A molecular chart of thymocyte positive selection

Margarida Correia-Neves, Diane Mathis and Christophe Benoist

Institut de Génétique et de Biologie Moléculaire et Cellulaire (CNRS/INSERM/ULP), Illkirch, C.U. de Strasbourg, France

As a new slant on T lymphocyte repertoire selection, we have examined batteries of TCR sequences in thymi from transgenic mice engineered to exhibit limited, focussed TCR diversity. We have tracked the fate of differentiating thymocytes expressing a set of particular TCR through the positive selection process. Subtly different TCR sequences can promote different maturation pathways and commitment choices. Two distinct routes are followed by CD8-lineage cells interacting with MHC class I molecules, via TCR^{hi} $CD4^+CD8^+$ or $CD4^+CD8^{int}$ intermediates, while CD4-lineage cells mature exclusively via a $CD4^+CD8^{int}$ stage. The CD8-lineage routes are partially exclusive, indicating that the latter cell type is not always preceded by the former. The distribution of sequences also indicates that CD4/CD8-lineage commitment is not strictly correlated with the class of MHC molecule engaged, and that some mechanism prevents mismatched intermediates from achieving full maturity.

Received	26/4/01
Accepted	15/6/01

Key words: T cell differentiation / TCR / T cell repertoire / TCR transgenic mice / Recombination

1 Introduction

Positive selection is a key step in thymocyte differentiation. It occurs after productive rearrangement of TCR α and β chain genes, and results from the interaction between TCR expressed on thymocytes and MHC molecules displayed by thymic stromal cells. It coincides with the transition from immature and uncommitted $CD4^+CD8^+$ (double-positive; DP) cells to mature, committed $CD4^+CD8^-$ and $CD4^-CD8^+$ (single-positive; SP) cells. Its outcomes are that the thymocyte survives and is allowed to continue to mature, and there is a close match between the phenotype of the mature cell and the class of MHC molecule with which its TCR can interact: $CD4^+$ T cells recognize MHC class II molecules, and $CD8^+$ T cells, MHC class I molecules.

In cellular terms, positive selection is a circuitous process, with multiple transitional intermediates appearing during the several days needed to complete the process. Passage from the DP to the SP stage does not simply involve shutting off one or the other coreceptor, but proceeds through complex permutations. For example, transitional populations of $CD4^+CD8^{int}$ phenotype contain precursors of both $CD4^+CD8^-$ and $CD4^-CD8^+$ cells [1, 2]. These complexities most likely reflect the regulation of coreceptor gene transcription – activation of enhancer elements specific to mature cells, binding of factors that silence the CD4 locus [3]. Other phenotypic changes also mark the DP to SP transition – for example, surface levels of TCR/CD3 complexes increase and the activation marker CD69 is induced [4].

How the match between TCR specificity and CD4/CD8 lineage is achieved remains a debated question [5–7]. Clearly is does not simply result from a direct sensing of the class of MHC molecule by the receptor/coreceptor complex, which would instruct lineage choice, because cells can mature in the absence of coreceptor, and under particular circumstances with mismatched coreceptors [8–13]. Yet it is not a random process either, and recent data indicate that quantitative and/or temporal aspects of TCR engagement may be the primary factor driving lineage choice. Recently, direct evidence was provided that the degree of *lck* activity elicited through TCR engagement can condition the outcome of selection [14], as had been hypothesized previously [12, 15]. The duration of a strong TCR engagement (1.5 *vs.* 14 h) also had a decisive influence on lineage choice in an *in vitro* system [16] – an observation one cannot immediately reconcile with that above.

The analysis of TCR transgenic mice has demonstrated that TCR clonotypes can promote quite different pathways of differentiation [17]. The final outcome varies, lineage matching TCR preference for MHC molecules. But the relative proportion of transitional intermediates also diverges between transgenic lines of similar class specificity – for example, thymi from different class I-restricted TCR transgenic lines show variable proportions of $CD4^+CD8^{int}$ intermediates ([18] and references

[I 21939]

Abbreviations: DP: Double positive **SP:** Single positive **DN:** Double negative **βgal:** β-galactosidase

 0014-2980/01/0909-2583$17.50+.50/0

therein). Such differences have also been observed when the same TCR transgene was bred onto backgrounds carrying different MHC alleles [19]. These observations suggest that variations in TCR affinity/avidity for selecting MHC : peptide ligands underlie distinct selection pathways. It is not possible, however, to rule out that the differences observed between transgenic lines are not artifactually due to variations in the timing and intensity of TCR transgene expression.

