IMMUNODIAGNOSTICS

Epigenetic immune cell counting in human blood samples for immunodiagnostics

Udo Baron[1]*, Jeannette Werner[1]*, Konstantin Schildknecht[1]*, Janika J. Schulze[1]*, Andargaschew Mulu[2,3], Uwe-Gerd Liebert[2], Ulrich Sack[4], Carsten Speckmann[5], Manfred Gossen[6,7], Ronald J. Wong[8], David K. Stevenson[8], Nina Babel[9], Dirk Schürmann[10], Tina Baldinger[1], Rosa Bacchetta[11], Andreas Grützkau[12], Stephan Borte[13,14†], Sven Olek[1†]

Immune cell profiles provide valuable diagnostic information for hematologic and immunologic diseases. Although it is the most widely applied analytical approach, flow cytometry is limited to liquid blood. Moreover, either analysis must be performed with fresh samples or cell integrity needs to be guaranteed during storage and transport. We developed epigenetic real-time quantitative polymerase chain reaction (qPCR) assays for analysis of human leukocyte subpopulations. After method establishment, whole blood from 25 healthy donors and 97 HIV$^+$ patients as well as dried spots from 250 healthy newborns and 24 newborns with primary immunodeficiencies were analyzed. Concordance between flow cytometric and epigenetic data for neutrophils and B, natural killer, CD3$^+$ T, CD8$^+$ T, CD4$^+$ T, and FOXP3$^+$ regulatory T cells was evaluated, demonstrating substantial equivalence between epigenetic qPCR analysis and flow cytometry. Epigenetic qPCR achieves both relative and absolute quantifications. Applied to dried blood spots, epigenetic immune cell quantification was shown to identify newborns suffering from various primary immunodeficiencies. Using epigenetic qPCR not only provides a precise means for immune cell counting in fresh-frozen blood but also extends applicability to dried blood spots. This method could expand the ability for screening immune defects and facilitates diagnostics of unobservantly collected samples, for example, in underdeveloped areas, where logistics are major barriers to screening.

INTRODUCTION

Quantitative abnormalities of lymphoid and myeloid immune cell subsets are indicative for several human diseases and therefore constitute important parameters for diagnosis and patient monitoring. Currently, immune cell quantification is mostly performed by flow cytometry, which provides flexibility with respect to the analyzed cell types and accuracy (*1*). However, although hematology analyzers used in diagnostic laboratories are highly developed and sample transportation times and conditions are optimized to prevent degradation, flow cytometry suffers from intrinsic limitations. The most critical challenge is that, flow cytometry–based cell counting requires intact leukocytes but fresh or well-preserved blood is not always available for all medical applications. Time to analysis influences results due to cell deterioration within a few hours after blood collection. Standardization remains a challenge due to biological, technical, and operational variations (*2*, *3*), and standardized protocols remain to be established, especially for samples with low numbers of certain cell populations, for example, in immunodeficiencies (*4*, *5*).

One situation that depends on immune profiling is monitoring HIV infection. Therapeutic decisions for HIV-infected patients depend on CD4$^+$ T cell counting. At frequencies below 500 CD4$^+$ T cells/μl of blood, antiretroviral therapy is recommended and becomes imperative below 200 cells/μl. In resource-poor regions, appropriate cell counting is hampered when blood collection and measurement cannot be performed in close succession. Therefore, treatment is initiated solely based on the occurrence of immunodeficiency-related clinical symptoms. In such situation, the chance to prevent AIDS-related conditions with potentially irreversible damage by CD4$^+$ T cell count–guided early antiviral therapy is missed (*6*, *7*).

Furthermore, flow cytometry is not applicable in newborn screening, which is routinely performed on dried blood spots (DBSs). Because of this sample type, quantitative deficiencies of specific leukocyte subpopulations are not detected, and most primary immunodeficiencies (PIDs) are not identified. The one exception is severe combined immunodeficiency (SCID), which is clinically characterized by the absence of T and/or B cells (*8*). Detection of SCID in newborns is currently based on quantitative polymerase chain reaction (PCR)–assisted T cell receptor excision circle (TREC) and immunoglobulin κ–deleting recombination excision circle (KREC) analyses (*9*). These methods reliably detect the lack of recent thymic T cell and bone marrow B cell emigrants, the predominant T and B cell subtypes present in neonatal blood. However, TREC/KREC analysis fails to detect other specific lymphocyte subsets defective in other severe PIDs, such as natural killer (NK) cells, regulatory T (T_{reg}) cells, or neutrophils. Despite this limitation, TREC newborn screening is effective and shows improved disease outcome due to early diagnosis (*10*).

To overcome current technological and diagnostic limitations and to broaden applicability of immune monitoring, we established DNA (un)

[1]Ivana Türbachova Laboratory for Epigenetics, Epiontis GmbH, Precision for Medicine Group, 12489 Berlin, Germany. [2]Institute of Virology, Faculty of Medicine, University Leipzig, 04009 Leipzig, Germany. [3]Armauer Hansen Research Institute, 1005 Addis Ababa, Ethiopia. [4]Institute of Clinical Immunology, Faculty of Medicine, University Leipzig, 04009 Leipzig, Germany. [5]Center for Chronic Immunodeficiency and Department of Pediatric and Adolescent Medicine, Division of Pediatric Hematology and Oncology, Faculty of Medicine, University of Freiburg, 79110 Freiburg, Germany. [6]Institute of Biomaterial Science, Helmholtz-Zentrum Geesthacht, 14513 Teltow, Germany. [7]Berlin-Brandenburg Center for Regenerative Therapies, 13353 Berlin, Germany. [8]Division of Neonatal and Developmental Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA 94305, USA. [9]Marienhospital Herne, Medizinische Klinik I, Universität Bochum, 44625 Herne, Germany. [10]Charité Universitätsmedizin Berlin, 13353 Berlin, Germany. [11]Division of Stem Cell Transplantation and Regenerative Medicine, Department of Pediatrics, Stanford School of Medicine, Stanford, CA 94305, USA. [12]Deutsches Rheuma-Forschungszentrum, an Institute of the Leibniz Association, Immune Monitoring Core Facility, 10117 Berlin, Germany. [13]ImmunoDeficiencyCenter Leipzig, Municipal Hospital St. Georg Leipzig, 04129 Leipzig, Germany. [14]Division of Clinical Immunology, Department of Laboratory Medicine, Karolinska University Hospital Huddinge at Karolinska Institutet, 14186 Stockholm, Sweden.
*These authors contributed equally to this work.
†Corresponding author. Email: sven.olek@epiontis.com (S.O.); stephan.borte@idcl.de (S.B.)

methylation-based, quantitative assessment of immune cells (epigenetic qPCR). This technique provides relative and absolute immune cell counts applicable to fresh-frozen or paper-spotted dried blood. Signals are digital; they indicate either one positive or negative value per cell rather than arbitrarily defined thresholds for "positivity" as in flow cytometric methods. Epigenetic qPCR can be performed in an automated, operator-independent manner and reduces susceptibility to reagent variability.

Cell type–specific DNA methylation markers (*11*–*13*) amplified in qPCR potentially allow immune cell quantification in samples of limited quantity and quality. The rationale for the identification of cell type–specific epigenetic markers has been described before (*12*, *14*–*16*). Alternative methods for DNA methylation–based immune cell quantification include the analysis of individual CpG sites on a genome-wide scale relying on microarray analysis (*17*). Such a method allows the estimation of leukocyte subpopulations based on calculated β values (that is, signal intensities).

For epigenetic qPCR, genomic DNA is treated with bisulfite. Unmethylated CpG dinucleotides are converted and amplified to TpGs, whereas methylated CpGs remain unaltered. Thus, bisulfite conversion translates epigenetic marks into sequence information, allowing discrimination and quantification of both variants. Epigenetic qPCR is resistant to loss of cell integrity because DNA is a stable substrate. It can be performed on fresh-frozen blood, DBS, or possibly other specimens without particular demands on preservation state. In addition, PCR components are synthetically produced, and standardization is easy to achieve. Nevertheless, immune cell counting via epigenetic qPCR has not yet been demonstrated. This is due to the absence of well-defined biomarkers specific for different immune cell types. In addition, definitive quantification, that is, unbiased assessment of cell counting after correction for method-inherent errors by a calibrator (*18*), and absolute immune cell counting (that is, cells per microliter) have not been shown for this method.

Here, we studied immune cell type–specific epigenetic qPCR for the quantification of leukocyte populations in human blood. For total $CD3^+$, $CD4^+$, and $CD8^+$ T cells, regulatory elements in the genes coding for the cell type determining proteins (*19*, *20*) were analyzed with respect to their methylation status. Epigenetic markers for neutrophils, B cells, and NK cells were identified from genome-wide discovery and subsequent profiling of resulting candidate genes. The T_{reg} cell–specific demethylated region has been described previously (*13*) and was shown to be overrepresented in its unmethylated state in immunodysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome, a severe PID (*21*). Determination of absolute cell numbers (cells per microliter of blood) constitutes the gold standard, for immune diagnostics in general.

We tested definitive and absolute quantification of immune cells based on their cell type–specific epigenetic signals in healthy donors and a cohort of HIV^+ patients and analyzed their equivalence to flow cytometry. For DBS, where the blood volume is difficult to define, copies of unmethylated immune cell type–specific marker genes were related to copies of a universal denominator (*GAPDH*). Moreover, the diagnostic potential of epigenetic qPCR was demonstrated by identifying PID cases in a cohort of clinically inconspicuous newborns using DBS.

