New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling

Willem A. Dik,[1] Karin Pike-Overzet,[1] Floor Weerkamp,[1] Dick de Ridder,[1,3]
Edwin F.E. de Haas,[1] Miranda R.M. Baert,[1] Peter van der Spek,[2]
Esther E.L. Koster,[1] Marcel J.T. Reinders,[3] Jacques J.M. van Dongen,[1]
Anton W. Langerak,[1] and Frank J.T. Staal[1]

[1]Department of Immunology and [2]Department of Bioinformatics, Erasmus MC, 3015 GE Rotterdam, Netherlands
[3]Information and Communication Theory Group, Faculty of Electrical Engineering, Mathematics and Computer Science,
Delft University of Technology, 2600 GA Delft, Netherlands

To gain more insight into initiation and regulation of T cell receptor (TCR) gene rearrangement during human T cell development, we analyzed TCR gene rearrangements by quantitative PCR analysis in nine consecutive T cell developmental stages, including CD34$^+$ lin$^-$ cord blood cells as a reference. The same stages were used for gene expression profiling using DNA microarrays. We show that TCR loci rearrange in a highly ordered way (*TCRD–TCRG–TCRB–TCRA*) and that the initiating Dδ2-Dδ3 rearrangement occurs at the most immature CD34$^+$CD38$^-$CD1a$^-$ stage. *TCRB* rearrangement starts at the CD34$^+$CD38$^+$CD1a$^-$ stage and complete in–frame *TCRB* rearrangements were first detected in the immature single positive stage. *TCRB* rearrangement data together with the *PTCRA* (pTα) expression pattern show that human TCRβ–selection occurs at the CD34$^+$CD38$^+$CD1a$^+$ stage. By combining the TCR rearrangement data with gene expression data, we identified candidate factors for the initiation/regulation of TCR recombination. Our data demonstrate that a number of key events occur earlier than assumed previously; therefore, human T cell development is much more similar to murine T cell development than reported before.

CORRESPONDENCE
Frank J.T. Staal:
f.staal@erasmusmc.nl

T cells develop from progenitors that migrate from the bone marrow into the thymus (1). Thymocytes are subdivided roughly as being double negative (DN), double positive (DP), or single positive (SP), based on the expression of the CD4 and CD8 coreceptors (1). The DN stage is heterogeneous and can be subdivided into four distinct subsets in mice based on the expression of CD44 and CD25. In human, three distinct DN stages can be recognized: a CD34$^+$CD38$^-$CD1a$^-$ stage that represents the most immature thymic subset and the consecutive CD34$^+$CD38$^+$CD1a$^-$ and CD34$^+$CD38$^+$CD1a$^+$ stages. Human DN thymocytes mature via an immature single positive (ISP CD4$^+$) and a DP stage into CD4$^+$ or CD8$^+$ SP T cells that express functional T cell receptors (TCR) and that exit the thymus (1).

A hallmark of T cell development is the generation of T cells that express a functional TCR, TCRαβ or TCRγδ. During T cell development, the variable domains of *TCRA*, *TCRB*, *TCRG*, and *TCRD* (located within *TCRA*) genes are assembled following rearrangement of variable (V), diversity (D), and joining (J) gene segments by a process called V(D)J recombination (2). V(D)J recombination uses the RAG1 and RAG2 enzymes that selectively target recombination signal sequences that flank V, D, and J segments (2).

Studies in T cell acute lymphoblastic leukemias suggest that recombinations of TCR genes are sequential between the different genes (*TCRD* > *TCRG* > *TCRB* > *TCRA*) as well as within a particular gene (e.g., *TCRD*: Dδ2-Dδ3, Dδ2-Jδ1, Vδ-Jδ1) (2, 3), which is supported by limited data that were obtained from normal human T cell subsets (4). Therefore, the timing and efficiency of rearrangement of

