of CD4 + CADM1+ cells may be associated with progression among carriers and patients with indolent ATLL [56]. The present immunophenotypic profile should therefore not yet be regarded as a diagnostic, prognostic, or predictive biomarker.

Future studies should prioritize longitudinal follow-up of asymptomatic carriers and patients with indolent ATLL to determine whether naïve-cell contraction, TCM reduction, or TEM expansion precedes clinical progression [22]. Multicenter validation will be essential to assess reproducibility across different populations, sample-processing protocols, and treatment settings. Age- and sex-matched controls and standardized peripheral-blood collection across all groups would reduce demographic and pre-analytical uncertainty.

Subset-resolved molecular analyses represent the most important next step. Sorting of naïve, TSCM, TCM, TEM, and TTE cells followed by proviral-load measurement, integration-site analysis, T-cell receptor sequencing, and genomic or single-cell transcriptomic profiling could determine which compartments contain the dominant infected and malignant clones [23]. Integration with CADM1/CD7 profiles, soluble interleukin-2 receptor concentrations, clonal measurements, and established clinical biomarkers may clarify whether the accessible flow-cytometric pattern adds independent value to multimodal risk assessment [54, 56–58].

CONCLUSIONS

ATLL was associated with a marked redistribution of circulating CD4 + T-cell differentiation states, characterized by contraction of the naïve and central memory compartments and expansion of effector memory cells. These differences persisted after adjustment for age and sex, whereas total CD4 + T-cell frequency was not a robust independent discriminator. T memory stem cell and terminally differentiated effector T-cell frequencies were comparatively preserved, and no distinct immunophenotypic profile was identified according to presumed HTLV-1 transmission route among asymptomatic carriers. These findings support further longitudinal and subset-resolved molecular studies to determine whether circulating CD4 + T-cell architecture contributes to the biological characterization and risk stratification of HTLV-1 infection.

Abbreviations

ATLL

Adult T-cell leukemia/lymphoma

CCR7

C-C chemokine receptor type 7

CHOP

Cyclophosphamide, doxorubicin, vincristine, and prednisone

CI

Confidence interval

FDR

False discovery rate

FMO

Fluorescence-minus-one

HBZ

HTLV-1 basic leucine zipper factor

HCFMUSP

Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo

HTLV-1

Human T-cell lymphotropic virus type 1

IQR

Interquartile range

LDH

Lactate dehydrogenase

PBMC

Peripheral blood mononuclear cell

PCA

Principal component analysis

PCR

Polymerase chain reaction

PERMANOVA

Permutational multivariate analysis of variance

REDCap

Research Electronic Data Capture

ROC

Receiver operating characteristic

TCM

Central memory T cells

TEM

Effector memory T cells

TEMRA

Terminally differentiated effector memory T cells re-expressing CD45RA

t-SNE

t-Distributed stochastic neighbor embedding

TSCM

T memory stem cells

TTE

Terminally differentiated effector T cells

UMAP

Uniform manifold approximation and projection.

Declarations

Ethics approval and consent to participate

This study was approved by the Research Ethics Committee of Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, under protocol CAAE 60994322.4.0000.0068. All procedures involving human participants were conducted in accordance with the Declaration of Helsinki and applicable Brazilian ethical and regulatory requirements. Written informed consent was obtained from all participants before enrollment, collection of biological samples, and access to clinical and laboratory data.

The study was approved by the Research Ethics Committee of Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo under protocol CAAE 60994322.4.0000.0068.

Consent for publication

Not applicable. The manuscript does not contain identifiable individual-level information, images, or clinical details requiring specific consent for publication.

Competing interests

The authors declare that they have no competing interests. The views expressed in this article are those of the authors and do not necessarily represent the official position of the Brazilian Health Regulatory Agency (ANVISA) or any other affiliated institution.

Author Contribution

JBC and PLB contributed equally to this work. JBC, PLB, LAPCL, JP, COR, and HFC conceived and designed the study. JBC, PLB, DFJ, GSG, KSM, ABR, VCC, LR, GP, JAP, SA, VF, NCA, SC, LSF, LMA, SOGM, CR, JSRC, VR, YN, LAPCL, JP, COR, and HFC contributed to participant recruitment, clinical characterization, sample collection, laboratory procedures, data acquisition, or interpretation of the findings. JBC, PLB, DFJ, GSG, KSM, ABR, LSF, LMA, and HFC performed or supervised sample processing and flow-cytometry experiments. JBC, PLB, COR, and HFC curated the data. COR, LR, GP and SA performed statistical analyses and reproducible reporting, with supervision from HFC. COR, LR, GP, SA and HFC prepared the figures and drafted the manuscript. JP, COR, and HFC supervised the study. All

authors critically reviewed the manuscript for important intellectual content, approved the final version, and agreed to be accountable for all aspects of the work.

Acknowledgement

The authors acknowledge the Graduate Program in Medical Sciences of Faculdade de Medicina da Universidade de São Paulo for its academic support. The authors also acknowledge Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP) and the Laboratory of Pathogenesis and Targeted Therapy in Onco-Immuno-Hematology (LIM-31), São Paulo, Brazil, for providing access to clinical samples, research infrastructure, and technical support essential to the conduct of this study. The authors thank the staff of the HTLV-1 Outpatient Clinic, the Flow Cytometry Facility, and Fundação Pró-Sangue Hemocentro de São Paulo for their assistance with participant recruitment, sample collection, laboratory procedures, and technical activities. The authors also acknowledge the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for their support of scientific training, research development, and institutional research activities. The authors are especially grateful to all participants who generously contributed to this study.

Data Availability

The datasets generated and analyzed during the current study are not publicly available because they contain clinical and immunophenotypic data from participants with a rare hematological malignancy, which could increase the risk of re-identification. De-identified data may be made available by the corresponding author upon reasonable request, subject to approval by the responsible institution and Research Ethics Committee and in accordance with applicable ethical and data-protection requirements.

References

1. Szabo PA, Levitin HM, Miron M, Snyder ME, Senda T, Yuan J, et al. Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat Commun*. 2019;10:4706. <https://doi.org/10.1038/s41467-019-12464-3>.
2. Lam N, Lee Y, Farber DL. A guide to adaptive immune memory. *Nat Rev Immunol*. 2024;24:810–29. <https://doi.org/10.1038/s41577-024-01040-6>.
3. Cano-Gamez K, Soskic B, Roumeliotis TI, So E, Smyth DJ, Baldrighi M, et al. Single-cell transcriptomics identifies an effectorness gradient shaping the response of CD4 + T cells to cytokines. *Nat Commun*. 2020;11:1801. <https://doi.org/10.1038/s41467-020-15543-y>.
4. Osum KC, Jenkins MK. Toward a general model of CD4 + T cell subset specification and memory cell formation. *Immunity*. 2023;56:475–84. <https://doi.org/10.1016/j.immuni.2023.02.010>.

