High pathogen burden in childhood promotes the development of unconventional innate-like CD8+ T cells

Yves T. Falanga,1,2 Michela Frascoli,2 Yasin Kaymaz,3 Catherine Forconi,2 John Michael Ong’echa,4 Jeffrey A. Bailey,3,5 Leslie J. Berg,1 and Ann M. Moormann2

1Department of Pathology, 2Program in Molecular Medicine, and 3 Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, Massachusetts, USA. 4Center for Global Health Research, Kenya Medical Research Institute, Kisumu, Kenya. 5Division of Transfusion Medicine, Department of Medicine, University of Massachusetts Medical School, Worcester, Massachusetts, USA.

Introduction

The adaptive immune system has evolved to mount effective and specific immune responses to pathogenic infections that provide protective memory against future exposures (1–4). Although animal models have been essential to decipher key immunological processes, they are limited in their ability to recapitulate the substantial heterogeneity within and between humans that contributes to the balance between protection and pathology (5, 6). Chronic or repeated infections pose a challenge to the host. An overexuberant response may lead to immune pathology, while a weak response may lead to pathogens overwhelming the host. Overall, there are substantial gaps in our understanding of how pathogens, particularly chronic or repeated infections, interplay and affect the development of the human immune system during infancy and early childhood (5, 6). Also, studies of the immune system in children are challenging due to inherent difficulties in collecting samples and the logical limitations imposed on quantity and accessible immune compartments.

Plasmodium falciparum (Pf) malaria infections of children residing in endemic areas initially present as acute, symptomatic infections that evolve into chronic, asymptomatic infections after repeated, high-dose exposures (7, 8). The immunologic transition that results in premunition, also referred to as clinical immunity or immune tolerance, is not fully understood and becomes challenging to deconvolute from immunologic maturation that naturally occurs as children age. Seminal studies have elegantly demonstrated the association of clinical immunity to malaria with a decreased number and function of Vγ2+Vδ2+ cells, a subset of γδ T lymphocytes that does not require antigen processing to induce T cell receptor–mediated (TCR-mediated) cytolytic death of infected red blood cells via granulysin (9–13). These findings support the premise that persistent malaria infection induces immunological adaptation in humans, favoring a less inflammatory environment. Although numerous studies have shown that children’s and adults’ immune systems differ in multiple aspects (5, 6, 14, 15), very little is known about how pathogen exposure affects the transition of naïve CD8+ T cells to effector and effector memory cells in humans.
Chronic parasitic and viral infections have been shown to cause T cell exhaustion, a phenomenon characterized by a hierarchical degradation of pathogen-specific T lymphocytes that lose their ability to mount cytolytic and other effector functions upon TCR stimulation (16–18). Furthermore, mechanisms by which chronically infected individuals with potent and nonexhausted T cells are able to sustain elevated pathogen burden without displaying life-threatening clinical signs remain poorly understood. To better understand the natural progression of T cells from naïve-like to effector-memory subsets and changes in functionality, we prospectively compared two cohorts of age-matched children residing in areas with divergent pathogen burden (notably Pf and Schistosoma mansoni [Sm]) in western Kenya (19–21). We hypothesized that elevated cumulative pathogen burden would lead to T cell exhaustion and a significantly diminished capacity to respond to cognate antigen. Interestingly, and to our surprise, we found that over time, cumulative malaria and other infections promoted the development of unconventional innate-like CD8^+ T cells (Granzyme B^+, IFNγ^+^, TNFα^+^, PLFZ^hi^, ID2^hi^, IKZF2^hi^) reminiscent of NK cells and type-1 innate lymphoid cells (ILC1) (22–24).

**Results**

*Children living in areas of high pathogen burden have elevated proportions of CD3^+CD8^{dim} T cells.* The study’s primary goal was to understand the impact of holoendemic Pf exposure on the development and function of CD4^+ and CD8^+ T cells. However, the high malaria transmission region (Kisumu, Kenya) from which study participants were enrolled also experience higher exposure to other chronic infections such as Sm and Epstein-Barr virus (EBV) compared with the low malaria exposure region (Nandi, Kenya) (19–21, 25). Nandi and Kisumu children are both infected with EBV early in life, and even though EBV is a persistent gammaherpesvirus, Nandi children have lower viral loads during childhood (26). Therefore, the two cohorts were defined as high pathogen burden and low pathogen burden cumulative childhood exposures for Kisumu and Nandi, respectively (Figure 1A and Supplemental Figure 1; supplemental material available online with this article; https://doi.org/10.1172/jci.insight.93814DS1) (19, 25, 26).