W.A. Dik and K. Pike-Overzet contributed equally to this work.
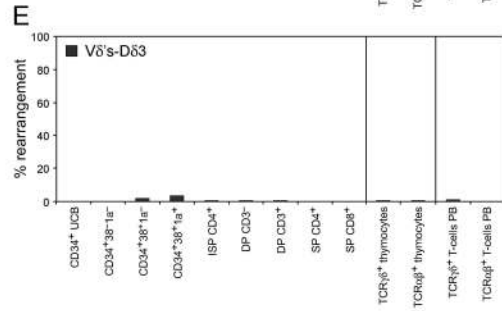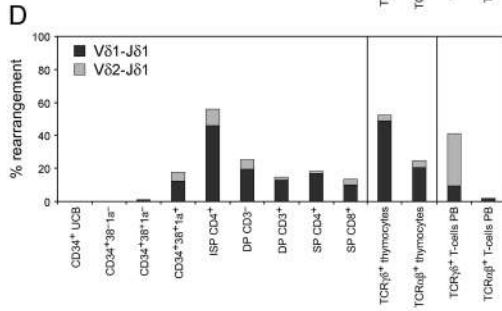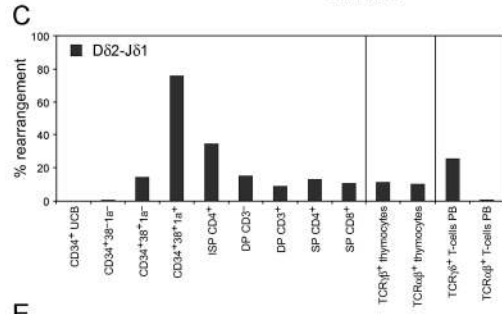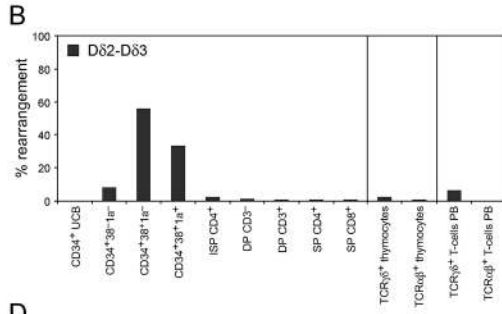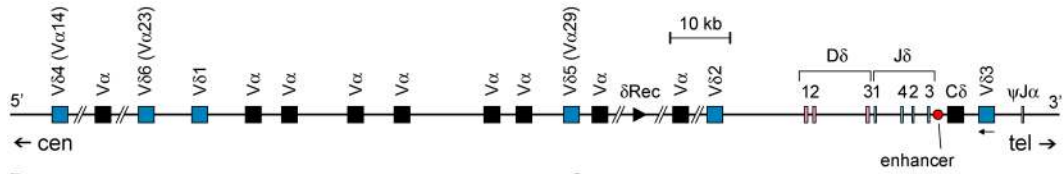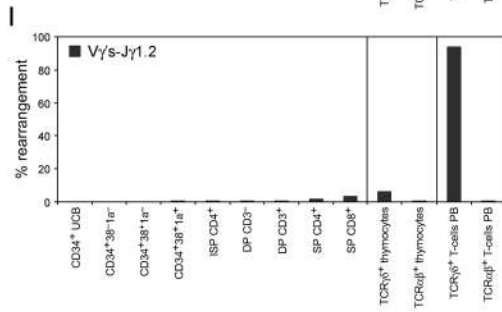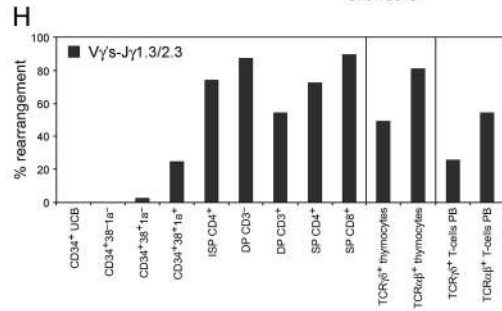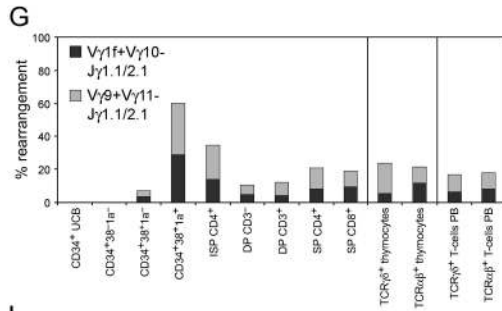
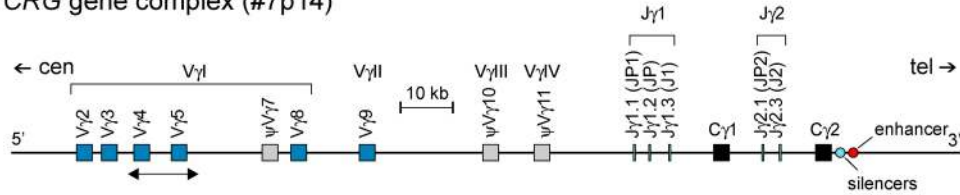**Figure 1. RQ–PCR analysis of *TCRD* and *TCRG* gene rearrangements in human T cell development and mature T cells.** (A) Schematic dia-

gram of the human *TCRD* gene complex. The six "classical" Vδ gene segments (blue) are scattered between Vα gene segments (black). Because

various TCR genes must be determined by the accessibility of gene segments to RAG enzymes. Evidence suggests that promoter and enhancer activity that is controlled by transcription factors regulate V(D)J recombination by modulating chromatin structures and rendering gene segments accessible to RAG cleavage (5, 6).

For obvious reasons T cell development mainly is studied in the mouse. Real-time quantitative PCR (RQ-PCR) and DNA microarray techniques allow careful analysis of small cell numbers. In this study we assessed the precise TCR gene configuration and the gene expression profiles of thymic subsets by RQ-PCR and Affymetrix DNA microarrays. By combining these two techniques we aimed for the identification of factors that play a role in regulating human TCR gene recombination.

RESULTS AND DISCUSSION
Definition of T cell populations
CD34$^+$ lineage negative "stem cell–like" cells were obtained from umbilical cord blood (UCB) and CD34$^+$CD38$^-$CD1a$^-$, CD34$^+$CD38$^+$CD1a$^-$, CD34$^+$CD38$^+$CD1a$^+$, ISP CD4$^+$, DP CD3$^-$, DP CD3$^+$, SP CD4$^+$, and SP CD8$^+$ subpopulations were obtained from thymi which represented consecutive stages of T cell development (Fig. S1, available at http://www.jem.org/cgi/content/full/jem.20042524/DC1). TCRαβ- and TCRγδ-expressing cells were obtained from thymic samples and peripheral blood mononuclear cells. Cells from five donors were pooled to reduce intrasample variation and all subsets were isolated twice from different donor pools.

Determination of TCR gene rearrangements by RQ–PCR and GeneScan analysis
TCR gene rearrangement analysis was performed in duplicate and on the two independently purified subsets (average is shown). The primers and TaqMan probes are listed in Table S1 (available at http://www.jem.org/cgi/content/full/jem.20042524/DC1) and do not amplify germline DNA.

TCRD
The Dδ2-Dδ3 rearrangement was detected first in the earliest CD34$^+$CD38$^-$CD1a$^-$ thymic subset, much earlier during development than in our previous Southern blotting–based report (4). Dδ2-Dδ3 rearrangements reached maximum levels in the subsequent CD34$^+$CD38$^+$CD1a$^-$ stage and declined to low levels in the later subsets (Fig. 1 B). The initial Dδ2-Dδ3 wave was followed by Dδ2-Jδ1 rearrangements which were detectable at low levels in CD34$^+$CD38$^-$CD1a$^-$ cells, increased in CD34$^+$CD38$^+$CD1a$^-$ cells, and peaked in CD34$^+$CD38$^+$CD1a$^+$ cells (Fig. 1 C). Complete Vδ1-Jδ1 or