5. Wells SB, Rainbow DB, Mark M, Szabo PA, Ergen C, Caron DP, et al. Multimodal profiling reveals tissue-directed signatures of human immune cells altered with age. *Nat Immunol.* 2025;26:1612–25. <https://doi.org/10.1038/s41590-025-02241-4>.
6. Masopust D, Awasthi A, Bosselut R, Brooks DG, Buggert M, Chamoto K, et al. Guidelines for T cell nomenclature. *Nat Rev Immunol.* 2026;26:298–313. <https://doi.org/10.1038/s41577-025-01238-2>.
7. Terekhova M, Swain A, Bohacova P, Aladyeva E, Arthur L, Laha A, et al. Single-cell atlas of healthy human blood unveils age-related loss of NKG2C+GZMB–CD8 + memory T cells and accumulation of type 2 memory T cells. *Immunity.* 2023;56:2836–e28549. <https://doi.org/10.1016/j.immuni.2023.10.013>.
8. Jameson SC, Hamilton SE. Functions and features of circulating memory T cells. *Immunity.* 2026;59:847–62. <https://doi.org/10.1016/j.immuni.2026.03.014>.
9. Truong K-L, Schlickeiser S, Vogt K, Boës D, Stanko K, Appelt C, et al. Killer-like receptors and GPR56 progressive expression defines cytokine production of human CD4 + memory T cells. *Nat Commun.* 2019;10:2263. <https://doi.org/10.1038/s41467-019-10018-1>.
10. Gattinoni L, Lugli E, Ji Y, Pos Z, Paulos CM, Quigley MF, et al. A human memory T cell subset with stem cell–like properties. *Nat Med.* 2011;17:1290–7. <https://doi.org/10.1038/nm.2446>.
11. Kared H, Tan SW, Lau MC, Chevrier M, Tan C, How W, et al. Immunological history governs human stem cell memory CD4 heterogeneity via the Wnt signaling pathway. *Nat Commun.* 2020;11:821. <https://doi.org/10.1038/s41467-020-14442-6>.
12. Sun M, Phan JM, Kieswetter NS, Huang H, Yu KKQ, Smith MT, et al. Specific CD4 + T cell phenotypes associate with bacterial control in people who ‘resist’ infection with *Mycobacterium tuberculosis*. *Nat Immunol.* 2024;25:1411–21. <https://doi.org/10.1038/s41590-024-01897-8>.
13. Nagai Y, Kawahara M, Hishizawa M, Shimazu Y, Sugino N, Fujii S, et al. T memory stem cells are the hierarchical apex of adult T-cell leukemia. *Blood.* 2015;125:3527–35. <https://doi.org/10.1182/blood-2014-10-607465>.
14. Bangham CRM. HTLV-1 persistence and the oncogenesis of adult T cell leukemia/lymphoma. *Blood.* 2023. <https://doi.org/10.1182/blood.2022019332>.
15. Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC. 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.* 1980;77:7415–9. <https://doi.org/10.1073/pnas.77.12.7415>
16. World Health Organization. Human T-lymphotropic virus-1. World Health Organization 2025. <https://www.who.int/news-room/fact-sheets/detail/human-t-lymphotropic-virus-type-1> (accessed July 17, 2026).
17. Weterings DA, Rowan AG, Cook LB. Immunological aspects of HTLV-1 persistence; for the prevention and treatment of Adult T-cell leukaemia-lymphoma (ATL). *Leuk Res.* 2025;148:107635. <https://doi.org/10.1016/j.leukres.2024.107635>.

18. Nunes da Silva A, Araújo THA, Boa-Sorte N, Farias G, Galvão-Barroso AK, de Carvalho A, et al. Epidemiological and molecular evidence of intrafamilial transmission through sexual and vertical routes in Bahia, the state with the highest prevalence of HTLV-1 in Brazil. *PLoS Negl Trop Dis*. 2023;17:e0011005. <https://doi.org/10.1371/journal.pntd.0011005>.
19. Rosadas C, Menezes MLB, Galvão-Castro B, Assone T, Miranda AE, Aragón MG, et al. Blocking HTLV-1/2 silent transmission in Brazil: Current public health policies and proposal for additional strategies. *PLoS Negl Trop Dis*. 2021;15:e0009717. <https://doi.org/10.1371/journal.pntd.0009717>.
20. Vieira BA, Bidinotto AB, Dartora WJ, Pedrotti LG, de Oliveira VM, Wendland EM. Prevalence of human T-lymphotropic virus type 1 and 2 (HTLV-1/-2) infection in pregnant women in Brazil: a systematic review and meta-analysis. *Sci Rep*. 2021;11:15367. <https://doi.org/10.1038/s41598-021-94934-7>.
21. Hleihel R, Skayneh H, de Thé H, Hermine O, Bazarbachi A. Primary cells from patients with adult T cell leukemia/lymphoma depend on HTLV-1 Tax expression for NF- κ B activation and survival. *Blood Cancer J*. 2023;13:67. <https://doi.org/10.1038/s41408-023-00841-7>.
22. Rowan AG, Dillon R, Witkover A, Melamed A, Demontis M-A, Gillet NA, et al. Evolution of retrovirus-infected premalignant T-cell clones prior to adult T-cell leukemia/lymphoma diagnosis. *Blood*. 2020;135:2023–32. <https://doi.org/10.1182/blood.2019002665>.
23. Yamagishi M, Kubokawa M, Kuze Y, Suzuki A, Yokomizo A, Kobayashi S, et al. Chronological genome and single-cell transcriptome integration characterizes the evolutionary process of adult T cell leukemia-lymphoma. *Nat Commun*. 2021;12:4821. <https://doi.org/10.1038/s41467-021-25101-9>.
24. Cook LB, Phillips AA. How I treat adult T-cell leukemia/lymphoma. *Blood*. 2021;137:459–70. <https://doi.org/10.1182/blood.2019004045>.
25. Shimoyama M. Diagnostic criteria and classification of clinical subtypes of adult T-cell leukaemia-lymphoma. *Br J Haematol*. 1991;79:428–37. <https://doi.org/10.1111/j.1365-2141.1991.tb08051.x>.
26. Legrand N, McGregor S, Bull R, Bajis S, Valencia BM, Ronnachit A, et al. Clinical and Public Health Implications of Human T-Lymphotropic Virus Type 1 Infection. *Clin Microbiol Rev*. 2022;35. <https://doi.org/10.1128/cmr.00078-21>.
27. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61:344–9. <https://doi.org/10.1016/j.jclinepi.2007.11.008>.
28. Swerdlow SH. WHO classification of tumours of haematopoietic and lymphoid tissues. International Agency for Research on Cancer; 2017.
29. Gansner JM, Rahmani M, Jonsson AH, Fortin BM, Brimah I, Ellis M, et al. Plateletpheresis-associated lymphopenia in frequent platelet donors. *Blood*. 2019;133:605–14. <https://doi.org/10.1182/blood-2018-09-873125>.
30. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational

- research informatics support. *J Biomed Inf.* 2009;42:377–81.
<https://doi.org/10.1016/j.jbi.2008.08.010>.
31. Maecker HT, Trotter J. Flow cytometry controls, instrument setup, and the determination of positivity. *Cytometry Part A.* 2006;69A:1037–42. <https://doi.org/10.1002/cyto.a.20333>.
 32. Wang F, Cheng F, Zheng F. Stem cell like memory T cells: A new paradigm in cancer immunotherapy. *Clin Immunol.* 2022;241:109078. <https://doi.org/10.1016/j.clim.2022.109078>.
 33. World Medical Association Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Participants. Published Online: October 19. 2024. 2025;333;(1):71–74.
[10.1001/jama.2024.21972](https://doi.org/10.1001/jama.2024.21972)
 34. Kogure Y, Kameda T, Koya J, Yoshimitsu M, Nosaka K, Yasunaga J, et al. Whole-genome landscape of adult T-cell leukemia/lymphoma. *Blood.* 2022;139:967–82.
<https://doi.org/10.1182/blood.2021013568>.
 35. Firouzi S, Farmanbar A, Nakai K, Iwanaga M, Uchimar K, Utsunomiya A, et al. Clonality of HTLV-1–infected T cells as a risk indicator for development and progression of adult T-cell leukemia. *Blood Adv.* 2017;1:1195–205. <https://doi.org/10.1182/bloodadvances.2017005900>.
 36. Kataoka K, Nagata Y, Kitanaka A, Shiraishi Y, Shimamura T, Yasunaga J, et al. Integrated molecular analysis of adult T cell leukemia/lymphoma. *Nat Genet.* 2015;47:1304–15.
<https://doi.org/10.1038/ng.3415>.
 37. Saeed Z, Rowan A, Greiller C, Taylor GP, Pollock KM. Enhanced T-Cell Maturation and Monocyte Aggregation Are Features of Cellular Inflammation in Human T-Lymphotropic Virus Type 1–Associated Myelopathy. *Clin Infect Dis.* 2020;70:1326–35. <https://doi.org/10.1093/cid/ciz369>.
 38. Sugata K, Rahman A, Niimura K, Monde K, Ueno T, Rajib SA, et al. Intragenic viral silencer element regulates HTLV-1 latency via RUNX complex recruitment. *Nat Microbiol.* 2025;10:1447–62.
<https://doi.org/10.1038/s41564-025-02006-7>.
 39. Liang Z, Dong X, Zhang Z, Zhang Q, Zhao Y. Age-related thymic involution: Mechanisms and functional impact. *Aging Cell.* 2022;21. <https://doi.org/10.1111/accel.13671>.
 40. Mittelbrunn M, Kroemer G. Hallmarks of T cell aging. *Nat Immunol.* 2021;22:687–98.
<https://doi.org/10.1038/s41590-021-00927-z>.
 41. Velardi E, Tsai JJ, van den Brink MRM. T cell regeneration after immunological injury. *Nat Rev Immunol.* 2021;21:277–91. <https://doi.org/10.1038/s41577-020-00457-z>.
 42. Sallusto F, Lenig D, Förster R, Lipp M, Lanzavecchia A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature.* 1999;401:708–12.
<https://doi.org/10.1038/44385>.
 43. Jimbo K, Yamagishi M, Suzuki Y, Suzuki K, Mizukami M, Yokoyama K, et al. Progression of adult T-cell leukemia/lymphoma from smoldering to acute type due to branched subclonal evolution. *EJHaem.* 2023;4:1188–90. <https://doi.org/10.1002/jha2.776>.
 44. Kameda T, Kataoka K, Kamiunten A, Hidaka M, Miyoshi H, Nakano N, et al. Integrated genetic and clinical prognostic factors for aggressive adult T-cell leukemia/lymphoma. *Haematologica.*

- 2023;108:2178–91. <https://doi.org/10.3324/haematol.2022.281510>.
45. Anderson MJ, Walsh DCI. PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? *Ecol Monogr*. 2013;83:557–74. <https://doi.org/10.1890/12-2010.1>.
46. Mohanty S, Suklabaidya S, Lavorgna A, Ueno T, Fujisawa J, Ngouth N, et al. The tyrosine kinase KDR is essential for the survival of HTLV-1-infected T cells by stabilizing the Tax oncoprotein. *Nat Commun*. 2024;15:5380. <https://doi.org/10.1038/s41467-024-49737-5>.
47. Yuan X, Yue Y, Chen L, Liu D, Liang Y, Lu C, et al. HTLV-1 subverts the innate immune effector gene IRF7 by viral HBZ protein for oncogenesis. *Nat Commun*. 2026;17:6321. <https://doi.org/10.1038/s41467-026-72980-x>.
48. Kameda T, Shide K, Kamiunten A, Kogure Y, Morishita D, Koya J, et al. CARD11 mutation and HBZ expression induce lymphoproliferative disease and adult T-cell leukemia/lymphoma. *Commun Biol*. 2022;5:1309. <https://doi.org/10.1038/s42003-022-04284-x>.
49. Maher AK, Aristodemou A, Giang N, Tanaka Y, Bangham CRM, Taylor GP, et al. HTLV-1 induces an inflammatory CD4 + CD8+ T cell population in HTLV-1–associated myelopathy. *JCI Insight*. 2024;9. <https://doi.org/10.1172/jci.insight.173738>.
50. Tan BJY, Sugata K, Reda O, Matsuo M, Uchiyama K, Miyazato P, et al. HTLV-1 infection promotes excessive T cell activation and transformation into adult T cell leukemia/lymphoma. *J Clin Invest*. 2021;131. <https://doi.org/10.1172/JCI150472>.
51. Six KR, Vertongen S, Seghers S, De Bleser D, Compennolle V, Feys HB. Differential composition and yield of leukocytes isolated from various blood component leukoreduction filters. *J Immunol Methods*. 2024;533:113733. <https://doi.org/10.1016/j.jim.2024.113733>.
52. Bergstedt J, Azzou SAK, Tsuo K, Jaquaniello A, Urrutia A, Rotival M, et al. The immune factors driving DNA methylation variation in human blood. *Nat Commun*. 2022;13:5895. <https://doi.org/10.1038/s41467-022-33511-6>.
53. Wolf SN, Haddow J, Greiller C, Taylor GP, Cook LBM, Rowan AG. Quantification of T cell clonality in human T cell leukaemia virus type-1 carriers can detect the development of adult T cell leukaemia early. *Blood Cancer J*. 2021;11:66. <https://doi.org/10.1038/s41408-021-00458-8>.
54. Saito M, Sato T, Hasegawa H, Tanabe K, Yoshida-Furihata H, Horibe E et al. Clonality analysis of HTLV-1-infected cells enhances prognostic precision in smoldering adult T-cell leukemia/lymphoma. *Haematologica* 2025. <https://doi.org/10.3324/haematol.2025.287521>
55. Nagasaka M, Yamagishi M, Yagishita N, Araya N, Kobayashi S, Makiyama J et al. Mortality and risk of progression to adult T cell leukemia/lymphoma in HTLV-1–associated myelopathy/tropical spastic paraparesis. *Proceedings of the National Academy of Sciences*. 2020;117:11685–91. <https://doi.org/10.1073/pnas.1920346117>
56. Makiyama J, Kobayashi S, Watanabe E, Ishigaki T, Kawamata T, Nakashima M, et al. CD4 + CADM1 + cell percentage predicts disease progression in HTLV-1 carriers and indolent adult T-cell leukemia/lymphoma. *Cancer Sci*. 2019;110:3746–53. <https://doi.org/10.1111/cas.14219>.