RESULTS

Cell type–specific bisulfite conversion

Methylation-dependent conversion of CpG dinucleotides was analyzed by bisulfite sequencing (*22*), aiming at marker identification for immune cell populations from human peripheral blood of healthy adults. Candidate loci were selected from the literature or discovered using Illumina's 450k array-based assay. Our data showed absence of methylation at individual CpG positions for $CD56^+$ NK cells, $CD19^+$ B cells, and $CD15^+$ neutrophils (target cell types), whereas the same CpGs were methylated in control cell types (table S1, A and B). On the basis of these findings, amplicons (AMPs) were designed for the methylation analysis of all CpGs in the identified regions. DNA methylation of the intergenic *CD3G* and *CD3D* regions (AMP1405, AMP1406, and AMP1408), constituting a marker for $CD3^+$ T cells, and the methylation profile of *GAPDH* (AMP1570) were published previously (*13*). As a likely candidate marker for $CD4^+$ T cells, we designed three AMPs (AMP1255, AMP2000, and AMP2001) for the bisulfite sequence analysis covering regulatory elements within the 5′ region of the first intron in the *CD4* gene (*19*, *20*). Unmethylated CpG sites were detected as TpG residues after bisulfite conversion and amplification occurred exclusively in target $CD4^+$ T cells. The same CpGs were inert to bisulfite conversion in control cell types, including $CD56^+$ NK cells, $CD8^+$ T cells, $CD19^+$ B cells, and $CD15^+$ neutrophils (Fig. 1). The same is true for $CD14^+$ monocytes, but the analysis of this cell type indicated a single unmethylated CpG site in AMP2000. Next, we investigated the *CD8B* gene as a potential epigenetic marker for $CD8^+$ T cells (*19*, *20*) by designing an AMP targeting regulatory elements within its third intron (AMP2007). Here, bisulfite-mediated conversion of CpGs was observed exclusively in $CD8^+$ T cells, whereas those CpGs were inert to conversion in control cells. We identified epigenetic marks that uniquely associated with B cells, NK cells, and neutrophils in the genes coding for

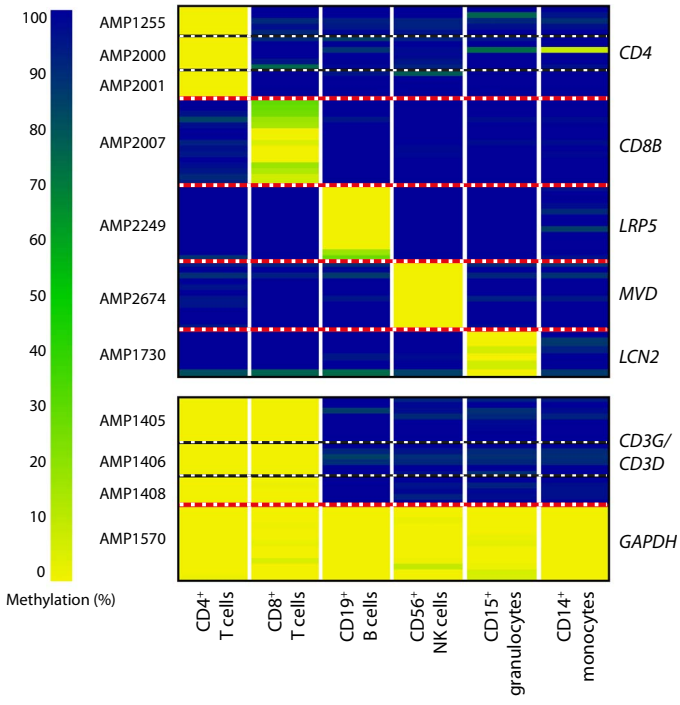

Fig. 1. Bisulfite sequencing–derived DNA methylation profiles of cell-specific marker genes in purified immune cells. Immune cell types, isolated from adult healthy donors, are arranged in columns with AMPs, and the associated gene names are arranged in rows. Different gene loci are separated by red dashed lines, and AMPs within the same locus are separated by black dashed lines. Each individual line represents a single CpG site. Methylation rates are color-coded ranging from yellow (0%) to blue (100%).

low-density lipoprotein receptor–related protein 5 (*LRP5*; AMP2249), mevalonate decarboxylase (*MVD*; AMP2674), and lipocalin 2 (*LCN2*; AMP1730), respectively. Each AMP was unmethylated in the target cell type and fully methylated in the corresponding control leukocyte populations (Fig. 1).

Locus-specific relative qPCR measurements

To target the differentially methylated CpG positions described above, we designed and characterized discriminating qPCR assay systems on synthetic template DNA cloned into plasmids. Templates corresponded to the bisulfite-modified genomic DNA, that is, all unmethylated cytosines (C) were replaced with thymidines (T). For the TpG template (mimicking unmethylated CpGs), we designed a plasmid carrying targets for all assays in an equimolar stoichiometry. A CpG plasmid (mimicking methylated CpGs) was similarly generated. Exclusive amplification of the desired DNA sequence without cross-reactivity with mutually antithetic templates was demonstrated for all qPCRs (Table 1). Assay specificity was tested on immune cell populations, which were purified from healthy adult human blood as described in Materials and Methods. For target cells, high copy numbers were observed in their respective TpG-specific system, low copy numbers were measured in the corresponding CpG system, and the converse was true in control cells. The original copy number of the target gene was determined by relating qPCR signals from the according amplification (f') to amplification of serially diluted standard plasmids (f; fig. S1), each with a defined concentration of the in silico–converted unmethylated version. Relative locus-specific unmethylated DNA (RD_{ls}) ranged from 89.9 to 100% in target cell types and from 0 to 3% in controls (Table 1). Exceptions were observed for $CD4^+$ T cells, showing 8.9% RD_{ls} at the *CD8B* locus and vice versa (that is, 9.6% *CD4* RD_{ls} in $CD8^+$ T cells), possibly due to residual cell contaminations and a small fraction of double-positive T cells.

Universal and definitive quantification of immune cells

Amplification efficiency and estimated copy numbers vary for each locus-specific qPCR system (*23*). Therefore, an invariably unmethylated regulatory region of the *GAPDH* (*24*) gene was used as a universal denominator to determine each cell type relative to all nucleated cells. This system was applied to purified $CD3^+$ T cells, $CD4^+$ T cells, $CD8^+$ T cells, $CD15^+$ neutrophils, $CD14^+$ monocytes, $CD56^+$ NK cells, and $CD19^+$ B cells (Table 1). Quantification of these subcell types is close to the expected values for purified cells (>98%) when using methylated and unmethylated amplification data at specific epigenetic loci (RD_{ls}), but shows assay-specific deviations from the expected values when analyzed using quantification of the unmethylated cell type–specific locus and the universally (that is, in all cell types) unmethylated *GAPDH* as the denominator (RD_u).

Because in silico–converted, double-stranded, GC-rich plasmids do not fully represent de facto bisulfite-converted, single-stranded, GC-depleted DNA (*25*, *26*), a "calibrator plasmid" was adopted harboring each one copy of all assay targets in their unconverted genomic (that is, unmethylated) state. This calibrator is bisulfite-converted in parallel to samples. When quantifying copy numbers as described in Materials and Methods, systematic amplification differences between the assays were detected and translated into an efficiency factor. We obtained this factor by calculating the quotient of TpG copies (of the respective immune cell type–specific assays) and *GAPDH* copies. We then used calculated efficiency factors to adjust for biases between cell type–specific assays and *GAPDH*. Cell type–specific efficiency factors were measured in about 25 experiments ranging between 0.53 [95% confidence interval (CI), 0.42 to 0.61] for *CD4* and 1.17 (95% CI, 0.95 to 1.31) for *CD3D/G* (see the "Epigenetic qPCR" section in Materials and Methods). Calculated efficiency factors provide universal definitive determination of unmethylated DNA (DD_u) for each assay (Table 1). Using this approach, we applied epigenetic qPCR for universal and definitive quantification of immune cells from biological samples. The concepts of immune cell quantification used in this work are illustrated in fig. S1.

Methodological comparison of flow cytometry and epigenetic qPCR

To allow absolute cell quantification comparable to flow cytometric measurements (that is, cells per microliter), we created a "spike-in plasmid" harboring an artificial *GAPDH*-derived sequence by inversing all CpG dinucleotides to GpC (GAP[GC]) and an epigenetic qPCR assay specific for that template. For absolute immune cell counting, this plasmid was added to blood samples in a defined concentration. The in silico bisulfite-converted, artificial GAP[GC] sequence was included in the quantification standard and the unconverted sequence into the calibrator plasmid.

To assess the overall performance of the epigenetic cell counting, we analyzed markers for B, NK, $CD3^+$ T, $CD4^+$ T, $CD8^+$ T, and $FOXP3^+$ T_{reg} cells and $CD15^+$ neutrophils in blood samples from 25 adult healthy donors in comparison with flow cytometry. Data from both methods were plotted either as relative (Fig. 2A) or absolute (Fig. 2B) cell counts. The joint comparison of all markers resulted in Spearman rank correlation coefficients (rho) of 0.96 and 0.97 ($P < 0.001$), respectively.

To challenge the individual epigenetic markers in a clinically relevant setting, we used blood from 97 HIV^+ subjects described in Materials and Methods and quantified $CD3^+$, $CD4^+$, and $CD8^+$ T cell counts by standard flow cytometry and by epigenetic qPCR using EDTA blood or DBS as samples. Method comparisons were conducted for all three approaches. For comparison of flow cytrometry data to epigenetic qPCR in liquid blood, correlation analyses for relative quantification of cell counts yielded Spearman rank correlation coefficient rho from 0.96 to 0.98 ($P < 0.001$; Fig. 3A). Leukocyte numbers per microliter of blood as determined by flow cytometry and epigenetic qPCR were highly correlated (Spearman rank correlation rho > 0.9; $P < 0.001$; fig. S2). Comparative analyses of relative cell counts from DBS and flow cytometry (Fig. 3B) yielded Spearman rank correlation between 0.74 and 0.95 ($P < 0.001$). Comparison of epigenetic measurements of relative cell counts from liquid blood and DBS yielded Spearman rank correlation rho between 0.8 and 0.95 ($P < 0.001$; Fig. 3C).