In order to evaluate the impact of chronic pathogen exposure on the development and effector function of CD8^+ T cells as children age, we analyzed peripheral blood mononuclear cells (PBMCs) from the same children at approximately 2.5 years (toddlers) and again at 6.5 years of age (school-age). In order to minimize investigator-based bias frequently encountered in flow cytometry gating strategies, in favor of a data-driven approach, we analyzed flow cytometry data with ACCENSE (27) and PhenoGraph (28), two platforms performing dimensionality reduction via t-distributed stochastic neighbor embedding (t-SNE) (29, 30) and partitioning of high-dimensional single cell data into subpopulations (Figure 1, B and C). Using this unsupervised approach on each study participant’s PBMCs revealed partitioning of CD3^+CD8^+ T cells into two groups based on surface expression of the CD8 molecule: CD8^{bright} and CD8^{dim} T cells (Figure 1B). Megacluster analysis of unstained individual samples on a population level using PhenoGraph (28) showed that PBMCs from Kisumu (high pathogen burden) school-age children were significantly enriched in CD8^{dim} clusters (Figure 1C) compared with their age-matched counterparts from Nandi (low pathogen burden).

While both groups displayed comparable proportions of CD8^{dim} T cells during toddlerhood (median 11.2% and 14.6%), we observed that over time, children from Kisumu had significantly higher proportions of CD8^{dim} T cells (46.1%) when they reached school age compared with their age-matched counterparts from Nandi (6.7%) (Figure 2, A and B). There was no significant difference in CD8^{bright} absolute cells count between Nandi and Kisumu school-age study participants (Figure 2B). Both CD8^{bright} and CD8^{dim} T cells displayed comparable survival and viability as previously reported (31–34). While CD8^+ T cell subsets (defined by CCR7 and CD45RA staining) from Nandi and Kisumu children showed no substantial differences in proportions, the majority (~60%) of CD8^{bright} T cells had a naïve phenotype, while over 80% of the CD8^{dim} cells were effector memory T cells (T_{EM}) and effector memory RA T cells (T_{EMRA}) (Figure 2, C and D), as observed previously (31, 32). Interestingly, there were no significant differences in the proportions of CD4^+ T cell functional subsets within our study populations as the children aged (Figure 2D). To our knowledge, our findings provide longitudinal epidemiological and immunological support for the uncharacteristic increase in CD8^{dim} T cell proportions in the context of chronic parasitic infections in humans.

*Elevated parasite-specific antibody titers are associated with increased proportions of CD8^{dim} T cells.* Although our cohorts were initially defined based on malaria transmission intensity, these children also had varied history of exposure to other common infections in this region (19–21, 25). In order to study the
history of past infections within our study participants, we assessed cumulative pathogen burden by measuring antibodies (IgG) directed against select liver- and blood-stage malaria antigens, EBV, and Sm, along with antibodies to vaccine antigens (tetanus toxoid and edmonston measles vaccine virus) (Supplemental Figure 1). Unsupervised clustering of serological data revealed coclustering of school-age children consistent with their geographic origin, suggesting that antibody titers reflect expected cumulative pathogen exposure. In contrast, toddlers displayed greater heterogeneity within study groups that was poorly associated with place of residence and prevalence of infectious diseases characteristic of the region (Figure 3A). This suggests that putative exposures attributed to residing in Kisumu or Nandi, defined as an ecological variable, may not be informative to characterize cumulative exposures.
for children at such a young age, and it suggests that interspersed longitudinal sample collection may miss detection of transient or subpatent infections. Not surprisingly, clusters in school-age children were driven by Pf and EBV antibody titers and were in accordance with previous studies (19, 25). Interestingly, we found that, in school-age children, antibody titers for Pf and Sm were positively correlated with the percentage of CD8 dim T cells (Figure 3B), while antibody titers to EBV antigens, measles vaccine virus, or tetanus toxoid were not.

Figure 2. Cumulative pathogen burden promotes the development of CD3+ CD8 dim T cells. (A) Gating strategy and representative bivariate plot of unstimulated PBMCs from Nandi and Kisumu. FSC, forward scatter. (B) Boxplot (median and 95% interquartile range [IQR]) representing the proportion and the absolute count of CD8 dim T cells in Nandi (n=14) and Kisumu (n=15) children. Black dots on boxplot represent individual patients. Outlier values are indicated with red dots. The median proportion of CD8 dim T cells (percent CD8 T cells) were as follows: Nandi = 6.7%, Kisumu = 46.1% with a significance level ***P < 0.001, ****P < 0.0001 (two-tailed unpaired t test with Welch’s correction). (C) Representative bivariate plot displaying flow cytometry gating of CD3+ CD8+ T cell functional subsets. (D) Pie charts showing the proportion of CD8+ and CD4+ T cell subsets comparing the same children over time from Nandi and Kisumu. Data accumulated from 9 independent experiments, n = 14 (Nandi), n = 15 (Kisumu). The proportion of T cell subsets are different between CD8 bright and CD8 dim but not over time (Welch’s t test).
Children living in areas of high pathogen burden have significantly elevated levels of plasma soluble CD163. Next, we examined the cytokine profiles in these children to determine if cytokine milieu was associated with the increased proportion of CD8dim T cells. We used multiplex panels to measure 51 plasma analytes (IL-1β, IL-1α, IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12(p70), IL-13, IL-15, IL-17, Eotaxin, GFG basic, G-CSF, GM-CSF, IFNγ, IP-10, MCP-1, MIP-1α, PDGF-BB, MIP-1β, RANTES, TNF-α, VEGF, APRIL, BAFF, sCD30, sCD163, Chitinase 3-like 1, gpl130, IFNβ, IL-11, IL-19, IL-20, IL-26, IL-27, IL-28, IL-29, IL-32, IL-34, IL-35, LIGHT, Osteocalcin, Pentraxin-3, sTNF-R1, sTNF-R2, TSLP, TWEAK). After adjusting for multiple comparison on the 51 analytes, we found that the increased proportion of CD8dim T cells was neither associated with a steady-state type 2 polarizing environment as reported in mouse models (32, 35), nor with differential expression of IL-10, IL-12, or TGFβ (data not shown). Instead, we discovered that children living in areas of high pathogen burden had significantly higher levels of soluble CD163 (sCD163) in their plasma (Figure 3C). sCD163 (a hemoglobin-haptoglobin scavenger receptor solely expressed on monocytes and macrophages) is released in the plasma upon activation via TLR ligation and serves as a biomarker for macrophage activation (36–39). Consequently, elevated levels of sCD163 observed in the plasma of Kisumu children, likely due to high pathogen burden (mainly malaria and schistosomiasis), might affect several branches of the immune system during infection, resulting in the generation of atypical cell subsets such as CD8dim T cells.