57. Nakamura H, Watanabe T, Sato A, Kawaguchi A, Uchimaru K, Satou Y, et al. A comprehensive assessment using multiple factors based on HAS-Flow analysis predicts ATL development and progression. *Sci Rep.* 2025;15:26618. <https://doi.org/10.1038/s41598-025-10822-4>.
58. Jimbo K, Kawamata T, Inamoto Y, Ito A, Yokoyama K, Sato A, et al. Flow cytometric profiles with CD7 and CADM1 in CD4 + T cells are promising indicators for prognosis of aggressive ATL. *Blood Adv.* 2024;8:3760–70. <https://doi.org/10.1182/bloodadvances.2024013089>.

Figures

Study design and analytical workflow

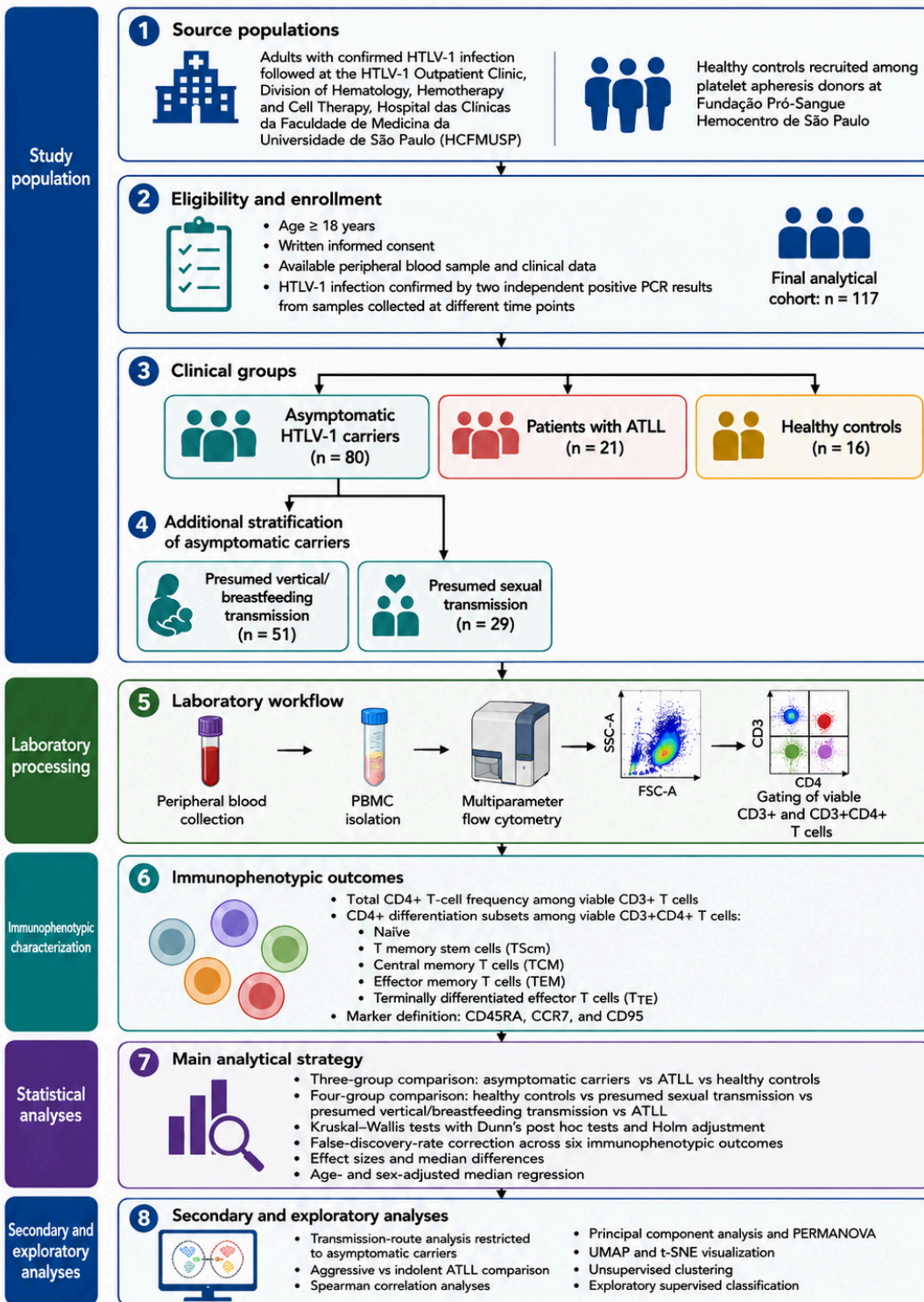


Figure 1

Study design and analytical workflow. Schematic overview of the study population, laboratory procedures, immunophenotypic characterization, and analytical strategy. Adults with confirmed human T-cell lymphotropic virus type 1 (HTLV-1) infection were recruited from the institutional outpatient clinic, whereas healthy controls were recruited among platelet apheresis donors. The final analytical cohort comprised 117 participants: 80 asymptomatic HTLV-1 carriers, 21 patients with adult T-cell

leukemia/lymphoma (ATLL), and 16 healthy controls. Asymptomatic carriers were further classified according to presumed vertical/breastfeeding transmission (n = 51) or presumed sexual transmission (n = 29). Peripheral blood mononuclear cells (PBMCs) were isolated and analyzed by multiparameter flow cytometry. Total CD4⁺ T-cell frequency was measured among viable CD3⁺ T cells, and naïve, T memory stem cells (TSCM), central memory T cells (TCM), effector memory T cells (TEM), and terminally differentiated effector T cells (TTE) were identified within viable CD3⁺CD4⁺ T cells using CD45RA, CCR7, and CD95. The main analyses included three- and four-group comparisons, correction for multiple testing, effect-size estimation, and age- and sex-adjusted median regression. Secondary and exploratory analyses included comparisons by presumed transmission route and ATLL clinical behavior, correlation analyses, principal component analysis, permutational multivariate analysis of variance, nonlinear projections, unsupervised clustering, and supervised classification.