Bland-Altman analysis determines relative systematic biases and precision (*27*). Here, it was used to compare $CD3^+$, $CD4^+$, and $CD8^+$ T cell counting between flow cytometry and epigenetic qPCR on liquid and dried blood (Fig. 3, A to C, right). Relative mean differences between the methods were tested for normality using the one-sample Kolmogorov-Smirnov test. Normality assumption was not rejected. Biological readouts of flow cytometric measurements and epigenetic counting from either substrate presented a high degree of method agreement for the tested cell types, with minor biases (4.3, −6.6, and 10.3 for $CD3^+$, $CD4^+$, and $CD8^+$ T cells, respectively) and high precision (all <20%) between flow cytometry and epigenetic qPCR of liquid blood. For counting $CD4^+$ T cells from DBS by epigenetic qPCR, a pronounced variation (>20%) was observed when compared to both flow cytometric and epigenetic measurements from liquid blood samples (table S2).

To investigate the influence of substrate stability in DBS, we used different storage times and conditions and sample dilutions mimicking

Table 1. Cell type–specific epigenetic qPCR systems. RD_{ls}, relative determination of unmethylated DNA (locus specific) in %; RD_u, relative determination of unmethylated DNA (universal) in %; EF, efficiency factor; DD_u, definitive determination of unmethylated DNA (universal) in %.

Cell type specificity	Target gene of qPCR assay	Amplification system	Quantification mode	Plasmid-based controls			Analyzed immune cell populations					
				TpG variant	CpG variant	Calibrator	$CD4^+$ T cells	$CD8^+$ T cells	$CD19^+$ B cells	$CD56^+$ NK cells	$CD15^+$ granulocytes*	$CD14^+$ monocytes
$CD4^+$ T cells	CD4	TpG-system [#TpG]		30,100	0	6443	4795	244	50	58	61	57
		CpG-system [#CpG]		0	29,650		8	2300	7990	5100	5335	3600
			RD_{ls} [%]	**100**	**0**		**99.8**	***9.6***	0.6	1.1	1.1	1.6
			RD_u [%]				53.4	2.7	0.6	0.6	0.7	0.6
			EF			***0.53***						
			DD_u [%]				**91.4**	**6.1**	0.6	1.1	1.3	1.4
$CD8B^+$ T Cells	CD8B	TpG-system [#TpG]		29,850	0	10,457	622	5845	51	36	37	19
		CpG-system [#CpG]		0	27,150		6400	608	11,100	7375	7985	5720
			RD_{ls} [%]	**100**	**0**		8.9	**90.6**	0.5	0.5	0.5	0.3
			RD_u [%]				6.9	65.1	0.6	0.4	0.4	0.2
			EF			***0.87***						
			DD_u [%]				7.3	**90.6**	0.4	0.4	0.5	0.3
$CD19^+$ B cells	LRP5	TpG-system [#TpG]		30,500	0	8723	2	2	9970	24	1	5
		CpG-system [#CpG]		0	31,500		4760	3205	1125	5105	5790	3655
			RD_{ls} [%]	**100**	**0**		0.0	0.1	**89.9**	0.5	0.0	0.1
			RD_u [%]				0.0	0.0	111.0	0.3	0.0	0.1
			EF			***0.72***						
			DD_u [%]				0.0	0.0	**91.7**	0.3	0.0	0.1
$CD56^+$ NK cells	MVD	TpG-system [#TpG]		27,750	0	12,400	150	169	170	10,550	172	95
		CpG-system [#CpG]		0	25,750		9585	6850	16,450	494	11,200	7220
			RD_{ls} [%]	**100**	**0**		1.5	2.4	1.0	**95.5**	1.5	1.3
			RD_u [%]				1.7	1.9	1.9	117.5	1.9	1.1
			EF			***1.03***						
			DD_u [%]				1.5	2.2	1.1	**101.2**	1.9	1.2
$CD15^+$ neutrophils	LCN2	TpG-system [#TpG]		29,100	0	4520	14	17	95	9	3230	65
		CpG-system [#CpG]		0	31,400		3485	2420	5170	3620	30	2110
			RD_{ls} [%]	**100**	**0**		0.	0.7	1.8	0.2	**99.1**	3.0
			RD_u [%]				0.1	0.2	0.6	0.1	36.	0.8
			EF			***0.3***						
			DD_u [%]				0.	0.6	1.7	0.2	**94.7**	2.2
$CD3^+$ T cells	CD3 D/G	TpG-system [#TpG]		33,350	0	14,133	12,050	8320	37	59	28.8	23
		CpG-system [#CpG]		0	29,450		4	1	13,800	9505	9125.0	6810
			RD_{ls} [%]	**100**	**0**		**100.0**	**100.0**	0.3	0.6	0.3	0.3
			RD_u [%]				122.8	112.1	0.2	0.6	0.3	0.3
			EF			***1.17***						
			DD_u [%]				**104.4**	**95.2**	0.2	0.5	0.2	0.3
Leukocytes	GAPDH	TpG-system [#TpG]				12,050	9815	7425	15,100	10,100	8980	7655

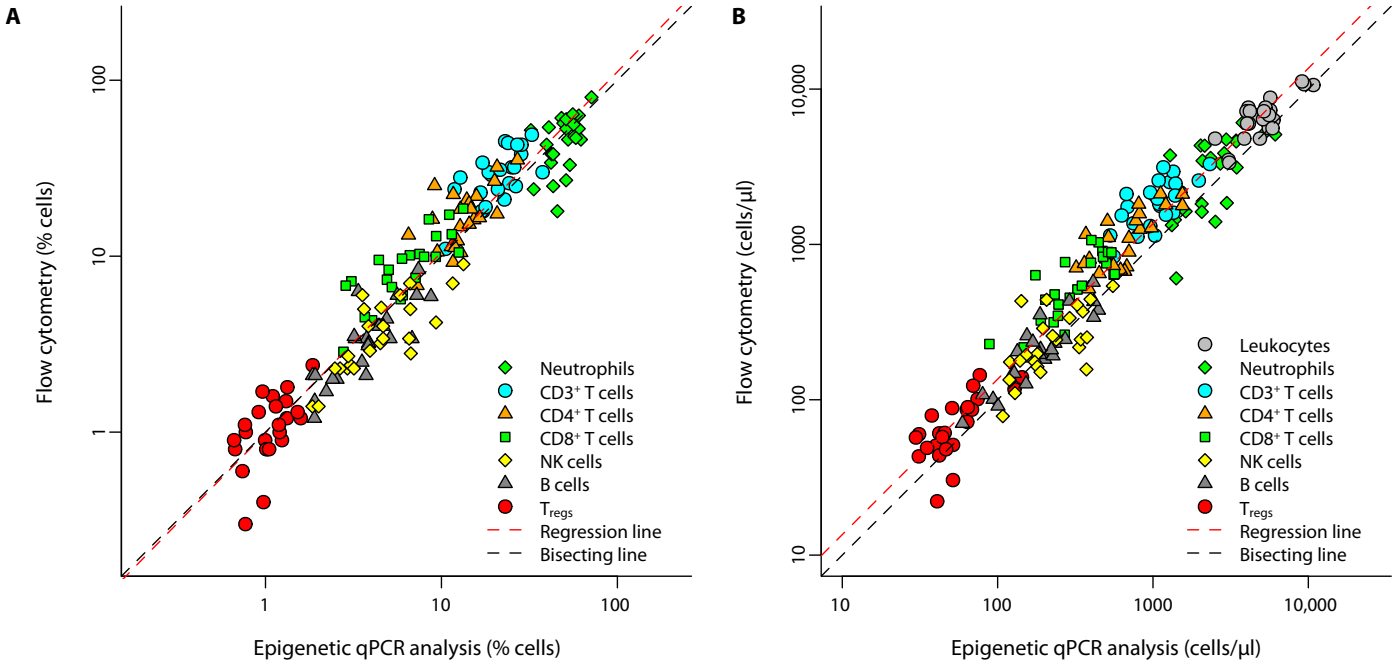

Fig. 2. Comparison of immune cell quantification by flow cytometry and epigenetic qPCR. Immune cells from blood samples of 25 independent donors were measured with flow cytometry (*y* axis) and epigenetic qPCR (*x* axis). (**A**) Relative immune cell counts are shown as percentage of total leukocytes. (**B**) Absolute immune cell counts are shown as cell number per microliter of whole blood. The regression line is depicted in red as computed from all data points, and the black line indicates the bisectrix.

unobservantly collected DBS. We did not observe degradation of template DNA at different storage conditions with coefficient of variation below 15% for all temperatures and time points (table S3). Coefficient of variation was below 30% down to blood dilutions of 1:9 (table S4). As previously shown by others, genomic DNA is a stable analyte and can be extracted and amplified from year-long stored DBS (*28*–*30*).

Epigenetic qPCR in neonatal screening samples

Epigenetic qPCR was applied in a case/control study consisting of original neonatal screening cards (DBS) from 24 PID patients and 250 randomly selected newborns, measuring total T, B, and NK cells (Fig. 4). PID cases included SCID patients with different gene defects and X-linked agammaglobulinemia (XLA) associated with *BTK* mutation (Table 2). Reference ranges were established using the joint distribution of all leukocytes (*GAPDH*-specific qPCR) and specific immune cell types. Copy numbers were log-transformed and used to estimate a bivariate normal distribution, whose confidence regions (99 and 99.9% curves) defined reference ranges for newborns. When testing for multivariate normality using the Henze-Zirkler test, we did not find evidence contradicting this assumption (*P* values for T, B, and NK cells were 0.21, 0.30, and 0.17, respectively). Because each of the three panels is tuned to 99 or 99.9% confidence regions, Bonferroni correction guarantees family-wise error rates below 3 or 0.3%, yielding final confidence of 99.7 or 97%, respectively.