Vδ2-Jδ1 rearrangements were detected first in CD34$^+$CD38$^+$CD1a$^-$ cells, increased to peak levels in CD4$^+$ ISP cells, and declined thereafter when the majority of thymocytes differentiated further into the TCRαβ lineage (Fig. 1 D). During TCRαβ T cell lineage development, the TCRD gene is deleted from the TCRA/D gene complex (2, 7) which results in the formation of a T cell receptor excision circle (TREC) that may contain Vδ–Jδ1 rearrangements. TRECs do not replicate on cell division; consequently, they are diluted out in proliferating, developing T cells (2). Ki67 staining demonstrated high percentages of proliferating cells within the human ISP and DP CD3$^-$ thymic subsets (unpublished data). This proliferation likely accounts for the observed decline in Vδ–Jδ1 rearrangements after the ISP stage. The Vδ1 gene segment seemed to be used preferentially in postnatal TCRγδ$^+$ thymocytes which is in sharp contrast to the well–described preferential Vδ2 usage by peripheral TCRγδ$^+$ T cells (Fig. 1 D; reference 8). We detected very low Vδ–Dδ3 levels in the thymic subsets; this represents a minor pathway to initiate TCRD gene rearrangement in postnatal thymus (Fig. 1 E).

TCRG
The first TCRG rearrangements (Vγ-Jγ1.1/2.1) were detected in the CD34$^+$CD38$^+$CD1a$^-$ thymic subset, earlier than we reported previously (4), and one stage after the initiation of TCRD rearrangement. These rearrangements peaked in CD34$^+$CD38$^+$CD1a$^+$ thymocytes and decreased to a relatively constant level in subsequent subsets (Fig. 1 G). Vγ-Jγ1.3/2.3 rearrangements were observed first in CD34$^+$CD38$^+$CD1a$^-$ cells after which they increased rapidly and by far exceeded Vγ-Jγ1.1/2.1 rearrangements from the ISP stage onward (Fig. 1 H). Peripheral TCRγδ$^+$ and TCRαβ$^+$ T cells revealed lower Jγ1.3/2.3 usage than their thymic counterparts (Fig. 1 H). In thymocytes, Jγ1.2 was used at very low frequency, but massive positive selection for the Jγ1.2 segment occurred in peripheral TCRγδ$^+$ T cells (Fig. 1 I; reference 8).

TCRB
Dβ1-Jβ1 rearrangements were detected first at low levels in the CD34$^+$CD38$^+$CD1a$^-$ population and increased thereafter from CD34$^+$CD38$^+$CD1a$^+$ to CD4+ ISP cells (Fig. 2 B). Dβ2-Jβ2 rearrangements were detected first at low levels in CD34$^+$CD38$^+$CD1a$^+$ cells, one differentiation stage after Dβ1-Jβ1 rearrangements (Fig. 2 B). The seemingly lower levels of Dβ-Jβ in DP CD3$^+$ probably are caused by variation within the lower range of detection of our assay and likely do not represent a true decrease.

Because of the complexity of the TCRB locus, Vβ-Jβ rearrangements were determined by using nonquantitative

Vδ4, Vδ5, and Vδ6 also are recognized as Vα gene segments, their Vα gene code is given in parentheses (adapted from reference 9). (B) Analysis of Dδ2-Dδ3 rearrangement. (C) Analysis of Dδ2-Jδ1 rearrangement. (D) Analysis of Vδ1-Jδ1 and Vδ2-Jδ1 rearrangements. (E) Analysis of Vδ-Dδ3 rearrangement. (F) Schematic diagram of the human TCRG gene complex.

Only the rearrangeable Vγ gene segments are depicted in blue (functional Vγ) or gray (nonfunctional Vγ). For the Jγ gene segments both nomenclatures are used (adapted from reference 9). (G) Analysis of Vγ to Jγ1.1 and Jγ2.1 rearrangements. (H) Analysis of Vγ to Jγ1.3 and Jγ2.3 rearrangements. (I) Analysis of Vγ to Jγ1.2 rearrangements.
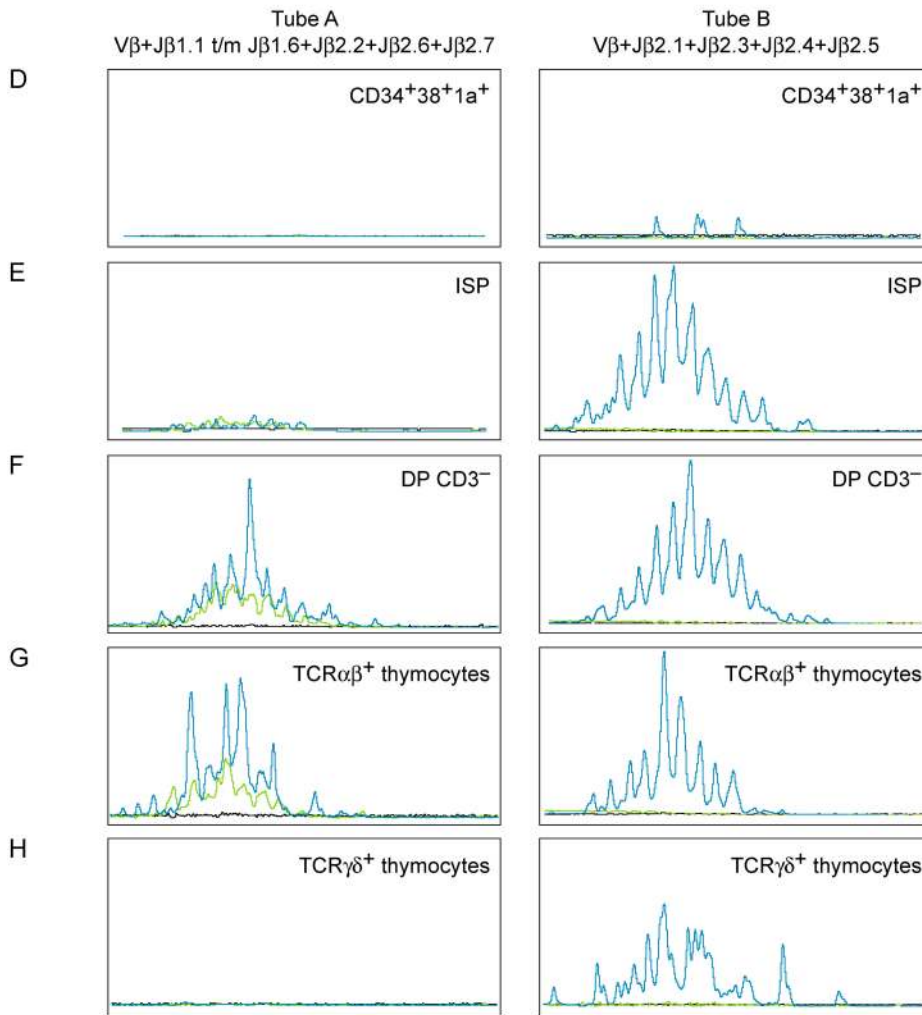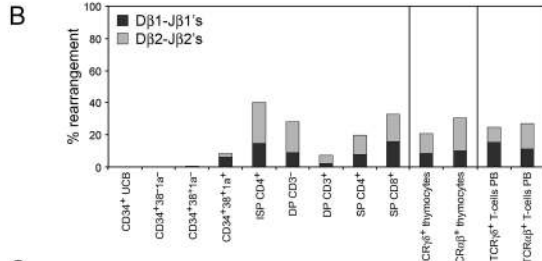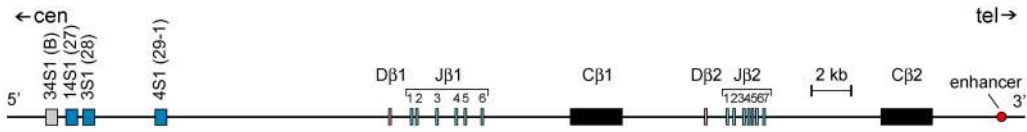
Figure 2.   RQ–PCR and GeneScan analysis of *TCRB* rearrangements in human T cell development and in mature T cells. (A) Schematic diagram of the human *TCRB* gene complex. (B) Analysis of Dβ to Jβ rearrangements. (C) Schematic diagram of Vβ-Jβ rearrangements as deter-