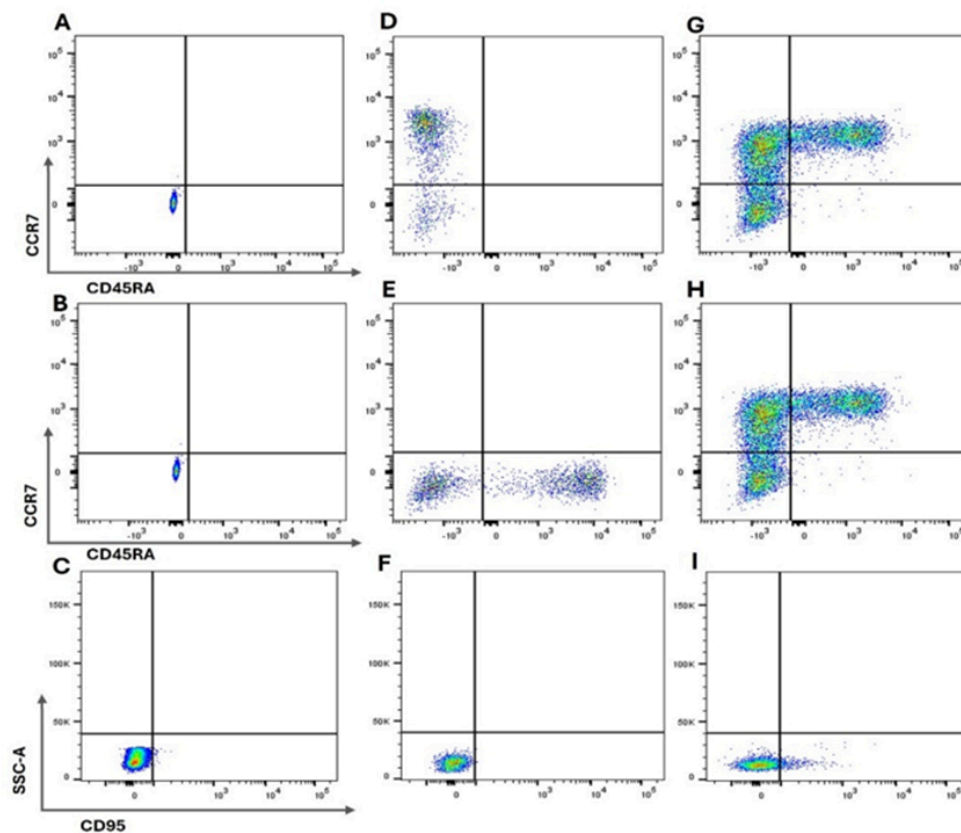


Figure 2

Fluorescence-minus-one controls. FMO control for anti-CD45RA-APC, anti-CCR7-ECD and anti-CD95-BB515 markers. Application of the FMO technique to define the boundaries of the positive population. **A**, **B**, and **C**, Unstained controls; **D**, FMO control without the anti-CD45RA-APC marker; **E**, FMO control without the anti-CCR7-ECD marker; **F**, FMO control without the anti-CD95-BB515 marker; **G**, **H**, and **I**, Dot plots representing a stained sample with the full panel. Analyses were performed using FlowJo v10.10 software (Becton Dickinson, USA).