For $CD3^+$ T cells and *GAPDH* measurements, 13 of 16 samples from SCID patients were outside the 99.9% confidence region, SCID15 was found outside the 99% but inside the 99.9% region, and SCID9 and SCID18 were presented as nonsuspicious. However, SCID15 and SCID18 were outside the 99.9% confidence region for NK cell and *GAPDH* measurements. Moreover, for B cell and *GAPDH* measurements, SCID18 was found outside the 99.9% confidence region, and SCID15 was found outside the 99% region (Fig. 4). Hence, 15 of 16 SCID patients were unambiguously identified as non-normal by epigenetic testing based on their newborn cards. SCID9 presented with maternal lymphocyte engraftment, as confirmed by flow cytometry and chromosomal analysis, and did not show a quantitative impairment of cell counts and was classified as normal despite the genetic defect. Epigenetic qPCR for $CD4^+$ and $CD8^+$ T cells and *GAPDH* confirmed the findings for all patients (fig. S3). DBSs of delayed-onset SCID (DO-SCID) associated with hypomorphic *JAK3* or *ADA* mutations were also analyzed. The JAK3-deficient delayed-onset patient (DO-SCID14) showed reduced CD3, NK, and B cell values (Fig. 4) outside the 99.9% confidence region. *ADA*-associated DO-SCID4 was outside the 99.9% confidence region for NK and B cells and outside the 99% region for $CD3^+$ T cells, whereas DO-SCID3 was outside the 99.9% confidence region in B cells and the 99% confidence region in NK but normal for $CD3^+$ T cells. Overall, all three DO-SCID samples were identified on the basis of the epigenetic analysis. DBS from four of five patients with XLA showed B cell counts outside of the 99.9% confidence region. Hypomorphic XLA24 was outside the 99% confidence region in B and NK cells. NK and T cell counts were at the borders of reference ranges for other XLA samples (XLA23 in NK cells and XLA20 and XLA23 in T cells). Together, the XLA phenotype was in accordance with B cell deficiency. Comparison with TREC/KREC values showed that epigenetic quantification detected all but one patient (SCID9), whereas TREC/KREC failed to detect two of five cases with delayed-onset or hypomorphic genetic background. However, maternal engraftment masked detection via epigenetic counting, whereas TREC analysis was not affected (Table 2).

Our screening classification bases on quantification of T, B, and NK cells. A sample was considered conspicuous if any of the three cell type–specific measurements were outside of the respective confidence

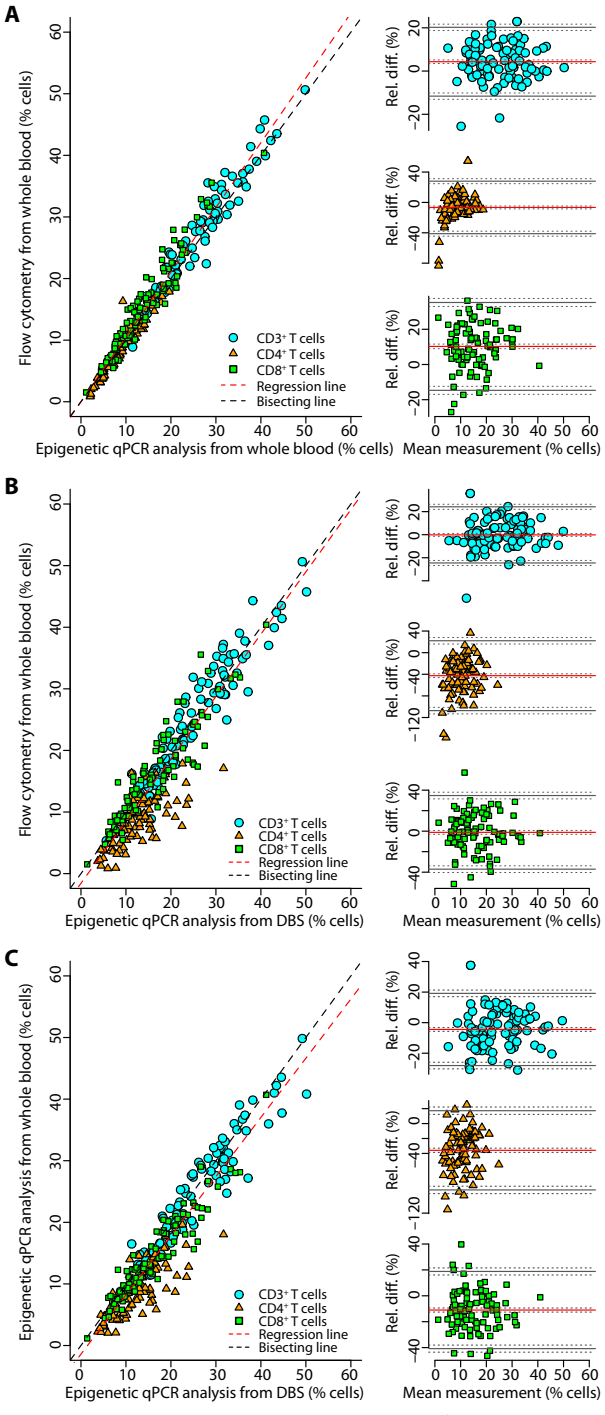

Fig. 3. Method comparison of T cell subsets in an HIV$^+$ cohort. Samples were analyzed from 97 HIV$^+$ subjects. Relative counts of CD3$^+$, CD4$^+$, and CD8$^+$ T cells in percentage of total nucleated cells determined by (**A**) flow cytometry and epigenetic qPCR in liquid whole blood, (**B**) flow cytometry as in liquid blood and epigenetic qPCR from DBS, and (**C**) comparison of epigenetic qPCR from liquid blood and DBS. On the left hand side, data are presented as scatterplots. The regression line is depicted in red as computed from all data points, and the black line indicates the bisectrix. On the right hand side, Bland-Altman plots show average cell counting of the respective analyses (*x* axis) plotted over their relative difference (*y* axis). Gray lines reflect limits of agreement. Central red lines illustrate the systematic bias. The respective 95% CIs are shown as dotted gray lines. Top, total CD3$^+$ T cells; middle, CD4$^+$ T cells; bottom, CD8$^+$ T cells.

region. This yielded sensitivity of 0.958 and specificity of 0.984 using the 99% confidence regions. For the 99.9% confidence regions, sensitivity was at 0.917, whereas specificity reaches 1. This compares to specificity of TREC/KREC analysis of 0.994 from data reported in (*31*) and with sensitivity of 0.917, as estimated from our data in Table 2.

IPEX and severe congenital neutropenia (SCN) are two other forms of severe PIDs with no currently available newborn screening. Given their severe early onset and morbidity, patients would benefit from neonatal diagnosis. In juvenile IPEX patients, peripheral T_{reg} cells are increased when compared to healthy age-matched donors and disease controls (*21*). Here, we tested DBS each from a newborn and a 2-year-old IPEX patient by epigenetic qPCR for T_{reg} and CD3$^+$ T cells (Fig. 5A). The percentage of T_{reg} cells among CD3$^+$ T cells of the two IPEX patients is increased compared to the nonaffected healthy newborns (n = 13). Applying neutrophil-specific epigenetic qPCR, we detected neonatal patients with SCN based on a significant reduction (Wilcoxon rank sum test, $P < 0.001$) of this cell type (Fig. 5B). The median percentage of neutrophils was at 55% in the control cards (n = 26) and at 17% in neutropenic patients (n = 6).

DISCUSSION

Current immune cell monitoring requires fresh or well-preserved blood, hampering diagnostics in medical settings where such samples are unavailable. Here, we describe immune cell type–specific epigenetic qPCR to allow the determination of immune cell counts from unobservantly preserved, paper-spotted dried blood or fresh-frozen samples.

Ideal DNA methylation markers for cell type identification are discriminative between target cells (near 0% methylation) and all control cells (near 100% methylation). In addition to analysis of T cell–associated genes *CD3G/D*, *CD4*, and *CD8B*, loci in genes *MVD*, *LRP5*, and *LCN2* were unmethylated only in NK cells, B cells, and neutrophils, respectively. *MVD* is a component of the mevalonate pathway (*32*) and is expressed in testis, duodenum, and colon. *LRP5* is involved in bone generation (*33*). *LCN2* is an extracellular transport protein and a major protein of the human tear fluid (*34*). Causes and consequences of methylation patterns in these regions remain unknown, but this does not affect their use for cell quantification in peripheral blood. All markers were validated by bisulfite sequencing, and only discriminatory CpG dinucleotides were selected for qPCR development and characterized on artificially generated methylated and unmethylated DNA. Quantitative amplification of target DNA was achieved with minimal background from nontarget templates detected. qPCR assay performance was robust, with small intra- and interassay coefficients of variation in fresh, frozen, or dried blood.

For the simultaneous quantification of different cell types in biological samples, we designed a calibrator plasmid containing the unmethylated genomic sequences of *GAPDH* as a reference quantifier along with cell type–specific markers. Whereas *GAPDH* was previously described as an unstable gene expression normalizer (*35*) containing segmental duplications (*36*), the *GAPDH* locus selected here is stably diploid and always unmethylated. Therefore, by normalizing the quantification of biological samples with the calibrator, assay-specific technical inefficiencies can be corrected to allow definitive quantification of the respective loci relative to unmethylated *GAPDH*. Hence, epigenetic qPCR indicates a direct proportional relation to cell types as determined by flow cytometry. The remaining observed biases between the methods may result from the biological and technical disparities between nucleic acid– and antibody-based methods. Homogeneous error distribution and precision were comparable to data from previously