GeneScan analysis (9). The CD34+CD38+CD1a+ thymocytes contained low levels of the first Vβ-Jβ rearrangements in which Jβ2 gene segments were used exclusively (Fig. 2 D), much earlier than we described before (4). From the ISP subset onward, Vβ-Jβ rearrangements were in-frame as shown by the triplet peaks (Fig. 2 E) and were retained throughout all subsequent stages of development (Fig. 2, E–H). Previous reports suggested that TCRβ-selection in humans is initiated at the ISP/DP stages of T cell development (4, 10, 11), but the three-nucleotide spacing of the peaks in our GeneScan analysis suggest that selection for in-frame *TCRB* already occurs at the transition from CD34+CD38+CD1a+ to ISP. From the DP CD3− fraction, Vβ-Jβ1 rearrangements also were present, although they were less abundant than Vβ-Jβ2 rearrangements (Fig. 2 F; tube A). For Vβ-Jβ rearrangements, usage of Jβ2 gene segments was preferred over Jβ1 (Fig. 2, D–H). TCRγδ+ thymocytes contained no Vβ-Jβ1 rearrangements but did contain Vβ-Jβ2 rearrangements (Fig. 2 H).

TCRA

Because of large numbers of rearrangable Vα (∼54) and Jα (61) gene segments (2), we could not design a multiplex RQ-PCR for reliable quantification of all Vα-Jα rearrangements. Instead, we aimed for an alternative approach in which we used different indirect measures to study *TCRA* recombination. *TCRA* recombination is initiated by the transcription of T-early α (TEA) in order to open the 5′ site of the Jα cluster, which is followed by *TCRD* deleting rearrangements, particularly the δREC-ψJα rearrangement. These initiating events are followed by multiple, consecutive Vα-Jα rearrangements (7, 12).

To study initiation of *TCRA* rearrangement we determined the level of TEA-Cα transcripts as well as the occurrence of δREC-ψJα rearrangements. TEA-Cα transcripts started to increase in CD34+CD38+CD1a+ cells and reached peak levels in ISP and DP cells after which they declined again (Fig. 3 B). δREC-ψJα rearrangements were detected first in ISP cells and reached peak levels in SP and TCRαβ+ thymocytes (Fig. 3 C). These data show that *TCRA* rearrangement already has started in the ISP cell population but that there are still cells within the CD3+ DP population that start rearrangement of the (most likely) second *TCRA* allele. Although TEA-Cα transcripts and δREC-ψJα rearrangements are good measures for initiation of *TCRA* rearrangement, they cannot be used for quantification of the actual *TCRA* rearrangements. TEA-Cα is an mRNA product that cannot be extrapolated simply to the actual level of *TCRA* rearrangements. Quantification of δREC-ψJα is complex because it is influenced strongly by ongoing Vα-Jα rearrangements and the consequently produced TRECs (containing δREC-ψJα), whereas the amount of TRECs (and their dilution) is heavily dependent on the fraction of proliferating cells within specific subsets. Therefore, extra accumulation of TRECs may explain the relatively high δREC-ψJα levels in nonproliferating SP cells as compared with the preceding proliferating stages.

In an attempt to quantify *TCRA* recombination, we determined loss of germline *TCRA* DNA based on the disappearance of germline ψJα as an indirect measure for *TCRA* rearrangements. For this approach we used DNA from CD34+ lin− UCB cells as 100% germline reference. Germline *TCRA* clearly declined from the CD34+CD38+CD1a+ stage onward with a major decline when ISP cells progress toward DP cells (Fig. 3 D). The low levels of germline *TCRA* in DP CD3− cells indicate that extensive *TCRA* rearrangement has occurred at this stage. The apparent inconsistency with the relatively low levels of δREC-ψJα in these DP CD3− cells can be explained by the fact that TRECs that contain δREC-ψJα are diluted-out rapidly in these heavily proliferating cells and that other *TCRD* deleting rearrangements can occur as well, such as δREC-Jα58 and Vδ-ψJα (7). Our approach is only suitable for quantification of major decreases in germline *TCRA*, which occur at early stages of T cell development (ISP/CD3- DP); however, it is much more difficult to measure additional *TCRA* rearrangements accurately at the later stages of T cell development because germline ψJα (e.g., due to δREC-Jα58 rearrangements) remains detectable on TRECs and the difficulty of detecting further decreases within the 5–10% germline *TCRA* that is present in these stages.