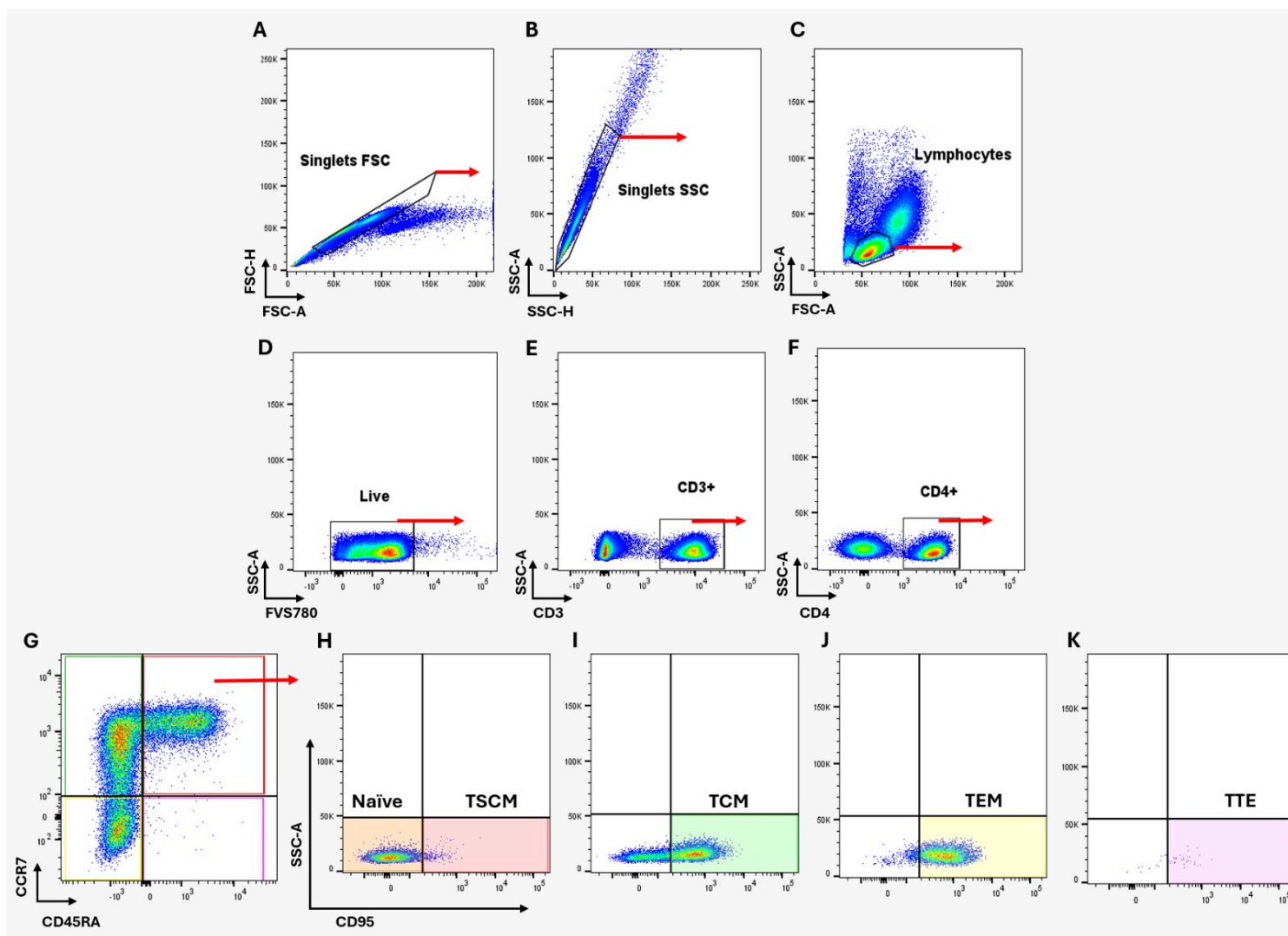


Figure 3

Representative flow-cytometry gating strategy for circulating CD4⁺ T-cell subsets. Representative sequential gating strategy used for immunophenotypic analysis. Cellular aggregates and doublets were excluded using forward-scatter height versus area and side-scatter area versus height parameters. Lymphocytes were identified according to forward- and side-scatter characteristics, followed by exclusion of non-viable cells. CD3⁺ T cells were selected within viable lymphocytes, and CD4⁺ T cells were subsequently identified within the viable CD3⁺ population. CD4⁺ differentiation subsets were defined within viable CD3⁺CD4⁺ T cells according to CD45RA, CCR7, and CD95 expression as naïve cells, CD45RA⁺CCR7⁺CD95⁻; T memory stem cells (TSCM), CD45RA⁺CCR7⁺CD95⁺; central memory T cells (TCM), CD45RA⁻CCR7⁺CD95⁺; effector memory T cells (TEM), CD45RA⁻CCR7⁻CD95⁺; and terminally differentiated effector T cells (TTE), CD45RA⁺CCR7⁻CD95⁺. Arrows indicate the sequentially gated populations used in subsequent analyses.

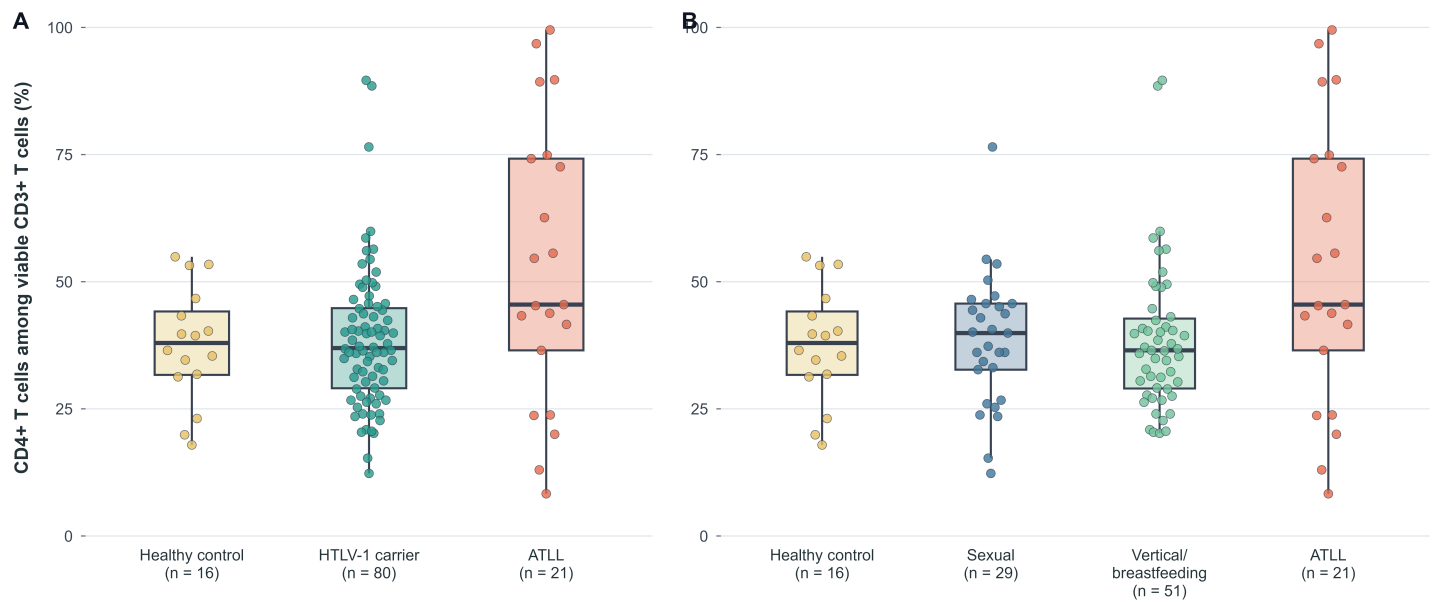


Figure 4

Circulating CD4+ T-cell frequencies across clinical groups and presumed HTLV-1 transmission routes. **A**, Frequency of circulating CD4+ T cells among asymptomatic HTLV-1 carriers, patients with adult T-cell leukemia/lymphoma, and healthy controls. **B**, Frequency of circulating CD4+ T cells after stratification of asymptomatic carriers according to presumed sexual or presumed vertical/breastfeeding transmission. Violin plots show the distribution of individual observations, and internal boxplots indicate the median and interquartile range. Overall comparisons were performed using the Kruskal–Wallis test. When applicable, pairwise comparisons were performed using Dunn’s test with Holm adjustment. Displayed pairwise values correspond to Holm-adjusted p values. ATLL, adult T-cell leukemia/lymphoma; HTLV-1, human T-cell lymphotropic virus type 1.