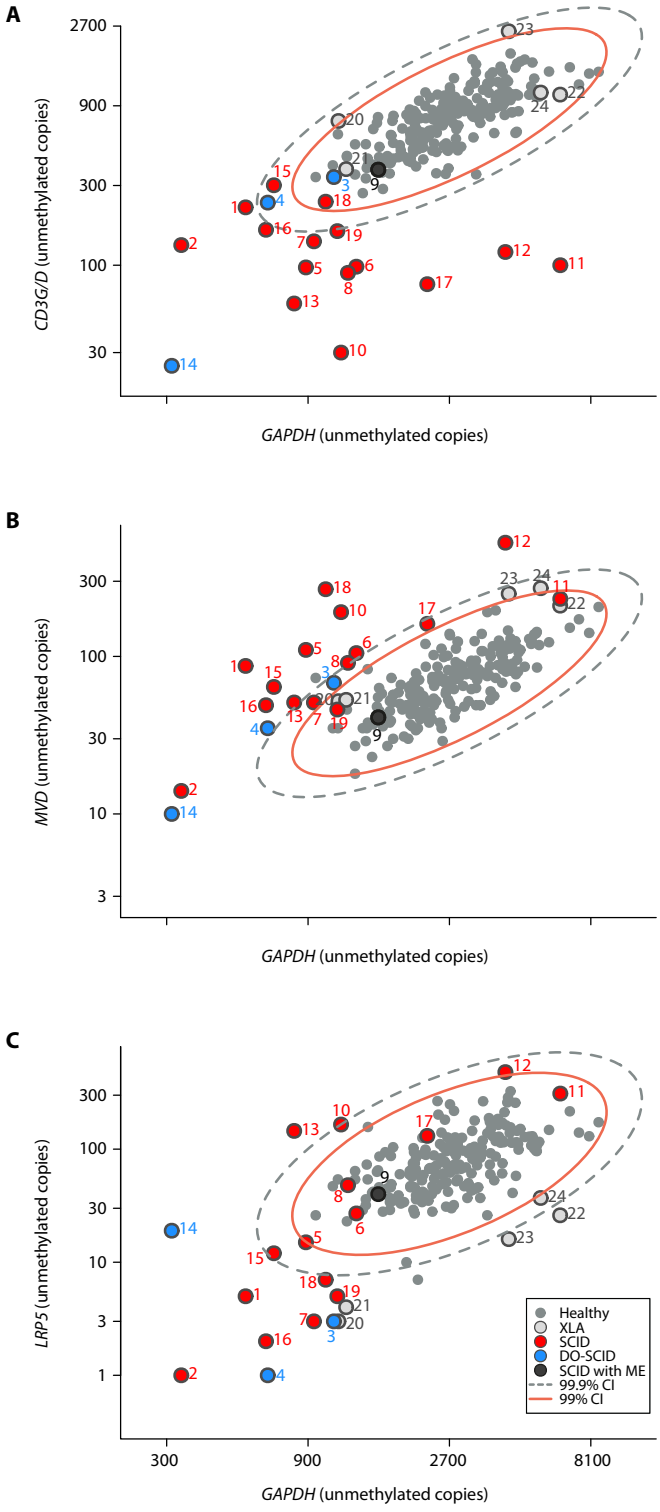

Fig. 4. Epigenetic qPCR on DBS from newborns. Copies from cell type–specific qPCRs (*y* axis) plotted against GAPDH (glyceraldehyde-3-phosphate dehydrogenase) copies (*x* axis). DBSs from healthy neonates (n = 250; gray circles) estimate reference ranges for each assay, as defined by 99% confidence region (red ellipse) and 99.9% confidence region (dashed and gray ellipse). Twenty-four DBSs from PID-diagnosed newborns are shown as colored circles, each referencing disease characteristics shown in Table 2. (**A**) Unmethylated *CD3G/D*, indicating T cells. (**B**) *MVD*, indicating NK cells. (**C**) *LRP5*, indicating B cells.

performed method comparisons among different antibody-based methods (*37*). Together, these data suggest that epigenetic qPCR, both from liquid and dried blood substrates, performs equivalently to flow cytometry for the relative quantification of immune cells.

With respect to clinical applications, relative cell quantification is accepted by the World Health Organization in HIV treatment guidelines, but in medical practice treatment, decisions depend on cell counting per volume (*38*, *39*). For epigenetic qPCR, this poses a problem because DNA recovery is not quantitative, and the relationship between DNA amount and blood volume is not fixed. For that reason, our experimental setup included the spiking of a defined concentration of artificial GAP[GC] into blood, allowing for an approximation of the original DNA content in a defined blood volume upon subsequent qPCR. Whereas different efficiencies of genomic and plasmid DNA have been described (*40*), such differences are possibly reduced upon bisulfite treatment and the resulting genomic DNA fragmentation. When applied to healthy donors and an HIV cohort, the comparison of immune cell counting via flow cytometry and epigenetic qPCR showed high correlation, low biases, and narrow limits of agreements similar to the data described for relative quantification method comparison. We concluded that immune cell counting per microliter can be performed by epigenetic qPCR equivalent to flow cytometry.

At present, neonatal screening is always performed from DBS. Because flow cytometry is not applicable to this substrate, TREC/KREC analysis is routinely used for SCID screening. Introducing epigenetic qPCR in such screening would therefore require equivalence testing to TREC/KREC. Because of different parameters tested, DNA excision circles versus genomic DNA, method comparison is not feasible. Instead, we estimated the specificity and sensitivity of TREC/KREC according to previously published work (*31*). Epigenetic qPCR reliably identified newborns suffering from different types of SCID and XLA with similar sensitivity and specificity when using the 99% confidence regions. It only failed to identify one newborn SCID patient with maternal cell engraftment, where the absence of T and B cells was masked by maternal cells. Such problems may be addressed by expanding the epigenetic qPCR portfolio to markers for memory T or B cells, which are absent in newborns and occur only after engraftment.

Unlike the analysis of excision circles, epigenetic analysis is not limited to the main lymphocyte subsets. Quantitative defects of other immune cell populations also affect neutrophils and T_{reg} cells. Our data indicate that the identification of patients with such defects based on epigenetic qPCR for neutrophils and T_{reg} cells is possible shortly after birth, allowing for early diagnosis of SCN and IPEX, respectively, which constitute potentially life-threatening PIDs (*41*, *42*). Moreover, the ability to quantify T_{reg} cells potentially opens the door to the early diagnosis of IPEX-like diseases recently described as T_{reg} cell deficiencies caused by genetic mutations (*43*). The importance of detecting and treating these severe immune disorders has been exemplified before (*44*).

The epigenetic qPCR approach and studies of rare disease in general have intrinsic limitations. For epigenetic qPCR, defects are only detected if they result in quantitative cellular aberrations, whereas functional deficiencies alone cannot be identified. More generally, scarceness of patients renders comprehensive studies of rare genetic diseases into major challenges. In the current study, this limitation is most apparent in the analysis of only two IPEX patients and six SCN patients. However, this is not true for the SCID analysis, where the number of patients with different genetic backgrounds is well comparable to previously published studies (*31*). The limited set of data provided in this study shows feasibility of the epigenetic qPCR approach but does not yet allow

Table 2. Genetic defects and diagnostic classification by TREC/KREC and epigenetic qPCR for PID patients.

Disease description				TREC/KREC newborn screening			Epigenetic qPCR analysis			
Identifier	Classification	Gene defect	Loss-of-function type	TREC* positive (yes/no)	KREC† positive (yes/no)	Screening classification	(CD3G/D, GAPDH)‡ conspicuous (yes/no)	(MVD, GAPDH)‡ conspicuous (yes/no)	(LRP5, GAPDH)‡ conspicuous (yes/no)	Screening classification
1	SCID	*ADA*	Amorph	Yes	Yes	Correctly identified	Yes	Yes	Yes	Correctly identified
2	SCID	*ADA*	Amorph	No	Yes	Correctly identified	Yes	Yes	Yes	Correctly identified
3	DO-SCID§	*ADA*	Hypomorph	No	Yes	Correctly identified	No	Yes	Yes	Correctly identified
4	DO-SCID§	*ADA*	Hypomorph	No	Yes	Correctly identified	Yes	Yes	Yes	Correctly identified
5	SCID	*AK2*	Amorph	Yes	No	Correctly identified	Yes	Yes	Yes	Correctly identified
6	SCID	*AK2*	Amorph	Yes	Yes	Correctly identified	Yes	Yes	No	Correctly identified
7	SCID	*Artemis*	Amorph	Yes	Yes	Correctly identified	Yes	Yes	Yes	Correctly identified
8	SCID	*CD3D*	Amorph	Yes	No	Correctly identified	Yes	Yes	No	Correctly identified
9	SCID w ME¶	*IL2RG*	Amoprh	Yes	No	Correctly identified	No	No	No	Not identified
10	SCID	*IL2RG*	Amorph	Yes	No	Correctly identified	Yes	Yes	Yes	Correctly identified
11	SCID	*IL7RA*	Amorph	Yes	No	Correctly identified	Yes	No	No	Correctly identified
12	SCID	*IL7RA*	Amorph	Yes	No	Correctly identified	Yes	Yes	Yes	Correctly identified
13	SCID	*IL7RA*	Amorph	Yes	No	Correctly identified	Yes	Yes	Yes	Correctly identified
14	DO-SCID§	*JAK3*	Hypomorph	No	No	Not identified	Yes	Yes	Yes	Correctly identified
15	SCID	*PNP*	Amorph	Yes	Yes	Correctly identified	Yes	Yes	Yes	Correctly identified
16	SCID	*PNP*	Amorph	Yes	Yes	Correctly identified	Yes	Yes	Yes	Correctly identified
17	SCID	*RAG1*	Hypomorph	Yes	Yes	Correctly identified	Yes	Yes	No	Correctly identified
18	SCID	*RAG1*	Amorph	Yes	Yes	Correctly identified	No	Yes	Yes	Correctly identified
19	SCID	*RAG2*	Amorph	Yes	Yes	Correctly identified	Yes	No	Yes	Correctly identified
20	XLA	*BTK*	Amorph	No	Yes	Correctly identified	Yes	No	Yes	Correctly identified
21	XLA	*BTK*	Amorph	No	Yes	Correctly identified	No	No	Yes	Correctly identified
22	XLA	*BTK*	Amorph	No	Yes	Correctly identified	No	No	Yes	Correctly identified
23	XLA	*BTK*	Amorph	No	Yes	Correctly identified	Yes	Yes	Yes	Correctly identified
24	XLA	*BTK*	Hypomorph	No	No	Not identified	No	Yes	Yes	Correctly identified

*TREC values ≤6 copies per dot were considered positive; †KREC values ≤4 copies per dot were considered positive; ‡Values outside the joint 99% reference range were considered conspicious (see Fig. 5); §Delayed onset SCID; ¶SCID with maternal engraftment