CD8dim T cells represent a transcriptionally and functionally distinct state of CD8+ T cells. Due to limitations in blood volumes obtainable from children, we decided to first examine transcriptional profiles comparing CD8bright with CD8dim T cell subsets in PBMCs collected from adults with different histories of malaria exposure. CD8dim T cells were found to a greater degree in healthy adult Africans (44.2% median, n = 11) and to a much lesser extent in malaria-naive North Americans (8.2% median, n = 9) (Figure 4A), suggesting that this CD8+ T cell stratification was not restricted to Kenyan children but may represent a distinct cell subset that persists into adulthood after cumulative pathogen exposure. Both CD8bright and CD8dim T cells had comparable viability potential, as only 3% of CD8dim T cells were Caspase 3–positive compared with none within the CD8bright T cell subset (Figure 4B). Analysis of the proliferative potential upon stimulation with anti-CD3 and anti-CD28 antibodies for 72 hours of sorted CD8dim T cells, and CD8bright CD45RA+ and CD45RA– T cells, from an adult individual previously exposed to malaria revealed that CD8dim T cells had stunted proliferation compared with CD8bright T cells (for both CD45RA+ and CD45RA– subsets). Of note, after 96 hours of stimulation, a subset of CD8dim cells appeared to upregulate CD8α expression (Figure 4C), suggesting a contextual plasticity in CD8dim T cells. Furthermore, we observed that CD8dim T cells from Kenyan school-age children and Kenyan adults, as well as from North American adults, were essentially αβ and not γδ T cells with a limited proportion of cells expressing surface CD56 (Supplemental Figure 2).

To better characterize these CD8+ T cell subsets, we performed mRNA sequencing on sorted CD8bright and CD8dim T cells, further gated into CD45RA and CCR7 subsets, collected from 3 healthy adults representing diverse cumulative pathogen-exposure histories. Although all functional subsets of CD8+ T cells shared a set of differentially expressed genes determined after unsupervised clustering analysis (Figure 5A), we found that T_EM and T_EMBRA contained the most differentially expressed genes when stratified as CD8dim or CD8bright (Figure 5, B and C). CD8dim T cells were not only characterized by differentially expressed genes involved in innate-like functions (Figure 6A), but they also were highly enriched in TNFα, NF-κB, and IL-6 signaling pathways when compared with total CD8bright T cells after gene set enrichment analysis (GSEA) (Figure 6B). Altogether, our data show that CD8dim T cells are transcriptionally distinct from CD8bright T cells.