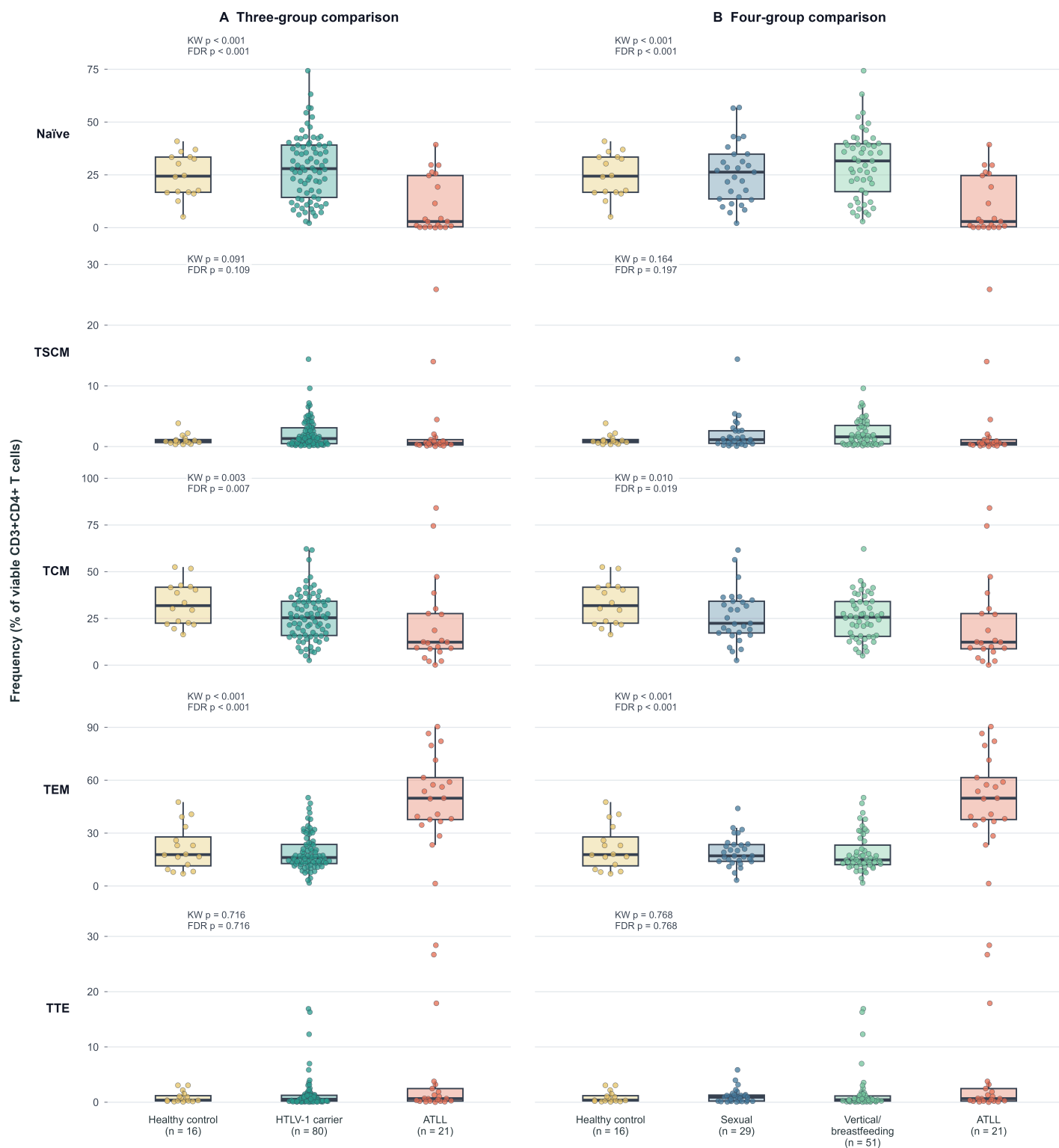


Figure 5

Circulating CD4+ T-cell subsets across the differentiation continuum. Frequencies of naïve, T memory stem cells (TSCM), central memory T cells (TCM), effector memory T cells (TEM), and terminally differentiated effector T cells (TTE) among viable CD3+CD4+ T cells. **A**, Comparisons among healthy controls, asymptomatic HTLV-1 carriers, and patients with adult T-cell leukemia/lymphoma. **B**, Comparisons after stratification of asymptomatic carriers according to presumed sexual or presumed

vertical/breastfeeding transmission. Boxplots indicate the median and interquartile range, and points represent individual participants. Overall *p* values were obtained using the Kruskal–Wallis test and adjusted across the six prespecified immunophenotypic outcomes using the Benjamini–Hochberg false-discovery-rate procedure. Significant pairwise comparisons were identified using Dunn’s test with Holm adjustment. ATLL, adult T-cell leukemia/lymphoma; HTLV-1, human T-cell lymphotropic virus type 1.