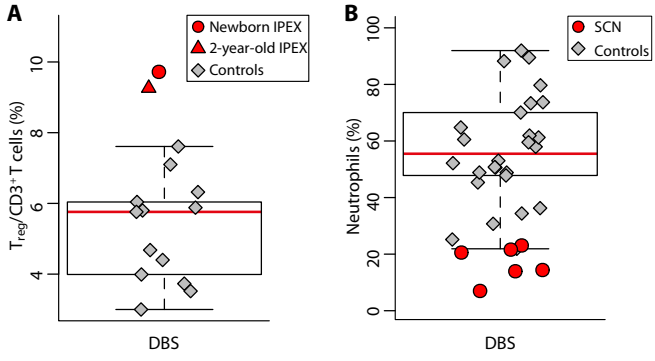

Fig. 5. Epigenetic qPCR on DBS from newborns with IPEX or SCN. DBS from healthy controls (gray-shaded boxes) and newborns with (**A**) confirmed IPEX and (**B**) SCN were subjected to epigenetic qPCR for quantification of the percentage of T_{reg} cells within $CD3^+$ T cells depicted in (A) or the percentage of $CD15^+$ neutrophils of all nucleated cells depicted in (B). Healthy cohorts [n = 13 in (A) and n = 26 in (B)] are represented in the boxplot, and results from diseased patients are depicted in red. Each box represents the interquartile range, and the central line shows the median.

translation into newborn screening. In addition, to validate the use of T_{reg} cell and neutrophil measurements for IPEX and SCN screening, respectively, larger studies are required to better determine potential issues regarding quality and storability of DBS and the efficiency of DNA extraction from DBS. Despite the strict limitations of this concept study, our data indicate that epigenetic qPCR may one day provide an option in medical screening procedures. Together, this study shows that epigenetic qPCR provides precise and accurate means for immune monitoring and it underscores that epigenetic qPCR may assist current immune diagnostics, particularly for unobservantly preserved blood or DBS.

MATERIALS AND METHODS

Study design

The research objective was to determine whether epigenetic qPCR can complement current methods for diagnostic immune cell counting. To test this, we identified and evaluated cell type–specifically unmethylated DNA loci for $CD15^+$ neutrophils and $CD19^+$ B, $CD56^+$ NK, $CD3^+$ T, $CD4^+$ T, $CD8^+$ T, and $FOXP3^+$ T_{reg} cells. On the basis of detected loci, epigenetic qPCRs were developed. Critical steps for their use were normalization and standardization of the different assays, which were performed with a universal denominator for the quantification of different target loci. Development of a calibrator system for compensation of differing amplification efficiencies for epigenetic qPCR at different loci and the establishment of a heterologous assay for normalization of DNA purification efficiency were basic prerequisites for absolute quantification. In a systematic methodological comparison of epigenetic qPCR with flow cytometry and using blood samples from 25 healthy donors and 97 HIV^+ patients, without AIDS under surveillance and/or standard treatment, relative and absolute quantifications were tested for equivalence of the differing methods. To further explore the use of

epigenetic qPCR in diagnostically undersupplied applications, we analyzed 250 DBSs from healthy newborns and 24 spots from newborn patients with PIDs. Anonymized or pseudonomized patient material was provided from German and Californian hospitals and blinded before data analysis.

Dried blood spots

Three 3.2-mm DBS punches of genetically confirmed IPEX, SCID, SCN, and XLA patients, and from 250 randomly selected anonymous, healthy newborns were analyzed. In addition, capillary blood of one patient with confirmed IPEX was obtained. The sequencing and genetic confirmation of the included PID patients were performed in compliance with the practitioner toolkit of the Clinical Sequencing Exploratory Research Consortium. Written parental consent was obtained for all participants. The study was approved by the Medical Association Chamber of Saxony ethics committee (protocol number EK-allg-37/10-1) or institutional review board at University of Freiburg, Germany (protocol number 281/11).

Peripheral whole blood

EDTA-anticoagulated peripheral blood was collected from 25 healthy subjects and 97 HIV^+ patients under treatment at Leipzig University with ethical consent (protocol number Az 301/16-ek). Samples were subjected to epigenetic qPCR and to standard flow cytometry (*45*). Experimenters were blinded to personal identifying information and to the results of the immunological analyses conducted in the study with the other technologies.

DNA preparation and bisulfite conversion

For purified cells, genomic DNA was isolated and bisulfite-treated using DNeasy tissue and EpiTect Fast Bisulfite Conversion kits (Qiagen), according to the manufacturer's instructions. For EDTA blood, 20 µl of substrate was supplemented with 16 µl of lysis buffer, 3 µl of proteinase K (30 mg/ml), and 1 µl of GAP[GC] plasmid (final concentration, 20,000 copies/µl) and lysed for 10 min at 56°C. For conversion, the EpiTect Fast Bisulfite Conversion Kit was used. DBS punches (3 mm × 3.2 mm) were added to 68.75 µl of lysis buffer, 10.75 µl of proteinase K (30 mg/ml), and GAP[GC] plasmid (final concentration, 20,000 copies/µl) and lysed for 60 min at 56°C. Conversion was performed for 45 min at 80°C, adding 180 µl of ammonium bisulfite [68 to 72% (pH 4.8 to 5.3); Chemos GmbH] and 60 µl of tetrahydrofurfuryl alcohol (Sigma-Aldrich). For purification, "My Silane Genomic DNA kit" (Invitrogen) was used following the manufacturer's instructions. Bisulfite conversion rates were analyzed previously and are provided in the manufacturer's manual with values above 98% (*46*). In addition, efficiency of conversion was routinely checked by bisulfite sequencing showing rates above 98%. As process control, the genomic calibrator included conversion controls in each individual qPCR. BioPerl was used for in silico bisulfite conversion of sequences (*47*).

Epigenetic qPCR

Thermal cycling was done as follows: 1× 95°C for 10 or 35 min, followed by 50× 95°C for 15 s, and 61°C for 1 min in 5 µl (DBS) or 10 µl (EDTA blood) using the Roche LightCycler 480 Probes Master. For the calculation of cell numbers from autosomal genes, a 2:1 allele-to-cell ratio was assumed.

For RD_{ls} [%], we divided TpG copy numbers by the sum of TpG and CpG copy numbers for each epigenetic qPCR. For RD_u [%], the quotient of TpG copy numbers (of the respective immune cell assay) and *GAPDH* copy numbers was calculated. To correct for qPCR assay–specific performance differences, we used a plasmid-based calibrator harboring the genomic target regions of all qPCRs including *GAPDH* as universal denominator. This calibrator was subjected to bisulfite conversion followed by qPCR. From this, an efficiency factor was calculated by dividing the TpG copy numbers for each cell type–specific assay by the *GAPDH* copy numbers measured from the same plasmid. Dividing RD_u by the efficiency factor results in definitive quantification (DD_u). Efficiency factors were derived from about 25 experiments. Ninety-five percent CIs were 0.90 to 1.19 (*CD3G/D*), 0.47 to 0.63 (*CD4*), 0.75 to 1.00 (*CD8B*), 0.58 to 0.77 (*LRP5*), 0.89 to 1.18 (*MVD*), and 0.38 to 0.48 (*LCN2*). For absolute quantification, an artificial *GAPDH* sequence inversing all CpG dinucleotides to GpC (GAP[GC]) and its corresponding epigenetic qPCR were designed without cross reactivity with endogenous *GAPDH*. The efficiency factor for GAP[GC] was 0.87 with a 95% CI of 0.75 to 1.00.

Combined TREC/KREC newborn screening assay

TREC/KREC screening was applied as described previously (*48*). Briefly, DNA from one 3.2-mm punch of the original DBS was extracted in a 96-well format, and quantitative triplex real-time qPCR for TREC, KREC, and β-*actin* (*ACTB*) was performed using the ViiA7 Real-Time PCR System (Applied Biosystems). TREC and KREC copy numbers were determined per 3.2-mm punch. *ACTB* was used to verify suitable DNA amounts per DBS and not for normalizing TREC/KREC copy numbers.

Plasmids

Sequences, corresponding to methylated or unmethylated bisulfite-converted genomic regions, were designed in silico and inserted into plasmid pUC57 (GenScript Inc.) and used for assay establishment and as qPCR quantification standard. Standard plasmids harbor all assay target sequences equimolarly. Plasmids were spectrophotometrically quantified, linearized by ScaI, and serially diluted in λ-phage DNA (10 ng/µl; New England Biolabs) to obtain 31,250, 6250, 1250, 250, 50, or 30 copies in the final reaction. The calibrator plasmid harbors all assay target sequences equimolarly in genomic unconverted, unmethylated version. The artificial spike-in plasmid carries unconverted *GAPDH* with inverted CpG dinucleotides (GAP[GC]).

Oligonucleotides

Oligonucleotides (Metabion AG) are described in table S5.

Flow cytometry

For leukocyte purification, peripheral blood from healthy adult donors was fractionated by flow cytometry into $CD15^+$ neutrophils, $CD14^+$ monocytes, and $CD56^+$ NK, $CD19^+$ B, $CD4^+$ T, and $CD8^+$ T cells with cell purities of >97% and viability of >99%, as described previously (*11*). For analytical cell quantification, absolute $CD45^+$ leukocyte counts were determined by a MACSQuant cytometer (Miltenyi Biotec). Frequencies and absolute counts of $CD15^+$ neutrophils and $CD19^+$ B, $CD56^+$ NK, $CD3^+$ T, $CD4^+$ T, $CD8^+$ T, and $FOXP3^+$ T_{reg} cells were calculated as previously described (*11*, *45*).

Statistical analysis

Crossing point of triplicate measurements was computed by second-derivative maximum applying LC480 software (Roche) to yield copy numbers (plasmid units) by interpolating amplification (*f*) from calibration curves generated with serial dilutions of plasmid-based standards.