We conclude that *TCRA* rearrangements are initiated when thymocytes progress from CD34+CD38+CD1a+ toward the ISP stage—which is much earlier than reported previously (10)—and apparently are ongoing until the CD3+ DP stage.
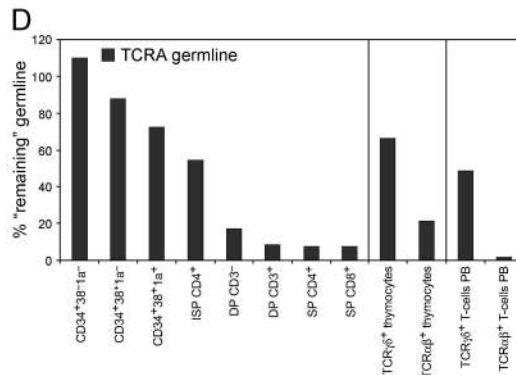
Microarray analysis

A total of 3848 probe sets underwent a significant change between any two successive stages of differentiation. Raw microarray data can be found at http://www.ebi.ac.uk/miamexpress, MIAME accession no. E-MEXP-337, and http://franklin.et.tudelft.nl, including a gene search browser. The expression levels of these probe sets were used to calculate a correlation coefficient between all possible pairs of microarrays and revealed high correlation between biological repeats (Fig. S2, available at http://www.jem.org/cgi/content/full/jem.20042524/DC1). This allowed us to use the average expression values of the two arrays that were performed per subset (obtained from five pooled thymi) for further analysis. Hierarchical clustering of the 3848 probe sets was performed and is described in Fig. S3 and Table S2 (available at http://www.jem.org/cgi/content/full/jem.20042524/DC1).

A *TCR A/D* gene complex (#14q11.2)
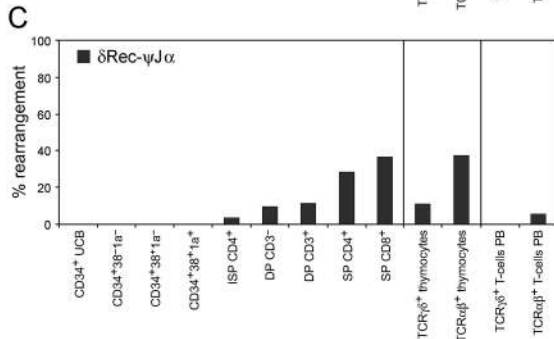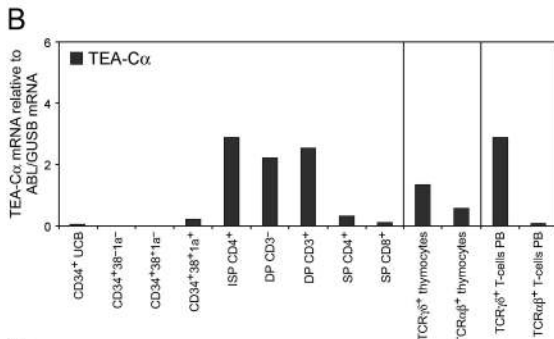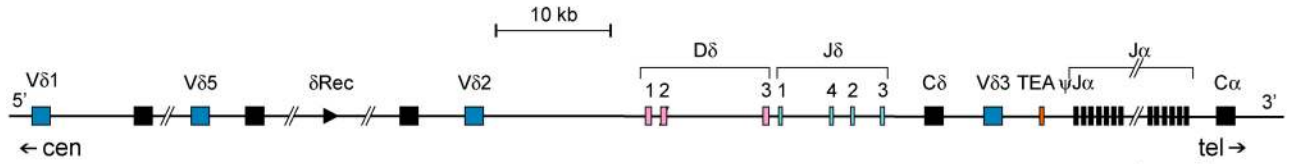
Figure 3. RQ-PCR analysis of initiating events around *TCRA* rearrangement in human T cell development and mature T cells. (A) Schematic diagram of the human *TCRA* gene complex. The TEA element (red) forms a sterile mRNA with Cα sequences. The *TCRD* deleting elements δRec and ψJα are also indicated. (B) Analysis of TEA-Cα mRNA expression. Due to shortage of material, TEA expression was only determined once. (C) Analysis of the *TCRD* deleting rearrangement δRec-ψJα. (D) Analysis of "remaining" *TCRA* germline DNA.

Of special interest for TCRβ-selection and initiation of *TCRA* rearrangement is that *PTCRA* (pTα) expression increased in the CD34⁺CD38⁺CD1a⁻ stage, peaked in the CD34⁺CD38⁺CD1a⁺ and ISP stages, after which it declined (Table S2). Mouse microarray data (13) have shown a similar expression pattern of pTα, with a minor peak at DN3 and a larger peak at the DP CD3⁻ stage. Experiments with pTα mutant mice indicate that TCRβ-selection in the mouse occurs at DN3 (14), and that *TCRA* recombination is initiated after TCRβ-selection has occurred (14). Here, we show that initiation of *TCRA* recombination starts in CD34⁺CD38⁺CD1a⁺ cells. Therefore, the analogous pTα expression between mice and men, our *TCRB* GeneScan, and our *TCRA* recombination data indicate that human TCRβ-selection occurs at the CD34⁺CD38⁺CD1a⁺ stage instead of the previously suggested ISP/DP stage (4, 10, 11).

To identify candidate transcription factors that are involved in regulating TCR recombination, the list of 3848 probe sets was filtered based on Gene Ontology annotation

(transcriptional regulation and DNA binding) which yielded a final list of 446 probe sets that encoded a total of 361 genes (Fig. S4, available at http://www.jem.org/cgi/content/full/jem.20042524/DC1). Expression of genes that are associated with T cell commitment/differentiation and/or V(D)J recombination, such as *NOTCH1*, *HES1*, *GATA3*, *BCL11B*, *RAG1*, *RAG2*, and *DNTT* (TdT), increased strongly in early T cell differentiation.