High pathogen burden promotes the development of Granzyme B+IFNγ+TNFα+CD3+CD8dim T cells. In order to validate transcriptional differences between CD8bright and CD8dim T cells subsets, we returned to samples from our Kenyan school-age cohorts and selected significantly differentially expressed genes that had commercially available antibodies for flow cytometry (Figure 7). While CD8bright T cells from high- and low-pathogen toddlers showed comparable proportions of T cells subsets (Figure 2D), CD8dim T cells from high-pathogen toddlers displayed significantly higher proportions of IFNγ, TNFα, and CCL4-producing cells upon in vitro stimulation with staphylococcal enterotoxin B (SEB) (Figure 7A). Interestingly, CD8dim T cells from these high pathogen–burden children (4 years later) developed severely impaired CCL4 and IFNγ production and displayed substantially decreased proportions of functional effector cells in both CD8bright and CD8dim compartments. Strikingly, and contrasting with some murine observations (32), we found that CD8dim T_EMBRA had significantly elevated levels of Granzyme B in steady state, which is reminiscent of ILC1s and type-1 innate-like T cells (ILTC1s) developing in the context of tumor-elicited immune surveillance (24) (Figure 7, B and C).
Figure 3. Children living in areas of elevated pathogen burden develop distinctive serological and plasma cytokine profiles. Serum antibody titers and plasma analytes were measured at a 4-year interval (toddlers to school-age) coinciding with T cell subset assays. Immunity to vaccine antigens, measles, and tetanus were measured as controls (Nandi, n = 33; Kisumu, n = 31). Antibody titers (IgG) specific for Pf (HPR-II, MSP1-FVO, CSP), measles virus (edmonston vaccine strain), Clostridium tetani (tetanus), Schistosoma mansoni (SWAP), and EBV (EAD, ZEBRA, VCA, EBNA1) were measured using multiplex conjugated-bead suspension assay. (A) Heatmap of scaled antibody titers (Z score). Pathogen burden is represented with orange (low, Nandi) and purple (high, Kisumu). Data generated from 1 experiment measuring plasma antibody titers from patients. (B) Dotplots (and 95% CI) representing the association between proportion of CD3+CD8dim T cells and pathogen-specific antibody titers in school-age children. Solid lines represent best-fit regression line and coefficient of determination ($r^2$), and P values are displayed (*P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001). (C) Steady-state plasma sCD163 levels from toddlers and school-age children. Boxplot (median and 95% IQR) displays the relative amount of sCD163 (pg/ml) (Nandi, n = 14; Kisumu, n = 15). Black dots are values from individual children. Two-way ANOVA with Sidak multiple comparison post test was used to analyze statistical significance for the 51 analytes measured in the two groups of divergent pathogen exposure. Data generated from one experiment measuring plasma analytes titers from 29 patients.
CD8dim T cells from school-age children living in areas of high pathogen burden have low NFAT1. In addition to having graded levels of NFAT1 (a key transcription factor involved in TCR signaling) in CD8+ T cells subsets, we found that cumulative pathogen burden negatively affected NFAT1 levels (Figure 8, A and B). In fact, decreased NFAT1 in addition to low surface CD8 expression is consistent with the defect in T cell effector function observed in school-age children from high-pathogen areas. In general, our data show that CD8bright T cells expressed higher levels of CXCR4, ABCB1, CD11a, KLF2, KLRK1, and CD2 compared with CD8dim T cells (Supplemental Figure 3A). Interestingly, the expression levels of some of these markers seemed to be independent of cumulative pathogen burden, as school-age children from both areas displayed varied expression (Supplemental Figure 3B). In addition, we examined transcriptional regulators of T cell differentiation (40) and found that high pathogen–exposure children had severely diminished frequencies of TEM and TEMRA CD8+ T cell subsets coexpressing the transcription factors T-bet and Eomesodermin (Eomes) compared with their age-matched, low pathogen burden counterparts (Figure 9). Taken together, our data demonstrate that parasitic infections during childhood induce transcriptional changes in CD8+ T cells favoring an innate-like circuitry of transcription factors and a pronounced steady state biosynthesis of Granzyme B.

Figure 4. Healthy African adults have greater proportions of CD3+ CD8dim T cells than healthy North American adults. (A) Peripheral blood mononuclear cells (PBMCs) were stained with a panel of fluorescently labeled antibodies to define T cell subsets and then analyzed by multiparameter flow cytometry. Data on the boxplot (median and 95% IQR) display the proportion of CD8dim T cells from healthy Kenyan adults (n = 11) and healthy North American adults (n = 9). Computed two-tailed unpaired t test with Welch’s correction P value is displayed at the bottom of the graph (***P < 0.001). Black dots on the boxplot display values from individual healthy adults. Data generated from 1 flow cytometry experiment. (B) Caspase 3 was detected in ~3% of CD8 dim and not in CD8 bright T cells, indicating that neither cell population is undergoing apoptosis. (C) Representative flow cytometric analysis of CFSE-labeled profile of CD8 Tem and CD8 dim T cells, either CD45RA- or CD45RA+ sorted from the peripheral blood of an adult sample that had previously been exposed to malaria. Cells were analyzed after 72 and 96 hours of stimulation with α-CD3 and α-CD28 antibodies. The number in each plot represents the percentage of cells that proliferated.
Discussion

Using a prospective cohort study of children exposed to different intensities of infectious diseases from 2–7 years of age, we were able to characterize their influence on the development of T cell subsets. We included malaria, Schistosomiasis, and EBV as exposures and observed an expansion of unconventional CD8<sup>dim</sup> T cells displaying innate-like characteristics (Granzyme B<sup>hi</sup>, IFNγ<sup>low</sup>, TNFα<sup>low</sup>, PLFZ<sup>hi</sup>, ID2<sup>hi</sup>, IKZF2<sup>hi</sup>) that develops in otherwise healthy children residing in areas of high pathogen burden. In contrast, longitudinal assessments of age-matched Kenyan children from a highland area with low pathogen burden showed that these children developed a more conventional polyfunctional IFNγ<sup>+</sup>, TNFα<sup>+</sup>, CCL4<sup>+</sup>CD3<sup>+</sup>CD8<sup>bright</sup> T cell profile as they aged. The fact that CD8<sup>dim</sup> T cells persisted into adulthood when there was a history of elevated pathogen burden and that their frequencies were largely associated with a maturing immune system for children in the face of cumulative pathogen burden suggests their potential contribution to protection from disease. Our study participants were in general good health; therefore, we speculate that CD8<sup>dim</sup> T cells may be an adaption consistent with premunition, thereby protecting an individual from immunopathology. However, future studies would be needed to determine if CD8<sup>dim</sup> T cells are directly involved in pathogen clearance or protection from infection, or if they are the result of immune tolerance mechanisms.
Consistent with our findings are reports that CD8+ T cells expressing Granzyme B have been identified as mediators of cerebral malaria in mouse models (41, 42). In contrast to contributing to immunopathology, Granzyme B produced by CD8+ and γδ T cells in experimental, controlled human malaria infections (CHMI) are associated with protection from liver-stage infection (41–44). Within the context of a CHMI study, adults from Tanzania that were preexposed to malaria were unable to produce IFN-γ recall responses from CD8+ and γδ T cells, compared with malaria-naive Dutch adults (45). Even though Granzyme B levels were not measured, IFN-γ levels were low prior to sporozoite challenge.