Sample sizes for method comparison were chosen as 100 to provide 95% CI for limits of agreement at ±0.34 times the underlying SD. Estimation of reference ranges demands a healthy population of at least 120 individuals for the nonparametric estimation of the 95% CI. The number of healthy newborns was increased until exhaustion of available samples to accommodate for multidimensionality and estimation of extreme quantiles. Henze-Zirkler test was used to check for multivariate normality. Method comparison between flow cytometric– and qPCR-based measuring technique was done as follows: Bivariate data from the two methods were illustrated in a scatterplot. Linear regression was performed testing (i) for a slope different from 1 and (ii) an intercept different from 0. Relative mean differences were tested for normality using the one-sample Kolmogorov-Smirnov test. Bland-Altman plots were inspected analyzing bias and precision statistics (*27*). Acceptable precision was regarded as average deviation from the bias in percent. The limit of quantification for qPCR assays defined by the interassay coefficient of variation (0.2) was used as precision criterion and acceptable limits of agreement of 0.4. Wilcoxon rank sum test was used to test for median differences. The estimated bias, precision statistics, and respective 95% CI are reported. For correlation, Spearman rank sum correlations were used. All *P* values are two-sided. Statistics software R 3.3.0 was used.

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/10/452/eaan3508/DC1

Fig. S1. Schematic illustration of the different quantification approaches for epigenetic cell counting.

Fig. S2. Correlation analysis of absolute quantification for T cell subsets in an HIV cohort.

Fig. S3. Epigenetic immune cell quantification on DBS from newborns.

Table S1A. Selected candidate regions for neutrophils and NK and B cells from genome-wide discovery on Illumina's Infinium methylation-specific array.

Table S1B. Calculated "β values" from genome-wide discovery on Illumina's Infinium methylation-specific array.

Table S2. Data from method comparison analysis.

Table S3. Stability testing of DBS.

Table S4. Epigenetic qPCR from DBS spotted with diluted blood.

Table S5. Oligonucleotides for bisulfite sequencing and epigenetic qPCR.

REFERENCES AND NOTES

1. A. Adan, G. Alizada, Y. Kiraz, Y. Baran, A. Nalbant, Flow cytometry: Basic principles and applications. *Crit. Rev. Biotechnol.* **8551**, 1–14 (2016).
2. L. Whitby, A. Whitby, M. Fletcher, D. Barnett, Current laboratory practices in flow cytometry for the enumeration of CD 4$^+$T-lymphocyte subsets. *Cytometry B Clin. Cytom.* **88**, 305–311 (2015).
3. L. A. Herzenberg, J. Tung, W. A. Moore, L. A. Herzenberg, D. R. Parks, Interpreting flow cytometry data: A guide for the perplexed. *Nat. Immunol.* **7**, 681–685 (2006).
4. H. T. Maecker, J. P. McCoy, R. Nussenblatt, Standardizing immunophenotyping for the Human Immunology Project. *Nat. Rev. Immunol.* **12**, 191–200 (2012).
5. H. T. Maecker, J. P. McCoy Jr.; FOCIS Human Immunophenotyping Consortium, M. Amos, J. Elliott, A. Gaigalas, L. Wang, R. Aranda, J. Banchereau, C. Boshoff, J. Braun, Y. Korin, E. Reed, J. Cho, D. Hafler, M. Davis, C. G. Fathman, W. Robinson, T. Denny, K. Weinhold, B. Desai, B. Diamond, P. Gregersen, P. Dimeglio, F. Nestle, M. Peakman, F. Villnova, J. Ferbas, E. Field, A. Kantor, T. Kawabata, W. Komocsar, M. Lotze, J. Nepom, H. Ochs, R. O'Lone, D. Phippard, S. Plevy, S. Rich, M. Roederer, D. Rotrosen, J. H. Yeh, A model for harmonizing flow cytometry in clinical trials. *Nat. Immunol.* **11**, 975–978 (2010).
6. World Health Organization, *Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Public Health Approach* (World Health Organization, ed. 2, 2016), 155 p.
7. L. Ryom, C. Boesecke, V. Gisler, C. Manzardo, J. K. Rockstroh, M. Puoti, H. Furrer, J. M. Miro, J. M. Gatell, A. Pozniak, G. Behrens, M. Battegay, J. D. Lundgren; EACS Governing Board, Essentials from the 2015 European AIDS Clinical Society (EACS) guidelines for the treatment of adult HIV-positive persons. *HIV Med.* **17**, 83–88 (2016).
8. J. van der Spek, R. H. H. Groenwold, M. van der Burg, J. M. van Montfrans, TREC based newborn screening for severe combined immunodeficiency disease: A systematic review. *J. Clin. Immunol.* **35**, 416–430 (2015).
9. A. Sottini, C. Ghidini, C. Zanotti, M. Chiarini, L. Caimi, A. Lanfranchi, D. Moratto, F. Porta, L. Imberti, Simultaneous quantification of recent thymic T-cell and bone marrow B-cell emigrants in patients with primary immunodeficiency undergone to stem cell transplantation. *Clin. Immunol.* **136**, 217–227 (2010).
10. J. King, J. Ludvigsson, L. Hammarström, Newborn screening for primary immunodeficiency diseases: The past, the present and the future. *Int. J. Neonatal Screen.* **3**, 19 (2017).
11. U. Baron, S. Floess, G. Wieczorek, K. Baumann, A. Grützkau, J. Dong, A. Thiel, T. J. Boeld, P. Hoffmann, M. Edinger, I. Türbachova, A. Hamann, S. Olek, J. Huehn, DNA demethylation in the human *FOXP3* locus discriminates regulatory T cells from activated FOXP3$^+$ conventional T cells. *Eur. J. Immunol.* **37**, 2378–2389 (2007).
12. G. Wieczorek, A. Asemissen, F. Model, I. Turbachova, S. Floess, V. Liebenberg, U. Baron, D. Stauch, K. Kotsch, J. Pratschke, A. Hamann, C. Loddenkemper, H. Stein, H. D. Volk, U. Hoffmüller, A. Grützkau, A. Mustea, J. Huehn, C. Scheibenbogen, S. Olek, Quantitative DNA methylation analysis of *FOXP3* as a new method for counting regulatory T cells in peripheral blood and solid tissue. *Cancer Res.* **69**, 599–608 (2009).
13. J. Sehouli, C. Loddenkemper, T. Cornu, T. Schwachula, U. Hoffmüller, A. Grützkau, P. Lohneis, T. Dickhaus, J. Gröne, M. Kruschewski, A. Mustea, I. Turbachova, U. Baron, S. Olek, Epigenetic quantification of tumor-infiltrating T-lymphocytes. *Epigenetics* **6**, 236–246 (2014).
14. S. Rapko, U. Baron, U. Hoffmüller, F. Model, L. Wolfe, S. Olek, DNA methylation analysis as novel tool for quality control in regenerative medicine. *Tissue Eng.* **13**, 2271–2280 (2007).
15. T. O. Kleen, J. Yuan, Quantitative real-time PCR assisted cell counting (qPACC) for epigenetic-based immune cell quantification in blood and tissue. *J. Immunother. Cancer* **3**, 46 (2015).
16. E. Houseman, W. P. Accomando, D. C. Koestler, B. C. Christensen, C. J. Marsit, H. H. Nelson, J. K. Wiencke, K. T. Kelsey, DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **13**, 86 (2012).
17. W. P. Accomando, J. K. Wiencke, E. Houseman, H. H. Nelson, K. T. Kelsey, Quantitative reconstruction of leukocyte subsets using DNA methylation. *Genome Biol.* **15**, R50 (2014).
18. J. W. Lee, V. Devanarayan, Y. C. Barrett, R. Weiner, J. Allinson, S. Fountain, S. Keller, I. Weinryb, M. Green, L. Duan, J. A. Rogers, R. Millham, P. J. O'Brien, J. Sailstad, M. Khan, C. Ray, J. A. Wagner, Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm. Res.* **23**, 312–328 (2006).
19. P. Kung, G. Goldstein, E. L. Reinherz, S. F. Schlossman, Monoclonal antibodies defining distinctive human T cell surface antigens. *Science* **206**, 347–349 (1979).
20. E. L. Reinherz, S. F. Schlossman, Regulation of the immune response—Inducer and suppressor T-lymphocyte subsets in human beings. *N. Engl. J. Med.* **303**, 370–373 (1980).
21. F. Barzaghi, L. Passerini, E. Gambineri, S. Ciullini Mannurita, T. Cornu, E. S. Kang, Y. H. Choe, C. Cancrini, S. Corrente, R. Ciccocioppo, M. Cecconi, G. Zuin, V. Discepolo, C. Sartirana, J. Schmidtko, A. Ikinciogullari, A. Ambrosi, M. G. Roncarolo, S. Olek, R. Bacchetta, Demethylation analysis of the FOXP3 locus shows quantitative defects of regulatory T cells in IPEX-like syndrome. *J. Autoimmun.* **38**, 49–58 (2012).
22. J. Lewin, A. O. Schmitt, P. Adorján, T. Hildmann, C. Piepenbrock, Quantitative DNA methylation analysis based on four-dye trace data from direct sequencing of PCR amplificates. *Bioinformatics* **20**, 3005–3012 (2004).
23. P. M. Warnecke, C. Stirzaker, J. R. Melki, D. S. Millar, C. L. Paul, S. J. Clark, Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res.* **25**, 4422–4426 (1997).
24. H. J. M. de Jonge, R. S. N. Fehrmann, E. S. J. M. de Bont, R. M. W. Hofstra, F. Gerbens, W. A. Kamps, E. G. E. de Vries, A. G. J. van der Zee, G. J. te Meerman, A. ter Elst, Evidence based selection of housekeeping genes. *PLOS ONE* **2**, e898 (2007).
25. D. Chen, P. S. Rudland, H. L. Chen, R. Barraclough, Differential reactivity of the rat *S100A4* (*p9Ka*) gene to sodium bisulfite is associated with differential levels of the *S100A4* (p9Ka) mRNA in rat mammary epithelial cells. *J. Biol. Chem.* **274**, 2483–2491 (1999).
26. J. Harrison, C. Stirzaker, S. J. Clark, Cytosines adjacent to methylated CpG sites can be partially resistant to conversion in genomic bisulfite sequencing leading to methylation artifacts. *Anal. Biochem.* **264**, 129–132 (1998).
27. D. Giavarina, Understanding Bland Altman analysis. *Biochem. Med.* **25**, 141–151 (2015).
28. S. Chaisomchit, R. Wichajarn, N. Janejai, W. Chareonsiriwatana, Stability of genomic DNA in dried blood spots stored on filter paper, Southeast Asian. *J. Trop. Med. Public Health* **36**, 270–273 (2005).
29. M. V. Hollegaard, J. Grauholm, B. Nørgaard-Pedersen, D. M. Hougaard, DNA methylome profiling using neonatal dried blood spot samples: A proof-of-principle study. *Mol. Genet. Metab.* **108**, 225–231 (2013).
30. M. V. Hollegaard, P. Thorsen, B. Norgaard-Pedersen, D. M. Hougaard, Genotyping whole-genome-amplified DNA from 3- to 25-year-old neonatal dried blood spot samples with reference to fresh genomic DNA. *Electrophoresis* **30**, 2532–2535 (2009).
31. S. Borte, U. Von Döbeln, A. Fasth, N. Wang, M. Janzi, J. Winiarski, U. Sack, Q. Pan-Hammarström, M. Borte, L. Hammarström, Neonatal screening for severe primary immunodeficiency diseases using high-throughput triplex real-time PCR. *Blood* **119**, 2552–2555 (2012).