To determine which genes may have a role in regulating TCR rearrangement, hierarchical cluster analysis was performed (Fig. 4 A). The 446 probe sets were divided into 15 clusters; the prototypic expression patterns are depicted in Fig. 4 B. Based on the TCR rearrangement patterns (as described above) and the prototypic gene expression patterns, we propose that clusters 3, 5, 6, 7, 14, and 15 (Fig. 4 B, asterisk) contain genes that may encode candidate factors for initiation/regulation of TCR rearrangements because the genes that are present in these clusters show higher expression at the moments at which active TCR gene rearrangement occurs. Genes that are
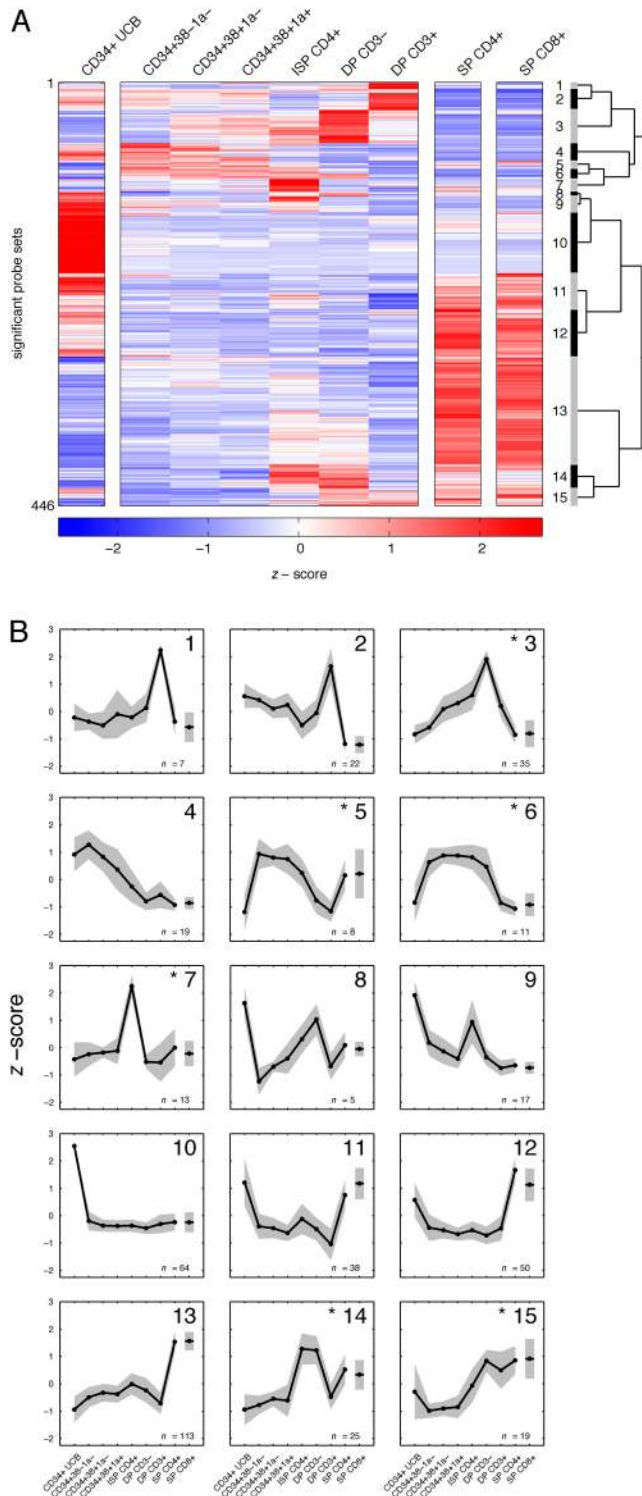
Figure 4. Analysis of probe sets related to transcriptional regulation. (A) Hierarchical clustering of 446 probe sets differentially expressed over the various stages and related to transcriptional regulation and DNA binding. (B) z-score trend representation for each cluster deduced from the 446 probe sets. Cluster codes (1–15) are shown in the top right corner, z-score on the y axis, subsets on the x axis. The gray area represents 1 SD. The asterisk in the top right corner indicates clus-

present in clusters 3, 5, 6, 7, 14, and 15 are presented in Table I; those of all 15 clusters are shown in Table S3 (available at http://www.jem.org/cgi/content/full/jem.20042524/DC1).

We propose that clusters 3, 5, and 6 contain genes that are common for the rearrangement of all four TCR genes. Cluster 5 likely contains genes that are important for the regulation of *TCRD*, *TCRG*, and *TCRB* rearrangements, whereas cluster 3 likely contains genes that need to be expressed at a higher level when *TCRA* rearranges. These clusters contain *DNTT* (cluster 6), *RAG 1/RAG 2* (cluster 3), and genes that encode factors, such as NOTCH1 (cluster 6), RORC (cluster 3), SMARCA4 (BRG1) (clusters 6 and 3), H2AFX (cluster 3), that previously were linked to regulation of *TCRG*, *TCRB*, or *TCRA* rearrangements (15–18).

Genes in cluster 7, 14, and 15, such as *SPIB*, *ICSBP1*, *TCF4*, *CREB1*, *ETS1*, and *LEF-1*, may encode factors that are involved in regulating *TCRA* rearrangements as well as allelic exclusion of the *TCRB* locus. These are discussed in the supplemental Results and Discussion (available at http://www.jem.org/cgi/content/full/jem.20042524/DC1).