Although downregulation of CD8 coreceptor is often subsequent to TCR-mediated signaling (46), a significantly high proportion of CD8dim T cells has been observed in African green monkeys (AGM) and was protective against SIV infection (47). Furthermore, in some contexts, CD8dim T cells substituted for CD4+ T cells and...
provided B cell help, leading to antibody production (48, 49). Other studies showed that type 2 polarization of CD8⁺ T cells induced epigenetic modifications on the CD8α locus, resulting in downregulation of CD8 surface expression in mice (32). That phenotype was sustained in vivo following adoptive transfer and was partially reversed in the presence of IFNγ (35). Altogether, our findings and these previous studies suggest that CD8dim T cells are not simply activated CD8⁺ T cells with downregulated CD8 coreceptors, but instead represent a cell state dynamically arising from a specific immunological context.

A limitation to human immunology studies is that conclusions are of necessity derived from circulating PBMCs from living study participants. Recent studies of organ donors elegantly demonstrate that the composition of human immune cells varies with their tissue localization, suggesting that immunological findings resulting from human PBMCs should be interpreted very cautiously (5, 6). Regardless of this understandable criticism, our longitudinal childhood cohort study was able to demonstrate remarkable differences in CD8⁺ T cell subsets associated with prolonged pathogen exposure, notably to malaria and Schistosomiasis. The role of Schistosomiasis in shaping human T cell lineage and fate decisions is speculative at this stage, but Schistosomiasis has been shown to activate CD8⁺ T cells (50). Additionally, the distinctive plasma composition observed in the Kismu compared with the Nandi children displays another layer of complexity generally overlooked in human immunology studies. Generally, very few in vitro assays conducted with human PBMCs replicate

---

**Figure 7. Atypical Granzyme B⁺ IFNγ-low TNFα-low CD3⁺ CD8dim T cells arise in school-age children living in areas of high pathogen burden.** (A) PBMCs were stimulated with Staphylococcus enterotoxin B (SEB) in vitro for 4 hours, and cytokines were measured in Temo and TEMRA. Pie charts generated with SPICE (Simplified Presentation of Incredibly Complex Evaluations) representing qualitative response of CD8bright and CD8dim T cells using Boolean combination of gates identifying TNFα, IFNγ, and CCL4 (MIP-1β). Pie chart colors represent number of effector molecules, while each cytokine is represented by an arc (Nandi, n = 14; Kisuμ n = 15). (B) Representative histograms of Temo cells gated for Granzyme B, IFNγ, or TNFα from unstimulated and PMA/ionomycin-stimulated CD8⁺ T cells (4 hours). (C) Boxplots of computed MFI from B (Nandi, n = 8; Kisuμ, n = 8). Welch's two-tailed t test P values are displayed (*P < 0.05, **P < 0.01, ***P < 0.001). Black dots on boxplot represents individual patients. Outlier values are indicated with red dots. Data generated from five independent experiments.
the impact of plasma cytokine found in the patient when interrogating the function of immune cells ex vivo. Further investigations are warranted to dissect not only what pathogen-induced mediators drive the observed plasma cytokine composition, but also to elucidate the long-term effects of the cytokine microenvironment on immune cell homeostasis and function in these children.

Structural and functional constituents of the human immune system are dynamically regulated in response to developmental and environmental cues (6, 14, 15, 46). Here, we have identified a subpopulation of CD8+ T cells with a distinct transcriptional profile showing significant expansion in individuals repetitively infected with parasites; it may represent a dynamic cell state in CD8+ T cell biology resulting from an adaptation of the immune system in response to persistent pathogen exposure, rather than reflecting a fixed cell type. In fact, although this subpopulation expressed detectable protein levels of CD3 and CD8, their transcriptome and in vitro stimulation signature suggest severely impaired TCR signaling (low CD2, CD8, ITK, NFAT1) in favor of an innate-like circuitry (high PLZF, ID2, IKZF2, Granzyme B) (22–24, 51). Our observations as well as recent publications (22–24) emphasize the study of immune cells as dynamic entities adapting to the microenvironment rather than classify them with a static set of cell surface markers that might merely reflect their status at the time of analysis.