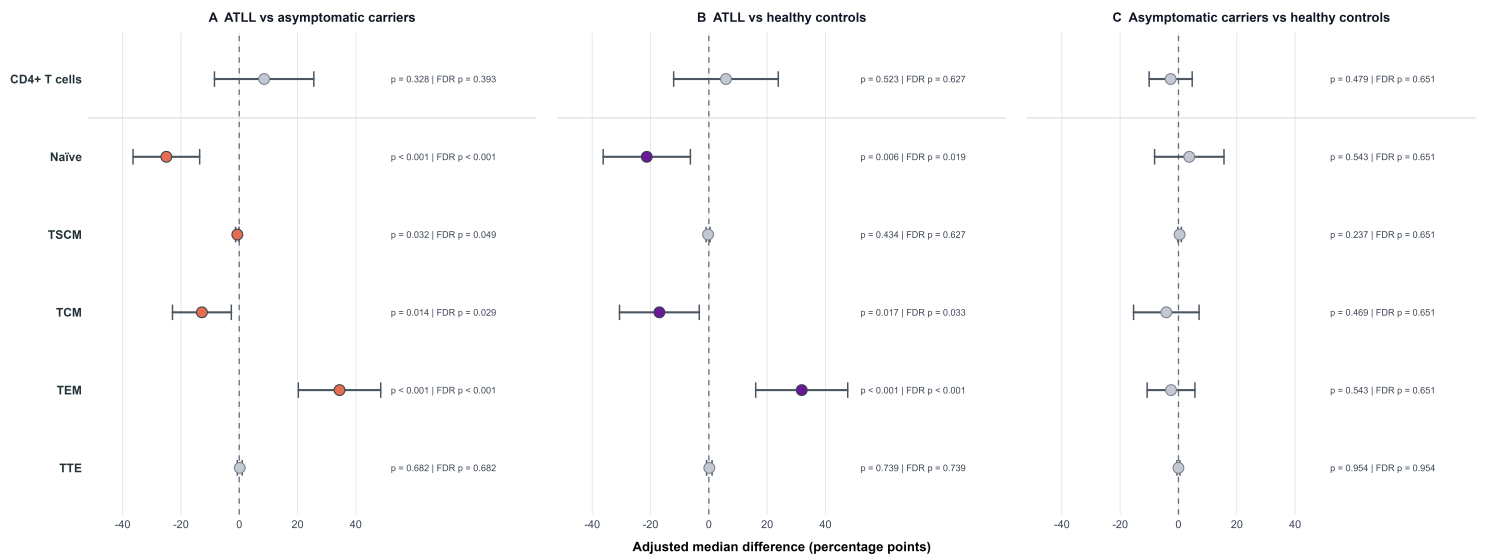


Figure 6

Age- and sex-adjusted differences in circulating CD4+ T-cell frequencies according to clinical groups. Forest plots show adjusted median differences and 95% confidence intervals derived from median regression models evaluating total circulating CD4+ T-cell frequency and the five operationally defined differentiation subsets. **A**, Patients with ATLL compared with asymptomatic HTLV-1 carriers. **B**, Patients with ATLL compared with healthy controls. **C**, Asymptomatic HTLV-1 carriers compared with healthy controls. Estimates represent the adjusted median frequency in the first-named group minus that in the reference group. Negative estimates indicate lower adjusted frequencies in the first-named group, whereas positive estimates indicate higher adjusted frequencies. Points represent adjusted median differences, horizontal lines indicate 95% confidence intervals, and the vertical dashed line at zero denotes no adjusted difference. Filled points indicate associations that remained statistically significant after FDR correction. Raw and FDR-adjusted *p* values are displayed for each estimate. Models were adjusted for age and sex. Total CD4+ T-cell frequency was expressed as a percentage of viable CD3+ T cells, whereas naïve, TSCM, TCM, TEM, and TTE frequencies were expressed as percentages of viable CD3+CD4+ T cells. ATLL, adult T-cell leukemia/lymphoma; CI, confidence interval; FDR, false discovery rate; HTLV-1, human T-cell lymphotropic virus type 1; TCM, central memory T cells; TEM, effector memory T cells; TSCM, T memory stem cells; TTE, terminally differentiated effector T cells.

Clinical severity, correlation structure, and multivariate organization of circulating CD4+ T-cell profiles

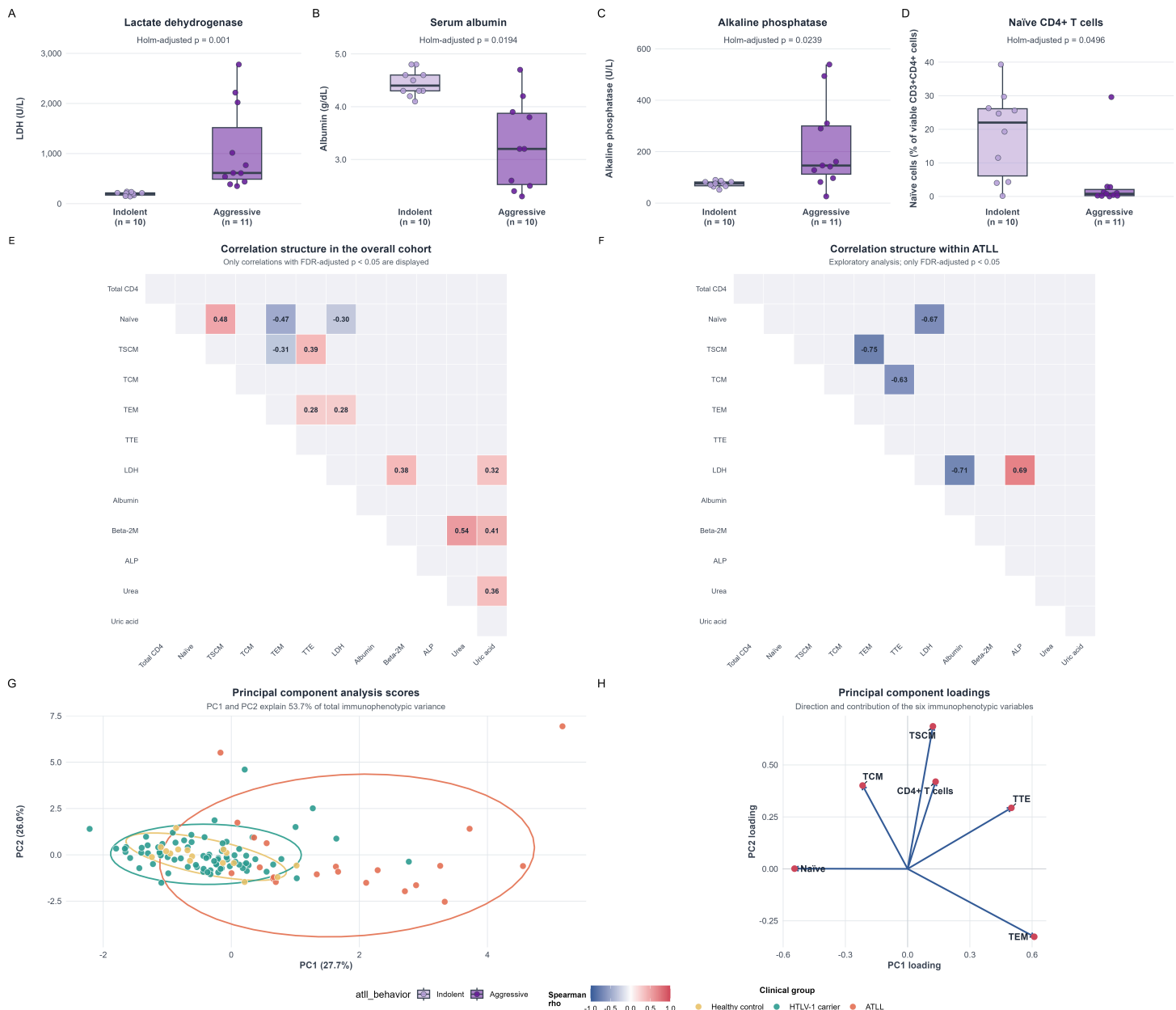


Figure 7

Clinical severity, correlation structure, and multivariate organization of circulating CD4+ T-cell profiles.

A–D, Clinical, biochemical, and immunophenotypic comparisons between patients with aggressive and indolent adult T-cell leukemia/lymphoma. Boxplots show lactate dehydrogenase concentrations (**A**), serum albumin concentrations (**B**), alkaline phosphatase concentrations (**C**), and naïve CD4+ T-cell frequencies (**D**). Boxes represent the median and interquartile range, whiskers extend to values within 1.5 times the interquartile range, and points represent individual patients. Multiplicity-adjusted p values are displayed. **E**, Spearman correlation matrix for immunophenotypic and laboratory variables in the overall cohort. **F**, Corresponding correlation matrix restricted to patients with adult T-cell leukemia/lymphoma. Cell color and intensity represent the direction and magnitude of Spearman's correlation coefficient, and only associations meeting the prespecified false-discovery-rate threshold are

highlighted or annotated. **G**, Principal component analysis score plot based on the six standardized immunophenotypic variables, according to clinical group. **H**, Principal component loading plot showing the contribution and direction of each immunophenotypic variable to the first two principal components. ATLL, adult T-cell leukemia/lymphoma; LDH, lactate dehydrogenase; PCA, principal component analysis; TCM, central memory T cells; TEM, effector memory T cells; TSCM, T memory stem cells; TTE, terminally differentiated effector T cells.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1Supplementarymaterial.docx](#)