Novel insights into human T cell development

We confirm that TCR loci rearrange in a highly ordered way (*TCRD-TCRG–TCRB-TCRA*) and defined sequential rearrangement steps of *TCRD*, *TCRG*, *TCRB*, and initiation of *TCRA* recombination to specific human thymic subsets. Importantly, our data show that recombination of the TCR genes occurs earlier during human T cell development than previously reported (4, 10). Given that *TCRD* rearrangement starts at DN1 in mice, followed by *TCRG* in DN2 and *TCRB* in DN2, but especially DN3 (1, 19), the human CD34+CD38−CD1a−, CD34+CD38+CD1a−, and CD34+CD38+CD1a+ subsets resemble murine early DN1 (CD44+CD25−CD117−), late DN1/DN2 (CD117+ DN1; CD44+CD25+), and DN3 (CD44−CD25+) stages, respectively. However, the relative frequency of DN1 cells in mice is higher than that of the corresponding human subset (CD34+CD38−CD1a−). We also demonstrate that TCRβ-selection and initiation of *TCRA* rearrangement already occur at the CD34+CD38+CD1a+ stage of human T cell development, instead of the ISP/DP stage (4, 10, 11), similar to the mouse (i.e., TCRβ-selection occurs at DN3 in the mouse; reference 14).

Based on the TCR rearrangement data and the expression profile of key recombination and differentiation genes (e.g., *RAG1*, *RAG2*, and *PTCRA*), we show that human and mouse T cell development are much more similar than assumed previously. In addition, candidate factors for regulation of TCR recombination are identified.

We propose an updated human T cell differentiation model as shown in Fig. S5 (available at http://www.jem.org/cgi/content/full/jem.20042524/DC1). These novel data help

ters that contain genes that are potentially important in regulating TCR rearrangements. N, indicates the number of probe sets in a cluster.

Table I. Transcription and DNA-binding–related genes with expression profiles that correlate with TCR gene rearrangements in human T cell development

Category	Common TCR	Common TCR D, G, B > A	Common TCR A > B, G, D	TCRA/allelic exclusion TCRB		
	Cluster 6	Cluster 5	Cluster 3	Cluster 7	Cluster 14	Cluster 15
Transcription factor/ coregulator of transcription	BAZ1B	BAZ2B	AEBP1	ACYP2	BRD8	ABCA7
	ETS2	FOS	CBFA2T3	BCL11A	CNOT3	ARNTL
	HMGB3	HES1	E2F1	CD86	CREB1	BCL6
	KLF4	IRF1	FOXM1	CUTL2	FALZ	CENTB1
	MYB	STAT4	GFI1	ICSBP1 (IRF8)	IFI16	ELF1
	NOTCH1	TCEAL1	MYBL1	IRF7	ILF3	ETS1
		TCFL1	MYBL2	MNDA	LEF1	GATA3
			NFATC3	MYCL1	LEREPO4	GMEB2
			NOTCH3	SPIB	M96	HTATIP
			ORC6L	TCF4	NFYA	LEF1
			RORC		NR2C1	LRRFIP1
			SAP30		PWP1	MXD4
			TFDP2		SOLH	TCF7
			TMPO		TAF5	ZNF297B
			TRIP13		TBDN100	
			TTF1		TCF12 (HEB)	
			WHSC1		TCF8	
			ZNF423		TFAM	
					ZNFN1A1 (IKAROS)	
Recombination	DNTT (TdT)		RAG1			
			RAG2			
Chromatin remodeling	H2AFY		CHAF1A		HDAC4	HDAC7A
	SMARCA4		EZH2		YY1	SATB1
			H2AFX			
			HMGB2			
			HMGN2			
			SMARCA4			
DNA repair	MCM2		BRCA1		BTG2	BTG2
	MCM6		MCM7			
			PTTG1			

to bridge gaps in our understanding of human T cell development, and also should provide insight into the development of T cell acute lymphoblastic leukemias and T-SCID.

MATERIALS AND METHODS

Purification of thymocyte subsets, CD34+ lineage− cells, and mature peripheral T cells. For the isolation of CD34+ lin− UCB cells and thymocyte subsets, total mononuclear cells or thymocytes from five donors were pooled to reduce intrasample variation. CD34+ subsets were prepurified by magnetic-activated cell-sorted beads and further purified by cell sorting to >96% purity.

RQ-PCR analyses of TCR gene rearrangements. This essentially was performed as described previously (20). For the various TCR recombinations, cell lines/diagnostic samples were selected as clonal control DNA.

GeneScan analysis for complete and in-frame Vβ-Jβ gene rearrangements. This was done by two multiplex PCR reactions as described by the BIOMED-2 Concerted Action (9).

TEA expression by RQ-PCR analyses. Expression of TEA was defined by calculating the ratio of TEA to the average value of the control genes ABL and GUSB.

Microarray analysis. Affymetrix microarray analysis essentially was done as described previously (21) and according to Minimum Information About a Microarray Experiment guidelines (www.mged.org/Workgroups/MIAME/miame.html).

Statistical analysis. Probe intensity background was removed using robust multichip analysis (22). The intensity levels were quantile normalized

(23). Array groups (two biologically independent arrays per group) that corresponded to the development stages were compared based on the perfect match probe intensity levels only (22), by performing a per-probe set two-way analysis of variance (with factors "probe" and "stage"). The p-values were adjusted for multiple testing using Šidák step-down adjustment (24) and all differences with adjusted p-values <0.05 were considered to be significant. For Fig. S2, expression values were $\log_2$-transformed and the per-probe set geometric mean was subtracted. Further analysis was performed using Genlab software, running under Matlab (http://www.genlab.tudelft.nl). After per-probe set normalization to zero mean and unit standard deviation (z-score), a hierarchical clustering (complete linkage) was calculated based on Pearson correlation. The number of clusters k was determined by looking for a local minimum of the Davies-Bouldin index calculated for $k = 1, \ldots, 30$ (25).

Online supplemental material. All methods are described in more detail as supplemental Materials and methods online. Additional Results and Discussion, as well as Figs. S1–S5 and Tables S1–S3, are provided online, as well. Online supplemental material is available at http://www.jem.org/cgi/content/full/jem.20042524/DC1.

REFERENCES

1. Staal, F.J., F. Weerkamp, A.W. Langerak, R.W. Hendriks, and H.C. Clevers. 2001. Transcriptional control of T lymphocyte differentiation. *Stem Cells*. 19:165–179.