While the chemokine receptor CXCR4 and the transcription factor KLF2 regulate T cell migration and memory formation (52–56), their downregulation is one of the hallmarks of ILC1s and ILTC1s (24). CD8bright and CD8dim T cells are not only metabolically dissimilar as described in murine studies (57, 58), but they may also have significantly distinct migratory and memory potential. Furthermore, while CD8+ TEM and TEMRA cells from school-age children living in Kisumu showed no exhaustion phenotype compared with Nandi children (as determined by Tim-3 expression; Figure 1B and Supplemental Figure 4), they predominantly coexpressed T-bet and Eomes at similar levels regardless of cell subset (Figure 9, C and D), an expression pattern often observed during NK maturation (59). This may suggest that these are short-lived effector cells, which would be consistent with mouse models (60), or alternatively that T-bet and Eomes ratios are not informative to predict CD8+ T cell fate in humans. More in-depth, longitudinal human immunology studies are necessary in order to fully explore transcriptional regulation of T cell differentiation and the relevance to protection from acute and chronic infections.

Figure 8. CD3+ CD8dim T cells from school-age children living in areas of high pathogen burden have low NFAT1. Unstimulated PBMCs from Nandi and Kisumu (low and high pathogen burden areas, respectively) were fluorescently labeled and analyzed with flow cytometry. (A) Representative histograms displaying graded NFAT1 in unstimulated CD8+ T cell subsets. (B) Boxplot (median and 95% IQR) representing computed MFI of NFAT1 in toddlers and school-age children (Nandi, n = 14; Kisumu, n = 15). Two-way ANOVA (***P < 0.01, ****P < 0.0001) comparing the amount of NFAT1 protein levels in CD8+ T cell subsets (red asterisks) or in paired samples from toddlers and school-age (black asterisks). Data generated from five independent experiments.
It is possible that the population of unconventional CD8dim T cells are NKT-like cells expressing low levels of CD8α (61). However, our data do not support this hypothesis, since we did not observe a significant difference in CD56 mRNA or protein expression between CD8bright and CD8dim T cells (Supplemental Figure 2, C and F, and data not shown). Altogether, our findings suggest that CD3+CD8dim Granzyme

Figure 9. School-age children from Kisumu (high pathogen burden) have lower proportions of CD8+ TEM and TEMRA cells coexpressing T-bet and Eomes transcription factors compared with Nandi (low pathogen burden). Peripheral blood mononuclear cells (PBMCs) were stained with antibodies to define T cell subsets and effector functions and then analyzed by multiparameter flow cytometry. (A) Gating strategy and (B) bivariate flow cytometry density plot displaying the proportion of CD8+ T cell subsets expressing T-bet and Eomes transcription factors. (C and D) Boxplots (median and 95% IQR) displaying the proportion of CD8bright and CD8dim T cell subsets coexpressing T-bet and Eomes from school-age children from Nandi (low pathogen burden, n = 14) and Kisumu (high pathogen burden, n = 15). Computed two-tailed unpaired t test with Welch’s correction P values are displayed on each graph (*P < 0.05, **P < 0.01, ***P < 0.001).
B*IFNγstim*TNFαstim T cells represent a distinct cell program in human CD8+ T cell biology. Further studies will elucidate their contribution in chronic diseases and cancer immunosurveillance, and their relevance to observed variation in vaccine efficacy in different populations.

**Methods**

**Study design.** We enrolled 140 children from the infant cohort previously described (19, 25) when they reached 2.5–3 years of age. Of those, we had 70% retention over the 4-year study period, until the age of 6.5–7 years. This was a normal, gradual rate of attrition expected for a longitudinal pediatric cohort study of healthy children. Those who continued to participate were not more or less likely to have been ill during the course of the study, aside from differences in episodes of uncomplicated malaria experienced in Kisumu children in contrast with children from Nandi. From those children with complete datasets, we randomly selected the subset for this immunologic study. For the current immunologic study, toddlers (2.4–3.5 years old) from Kisumu (n = 31), a high malaria transmission area on the shores of Lake Victoria, which is known to be a source of Sm (62) infections (high pathogen burden area) and Nandi (n = 33), a highland area with little malaria and no Schistosomiasis (low pathogen burden area) were followed prospectively until school age (5.5–6.5 years old). Healthy adult donors from Kisumu (n = 11) and the USA (n = 9) were also enrolled. Only HIV-negative individuals were enrolled.

**PBMC isolation.** PBMC collected in sodium heparin anticoagulated vacutainer tubes (BD Biosciences) were isolated using Ficoll-Hypaque gradient centrifugation. PBMCs were cryopreserved in 10% DMSO (90% heat-inactivated FBS). Plasma fractions were also collected. Samples were shipped from Kenya to the USA maintaining the cold chain. Cells were thawed and rested overnight in a 5% CO2 incubator (Forma Scientific - Co2 Water Jacketed Incubator), and viability was checked prior to commencing experiments.