32. H. M. Miziorko, Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch. Biochem. Biophys.* **505**, 131–143 (2011).
33. T. Mizuguchi, I. Furuta, Y. Watanabe, K. Tsukamoto, H. Tomita, M. Tsujihata, T. Ohta, T. Kishino, N. Matsumoto, H. Minakami, N. Niikawa, K. I. Yoshiura, LRP5, low-density-lipoprotein-receptor-related protein 5, is a determinant for bone mineral density. *J. Hum. Genet.* **49**, 80–86 (2004).
34. B. Redl, Human tear lipocalin. *Biochim. Biophys. Acta* **1482**, 241–248 (2000).
35. M. Krzystek-Korpacka, D. Diakowska, J. Bania, A. Gamian, Expression stability of common housekeeping genes is differently affected by bowel inflammation and cancer: Implications for finding suitable normalizers for inflammatory bowel disease studies. *Inflamm. Bowel Dis.* **20**, 1147–1156 (2014).
36. M. Ghani, C. Sato, E. Rogaeva, Segmental duplications in genome-wide significant loci and housekeeping genes; warning for GAPDH and ACTB. *Neurobiol. Aging* **34**, 1710. e1–1710.e4 (2013).
37. W. R. Rodriguez, N. Christodoulides, P. N. Floriano, S. Graham, S. Mohanty, M. Dixon, M. Hsiang, T. Peter, S. Zavahir, I. Thior, D. Romanovicz, B. Bernard, A. P. Goodey, B. D. Walker, J. T. McDevitt, A microchip CD4 counting method for HIV monitoring in resource-poor settings. *PLOS Med.* **2**, e182 (2005).
38. D. M. Moore, R. S. Hogg, B. Yip, K. Craib, E. Wood, J. S. G. Montaner, CD4 percentage is an independent predictor of survival in patients starting antiretroviral therapy with absolute CD4 cell counts between 200 and 350 cells/μL. *HIV Med.* **7**, 383–388 (2006).
39. L. M. Yu, P. J. Easterbrook, T. Marshall, Relationship between CD4 count and CD4% in HIV-infected people. *Int. J. Epidemiol.* **26**, 1367–1372 (1997).
40. S. J. Read, Recovery efficiencies of nucleic acid extraction kits as measured by quantitative LightCyclerTM PCR. *Mol. Pathol.* **54**, 86–90 (2001).
41. F. Hauck, C. Klein, Pathogenic mechanisms and clinical implications of congenital neutropenia syndromes. *Curr. Opin. Allergy Clin. Immunol.* **13**, 596–606 (2013).
42. K. Bin Dhuban, C. A. Piccirillo, The immunological and genetic basis of immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome. *Curr. Opin. Allergy Clin. Immunol.* **15**, 525–532 (2015).
43. R. E. Schmidt, B. Grimbacher, T. Witte, Autoimmunity and primary immunodeficiency: Two sides of the same coin? *Nat. Rev. Rheumatol.* **14**, 7–18 (2018).
44. L. Brown, J. Xu-Bayford, Z. Allwood, M. Slatter, A. Cant, E. G. Davies, P. Veys, A. R. Gennery, H. B. Gaspar, Neonatal diagnosis of severe combined immunodeficiency leads to significantly improved survival outcome: The case for newborn screening. *Blood* **117**, 3243–3246 (2011).
45. A. Boldt, S. Borte, S. Fricke, K. Kentouche, F. Emmrich, M. Borte, F. Kahlenberg, U. Sack, Eight-color immunophenotyping of T-, B-, and NK-cell subpopulations for characterization of chronic immunodeficiencies. *Cytometry B Clin. Cytom.* **86**, 191–206 (2014).
46. E. E. Holmes, M. Jung, S. Meller, A. Leisse, V. Sailer, J. Zech, M. Mengdehl, L.-A. Garbe, B. Uhl, G. Kristiansen, D. Dietrich, Z. Zuo, Performance evaluation of kits for bisulfite-conversion of DNA from tissues, cell lines, FFPE tissues, aspirates, lavages, effusions, plasma, serum, and urine. *PLOS ONE* **9**, e93933 (2014).
47. J. E. Stajich, An introduction to BioPerl. *Methods Mol. Biol.* **406**, 535–548 (2007).
48. M. Barbaro, A. Ohlsson, S. Borte, S. Jonsson, R. H. Zetterström, J. King, J. Winiarski, U. von Döbeln, L. Hammarström, Newborn screening for severe primary immunodeficiency diseases in Sweden—A 2-year pilot TREC and KREC screening study. *J. Clin. Immunol.* **37**, 51–60 (2017).

Acknowledgments: We thank N. Brodszki, A. Fasth, and A. Yildiran for allocating SCID samples and G. M. Shawn and the March of Dimes Prematurity Research Center at Stanford University for allocating normal blood. Ivana, it is better to have loved and lost than never to have loved at all. **Funding:** This work was supported by grants from the German Bundesministerium für Bildung und Forschung (EPILYZE, 031A191; TransScan, 01KT1305; Prägung der Immunität am Lebensbeginn, 01GL1746A, eKID 01ZX1312, and 01ZX1612), Bundesministerium für Wirtschaft und Energie (Zentrales Innovationsprogramm Mittelstand, 16KN041880), the Child Health Research Institute, and SPARK (Stanford University) to R.B. Immunodeficiency Center Leipzig received support from the Jeffrey Modell Foundation. **Author contributions:** U.B. and T.B. performed the epigenetic qPCR on healthy and HIV samples. J.W. and J.J.S. performed the epigenetic qPCR on DBS. K.S. performed the statistical analysis. A.M., U.-G.L., U.S., D.S., and N.B. provided the HIV samples and performed the flow cytometry. C.S., M.G., R.J.W., D.K.S., and R.B. provided the DBS and the clinical background from controls and IPEX patients. A.G. fluorescence-activated cell–sorted all cells. S.B. provided the clinical data and samples from SCID, XLA, SCN, and control DBS. S.O. designed the study and wrote the manuscript. **Competing interests:** U.B. and S.O. are inventors to patent applications (EP1826279, EP2199411, WO2013135454, WO2014/080017, WO2017/050916, and WO2017/050882) for epigenetic immune cell markers. S.O. has equity interest in Precision for Medicine, the owner of the epigenetic markers described here. **Data and materials availability:** All data associated with this study are present in the paper or Supplementary Materials. All materials can be requested from S.O.

Submitted 5 April 2017
Resubmitted 23 February 2018
Accepted 18 June 2018
Published 1 August 2018
10.1126/scitranslmed.aan3508

Citation: U. Baron, J. Werner, K. Schildknecht, J. J. Schulze, A. Mulu, U.-G. Liebert, U. Sack, C. Speckmann, M. Gossen, R. J. Wong, D. K. Stevenson, N. Babel, D. Schürmann, T. Baldinger, R. Bacchetta, A. Grützkau, S. Borte, S. Olek, Epigenetic immune cell counting in human blood samples for immunodiagnostics. *Sci. Transl. Med.* **10**, eaan3508 (2018).

Science Translational Medicine

Epigenetic immune cell counting in human blood samples for immunodiagnostics

Udo Baron, Jeannette Werner, Konstantin Schildknecht, Janika J. Schulze, Andargaschew Mulu, Uwe-Gerd Liebert, Ulrich Sack, Carsten Speckmann, Manfred Gossen, Ronald J. Wong, David K. Stevenson, Nina Babel, Dirk Schürmann, Tina Baldinger, Rosa Bacchetta, Andreas Grützkau, Stephan Borte and Sven Olek

Sci Transl Med **10**, eaan3508.
DOI: 10.1126/scitranslmed.aan3508

Counting by chromatin

Peripheral immune cell counts can be wielded to diagnose a variety of disorders. There are limitations to traditional methods such as flow cytometry, including the type of sample needed for analysis. Baron *et al.* took advantage of distinct immune cell epigenetic signatures and devised a real-time quantitative polymerase chain reaction (qPCR) method to perform immune cell counting without the requirement of viable cells. They examined different types of samples from healthy adults or those that were infected with HIV (and consequently had fewer $CD4^+$ T cells). The epigenetic qPCR method correlated well with flow cytometry and could also be applied to dried blood spots to identify newborns with primary immunodeficiencies. Although it is not ready to be deployed clinically, the advantages of using epigenetic qPCR make this an intriguing approach to develop further.