2. van Dongen, J.J.M., F.J. Staal, and A.W. Langerak. 2004. Developmental and functional biology of T lymphocytes. *In* Non-Hodgkin's Lymphoma's. P.M. Mauch, J.O. Armitage, B. Coiffier, R. Dalla-Favera, and N.L. Harris, editors, Lippincott, Williams & Wilkins, Philadelphia. 787-808.

3. Asnafi, V., K. Beldjord, E. Boulanger, B. Comba, P. Le Tutour, M.H. Estienne, F. Davi, J. Landman-Parker, P. Quartier, A. Buzyn, et al. 2003. Analysis of TCR, pT alpha, and RAG-1 in T-acute lymphoblastic leukemias improves understanding of early human T-lymphoid lineage commitment. *Blood*. 101:2693–2703.

4. Blom, B., M.C. Verschuren, M.H. Heemskerk, A.Q. Bakker, E.J. van Gastel-Mol, I.L. Wolvers-Tettero, J.J.M. van Dongen, and H. Spits. 1999. TCR gene rearrangements and expression of the pre-T cell receptor complex during human T-cell differentiation. *Blood*. 93:3033–3043.

5. Schlissel, M.S. 2003. Regulating antigen-receptor gene assembly. *Nat. Rev. Immunol*. 3:890–899.

6. Langerak, A.W., I.L. Wolvers-Tettero, E.J. van Gastel-Mol, M.E. Oud, and J.J.M. van Dongen. 2001. Basic helix-loop-helix proteins E2A and HEB induce immature T-cell receptor rearrangements in nonlymphoid cells. *Blood*. 98:2456–2465.

7. Verschuren, M.C., I.L. Wolvers-Tettero, T.M. Breit, J. Noordzij, E.R. van Wering, and J.J.M. van Dongen. 1997. Preferential rearrangements of the T cell receptor–delta–deleting elements in human T cells. *J. Immunol*. 158:1208–1216.

8. Breit, T.M., I.L. Wolvers-Tettero, and J.J.M. van Dongen. 1994. Unique selection determinant in polyclonal V delta 2-J delta 1 junctional regions of human peripheral gamma delta T lymphocytes. *J. Immunol*. 152:2860–2864.

9. van Dongen, J.J.M., A.W. Langerak, M. Bruggemann, P.A. Evans, M. Hummel, F.L. Lavender, E. Delabesse, F. Davi, E. Schuuring, R. Garcia-Sanz, et al. 2003. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*. 17:2257–2317.

10. Spits, H. 2002. Development of alphabeta T cells in the human thymus. *Nat. Rev. Immunol*. 2:760–772.

11. Carrasco, Y.R., C. Trigueros, A.R. Ramiro, V.G. de Yebenes, and M.L. Toribio. 1999. Beta-selection is associated with the onset of CD8beta chain expression on CD4(+)CD8alphaalpha(+) pre-T cells during human intrathymic development. *Blood*. 94:3491–3498.

12. Villey, I., D. Caillol, F. Selz, P. Ferrier, and J.P. de Villartay. 1996. Defect in rearrangement of the most 5′ TCR-J alpha following targeted deletion of T early alpha (TEA): implications for TCR alpha locus accessibility. *Immunity*. 5:331–342.

13. Hoffmann, R., L. Bruno, T. Seidl, A. Rolink, and F. Melchers. 2003. Rules for gene usage inferred from a comparison of large-scale gene expression profiles of T and B lymphocyte development. *J. Immunol*. 170:1339–1353.

14. von Boehmer, H., and H.J. Fehling. 1997. Structure and function of the pre-T cell receptor. *Annu. Rev. Immunol*. 15:433–452.

15. Wolfer, A., A. Wilson, M. Nemir, H.R. MacDonald, and F. Radtke. 2002. Inactivation of Notch1 impairs VDJbeta rearrangement and allows pre-TCR–independent survival of early alpha beta lineage thymocytes. *Immunity*. 16:869–879.

16. Villey, I., R. de Chasseval, and J.P. de Villartay. 1999. RORgammaT, a thymus-specific isoform of the orphan nuclear receptor RORgamma / TOR, is up-regulated by signaling through the pre-T cell receptor and binds to the TEA promoter. *Eur. J. Immunol*. 29:4072–4080.

17. Morshead, K.B., D.N. Ciccone, S.D. Taverna, C.D. Allis, and M.A. Oettinger. 2003. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. *Proc. Natl. Acad. Sci. USA*. 100:11577–11582.

18. Chen, H.T., A. Bhandoola, M.J. Difilippantonio, J. Zhu, M.J. Brown, X. Tai, E.P. Rogakou, T.M. Brotz, W.M. Bonner, T. Ried, and A. Nussenzweig. 2000. Response to RAG-mediated VDJ cleavage by NBS1 and gamma-H2AX. *Science*. 290:1962–1965.

19. Capone, M., R.D. Hockett Jr., and A. Zlotnik. 1998. Kinetics of T cell receptor beta, gamma, and delta rearrangements during adult thymic development: T cell receptor rearrangements are present in CD44(+)CD25(+) Pro-T thymocytes. *Proc. Natl. Acad. Sci. USA*. 95:12522–12527.

20. Langerak, A.W., B. Nadel, A. De Torbal, I.L. Wolvers-Tettero, E.J. van Gastel-Mol, B. Verhaaf, U. Jager, and J.J.M. van Dongen. 2004. Unraveling the consecutive recombination events in the human IGK locus. *J. Immunol*. 173:3878–3888.

21. Staal, F.J., F. Weerkamp, M.R. Baert, C.M. van den Burg, M. van Noort, E.F. de Haas, and J.J.M. van Dongen. 2004. Wnt target genes identified by DNA microarrays in immature CD34+ thymocytes regulate proliferation and cell adhesion. *J. Immunol*. 172:1099–1108.

22. Irizarry, R.A., B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, and T.P. Speed. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 4:249–264.

23. Bolstad, B.M., R.A. Irizarry, M. Astrand, and T.P. Speed. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*. 19:185–193.

24. Ge, U., S. Dudoit, and T.P. Speed. 2003. Resampling-based multiple testing for microarray analysis. *Test*. 12:1–77.

25. Davies, D.L., and D. Bouldin. 1979. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell*. 1:224–227.