**Multiplex bead-based assays (Bioplex).** Lucene bead–based multiplex immunoassay was used to simultaneously measure plasma antibody titers using previously published methods (25). In brief, recombinant proteins or synthetic peptides covering epitopes of Pf antigens (preerythrocytic-stage antigens CSP, CelTos, LSA1, and rPfSEA-A1; blood-stage antigens AMA1 [FVO], HRP-II, MSP1 [FVO, 3D7, FUP]; edmonston measles vaccine virus; tetanus toxoid; Sm worm antigen protein [SWAP]; and EBV antigens EAD, ZEBRA, VCA, EBNA1, BHFR1 p17 [otp 359, 360, 364]) were conjugated to beads to detect pathogen-specific antibodies following manufacturer’s instructions (Bio-Rad). Antibodies were gifts from Jaap Middeldorp (VU University Medical Center, Amsterdam, The Netherlands) for EBV serology profiles, Greg Poland and Inna Ovsyannikova from the Mayo Clinic (Rochester, Minnesota, USA) for the Measles Vaccine virus extract, Evelina Angov and Sheetij Dutta at Walter Reed Army Institute of Research (Silver Spring, Maryland, USA) for malaria recombinant proteins, and Jonathan Kurtis at Rhode Island Hospital for the Schistosome (Brown University, Providence, Rhode Island, USA) worm antigen preparations and malaria rPfSEA-A1.

Plasma cytokines, chemokines, and growth factors were measured using Bio-Plex Pro Human Cytokine 27-plex Assay (Bio-Rad, M500KCAF0Y) and Pro Hu INF Panel 1 24-Plex (Bio-Rad, 171AL002M) according to manufacturer instructions. Plasma was thawed on ice, filtered, and diluted prior to being run on the Bio-Rad platform.

Relative antibody titers were scaled (Z score normalization) for heatmap generation using R statistical packages. Spearman coefficient of correlation r between CD8stim and antibody titers, as well as statistical significance of antibody titers between children from low pathogen and high pathogen burden, was calculated with GraphPad PRISM using Welch’s Student t test assuming unequal variances. Statistical significance and data visualization were calculated with GraphPad PRISM and R statistical packages (CRAN).

**Flow cytometry analysis.** The following reagents were used to assess cell viability: live/dead aqua (Invitrogen, L34957) and cleaved Caspase-3 (Alexa Fluor 488, Cell Signaling Technologies). The cell staining method is a modified version of Lugli et al. protocol (63). The following fluorochrome-conjugated antibodies were used for surface staining: anti-human CD2 (BB515, RPA-2.10, BD Biosciences), CD3 (PE-Cy5, BD Biosciences), CD4 (PE-Texas Red, S3.5, Invitrogen), CD8 (Pacific Blue, RPA-T8, BD Biosciences), CD11a (FITC, BioLegend), CD45RA (PE-Cy7, L48, BD Biosciences), CD122/Tim3 (BV650, 7D3, BD Biosciences), ABCB1 (APC, UIC2, BioLegend), CCL4/MIP1β (APC, D21-1351, BD Pharmingen), CCR7 (PerCP-Cy5.5, 150503, BD Pharmingen), CXCRC4 (APC-Cy7, 12G5, BioLegend), KLRK1 (APC-Cy7, 1D11, BioLegend). Intracellular staining was performed following manufacturer instructions from fixation/permeabilization concentrate and diluent (eBiosciences). Surface-stained cells were fixed for
30 minutes at 4°C, washed twice in cold washing buffer, and stained 30 minutes at 4°C with the following antibodies: IFNγ (Alexa-Fluor 780, 4S.B3, ebiosciences), Granzyme B (PE, Invitrogen), KLF2 (PE, R&D Systems), NFAT1 (PE, 1:100, 14335, Cell Signaling Technologies), T-bet (FITC, eBio4B10, ebiosciences), Eomes (eFluor 660, WD1928, ebiosciences), TNFα (FITC, MAb11, ebiosciences). For the proliferation assay, sorted subsets of PBMCs were labeled with 1 M CFSE (Invitrogen) in PBS at 37°C for 8 minutes. Labeled cells were added to 96-well flat-bottom plates coated with plate-bound anti-CD3 (1 μg/ml) and anti-CD28 (2 μg/ml) (Invitrogen), at a concentration of 25,000 cells/well in 200 μl of RPMI (GIBCO) culture media supplemented with 2 mM glutamine, 1% (vol/vol) nonessential amino acids, 1% (vol/vol) sodium pyruvate, penicillin (50 U/ml) and streptomycin (50 μg/ml), and containing 10% (vol/vol) HI human serum (GEMINI Bio-Products). Stained cells were acquired on a 4-laser LSRII SORP analytical flow cytometer (BD Biosciences) and analyzed using FlowJo software (Tree Star Inc.). Compensated FCS files from each experimental conditions were later used as input files for ACCENSE (27) and PhenoGraph.

**Cell sorting and RNA sequencing.** CD8+ T cells were enriched from PBMCs (Stem Cell Technologies, 19053). Cells were stained with live/dead aqua (Invitrogen), propidium iodide (Invitrogen), CD3 (PE-Cy5, BD Biosciences), CD4 (PE-Texas Red, Invitrogen), CD8 (Pacific Blue, BD Biosciences), CCR7 (PerCP-Cy5.5, BD Pharmingen), and CD45RA (PE-Cy7, BD Pharmingen). Functional subsets defined by CCR7 and CD45RA were sorted in individual tubes from CD8bright and CD8dim T cells using a 4-laser BD FACSAria. Cells were collected in chilled 2× Buffer TCL with (2%) β-mercaptoethanol (Qiagen), and mRNA were isolated with oligo-dT using Dynabeads mRNA purification kit (Invitrogen). We followed standard protocol of SMARTer Stranded Total RNA-Seq Kit - Pico Input Mammalian for sequencing library preparation (Clontech). We performed the first-strand synthesis of purified mRNAs on Dynabeads using SMART pico oligos, which was followed by indexed Illumina Adapter ligation. We purified libraries using XP Ampure magnetic beads (Beckman Coulter Inc.) after each reaction step. Final libraries were amplified using SeqAmp DNA polymerase, and qualities were confirmed with a Bioanalyzer Agilent High sensitivity DNA kit and sequenced with paired-end read (2 × 100 bp) using an Illumina HiSeq 4000 (Illumina Inc.). Sequencing files were deposited in the NCBI’s database of Genotypes and Phenotypes (dbGaP phs001282.v2).

**Gene expression quantification and differential expression analysis.** The qualities of raw sequencing reads were checked using FastQC (64). Multiplexed reads were sorted by sample based on unique sample identifiers by Novobarc ode (Novocraft Technologies). Residual Illumina adaptor sequences on the 3′ end and sequences of template switching oligos introduced during the cDNA synthesis were trimmed using Fastx Toolkit (65). After preprocessing, paired reads were aligned to a transcriptome index built by RNA-Seq by expectation-maximization (RSEM) (66) using Gencode annotation version 19 for protein coding transcripts and GRCh37/hg19 genomic sequence. RSEM calculated the expected read counts for each gene with strand-specific settings. We merged all possible isoform-specific counts into a union gene model, which has the longest transcript structure. Libraries were assessed from separate batches and/or sample collection with principal component analysis. To perform differential gene expression analysis, we used Bioconductor package DESeq2 (67), which fits the read counts to a negative binomial distribution and tests for significance using Wald test. After adjusting the P values for multiple testing with Benjamini-Hochberg procedure, genes that are differentially expressed between cell subsets were considered significant with 10% FDR cut off.

**Gene set enrichment and pathway analysis.** GSEA was performed on normalized expression data of subsets using the GSEA program (http://software.broadinstitute.org/gsea/index.jsp) (68). For a ranking metric, we used signal to noise value of each gene and performed a permutation test for FDR by permuting phenotype assignments 1,000 times. Gene sets examined included the gene sets of hallmark, oncogenic, and immunologic signatures, as well as the curated C2 gene sets such as Reactome and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways from the Molecular Signatures Database (MSigDB version 5.0 datasets; http://software.broadinstitute.org/gsea/msigdb/). A FDR q value less than 0.1 was used as a cutoff for inclusion.

**Statistics.** Statistical analyses and statistical significance for each group were computed using a two-tailed Student t test with Welch’s correction for unequal variance (paired or unpaired). Two-way ANOVA with Sidak multiple comparison post test was used to analyze statistical significance for the 51 analytes measured in the two groups of divergent pathogen exposure. Significance levels are indicated as follows: \(* P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001.\) Linear regression analysis, as well as all the aforementioned statistical analyses, were performed using GraphPad PRISM, SPICE Software (Simplified Presentation of Incredibly Complex Evaluations), and R statistical packages.
Study approval. This study was approved by the IRB at the University of Massachusetts Medical School, and the Scientific & Ethics Review Unit at the Kenya Medical Research Institute. Written informed consent was obtained from adult study participants and parents of minors prior to inclusion in the study.

Author contributions
YTF, YK, MF, JAB, LJB, and AMM designed the experiments; YTF, MF, and YK carried out the experiments, analyzed the data, and interpreted the results; JAB, LJB, CF, and AMM also contributed to the interpretation of the results; YTF, MF, and YK prepared the figures; AMM and JMO carried out the epidemiologic study design; YTF, YK, and JAB processed and analyzed transcriptome data; YTF wrote the manuscript; MF, YK, CF, JAB, LJB, and AMM edited the manuscript; and all authors read the manuscript and approved submission.

Acknowledgments
This work was supported by grants from the US NIH (CA134051 and CA189806 to AMM and AI099473 to JAB, UMCCTS Pilot Project Program U1 LTR00161 to AMM and JAB, and AI106833 to LJB) and the Turkish Ministry of National Education Graduate Study Abroad Program to YK. We would like to thank Rosemary Rochford for donating some of the early-age PBMC samples (NIH CA102667) (25), and we thank the children and their families for their participation in this study. We would also like to thank Joslyn Foley and Christina Nixon for assistance with sample management and generating the serology data. In addition, we thank the flow cytometry and genomics core facilities at the University of Massachusetts Medical School for their assistance in sorting the cells and performing mRNA sequencing, respectively. This manuscript was approved for publication by the Kenya Medical Research Institute.

Address correspondence to: Ann M. Moormann, 373 Plantation Street, Biotech 2, Suite 318, Worcester, Massachusetts 01605, USA. Phone: 508.856.8826; Email: ann.moormann@umassmed.edu.