Therefore, it seemed desirable to track the differentiation pathways of T cells expressing TCR derived from the same transgene. To this end, we employed a line of transgenic mice carrying a TCR minilocus described recently [20]. This Va2Var transgene requires rearrangement to be expressed functionally, leading to a very focused V-J junctional diversity in TCRα chains, which otherwise use identical Vα and Jα elements, and always pair with the same Vβ chain encoded by a TCRβ transgene. From the diverse set of sequences generated by random rearrangement of the transgene in immature thymocytes, positive selection exaggerates the contribution of a restricted set of thirty or so dominant sequences (up to 20 % of mature thymocytes). Each is reproducibly selected into either the CD4 or CD8 lineages, indicative of restriction by class II or class I MHC molecules [20]. Because these repeated sequences make up such a significant proportion of the repertoire after positive selection, it is possible to trace their presence in transitional intermediates, thereby charting the selection pathways of cells bearing highly related TCR, differing only in their recognition of self-MHC : peptide ligands. The approach is conceptually similar to that of Sant'Angelo et al. [21], who charted the traffic of a dominant sequence through positive selection in single-transgene mice. It affords a broader view, leading to novel conclusions.

2 Results

The "Limited" mouse line that is the focus of this study is described in detail elsewhere [20]. Briefly, these animals carry a TCRα minilocus, called Va2Var, with a single Vα and two Jα segments separated by RSS motifs and an artificial 500-bp spacer (the Vα2 and Jα26 sequences, derived from the classical K^b-restricted OT-1 TCR [22, 23] and the natural Jα2 contributed by the vector [24]). The VJ joining process converts this template to a functional TCRα gene. A null mutation of the endogenous TCRα locus has been bred into these mice, ensuring that TCRα chains can only originate from productive rearrangements at the transgenic minilocus. Finally, these animals also carry a TCRβ transgene (Vβ5 from the same OT-1 TCR) which is expressed in a quasi-exclusive fashion, inhibiting the rearrangement of endogenously encoded TCRβ elements [25]. Thus, in the Limited mice, all mature T cells express the product of the rearranged minilocus coupled with the transgene-encoded Vβ5 chain. TCR variability is thus restricted to the junctional CDR3α region, using either one of the two Jα elements. Previous analyses have established that TCR restricted to both class I and class II MHC molecules are generated in Limited mice, promoting the generation of lymphocytes of both the $CD8^+$ and $CD4^+$ lineages.

We first verified that thymocyte maturation followed the usual pathways. Since the Va2Var transgene is not functional without RAG-mediated rearrangement, its expression would be expected to follow normal patterns, in particular avoiding the precocious expression in immature $CD4^-CD8^-$ double negatives (DN) that is commonly observed in TCR transgenic mice. This was confirmed in the analysis of Fig. 1 A: while no T cells matured in the TCRα-deficient mouse (Cα°), the presence of the Va2Var and OT-1β transgenes promoted the maturation of CD4 and CD8 SP the majority of which expressed Vα2 at densities comparable with those of wild-type mice (Fig. 1 A, middle panels). In contrast, there was little or no expression of the transgene in DN. These SP had normal maturation profiles, as evidenced by the four-color cytofluorimetric analyses of Fig. 1 B: most DP expressed no or little CD3, with only a minor proportion displaying higher CD3 (not quite SP levels) and lower PNAr. SP exhibited the highest level of CD3 and down-modulated PNAr in Limited and control mice. Similar concordance with wild-type patterns was observed when thymocytes from Limited mice were stained for CD24 or CD69 (not shown). As previously discussed [20], the number of cells that rearrange a productive TCRα and go on to maturation is lower in Limited mice than in wild-type animals, most likely because there is a single Vα rearrangement target instead of the normal array of Vα segments.

Having established that thymocytes mature normally in Limited mice, we then asked which CDR3α sequences from the Va2Var transgene could be found in the various thymic subpopulations.

As previously, we determined the representation of cells with particular CDR3α sequences by sorting single cells on a flow cytometer, amplifying the TCRα mRNA of each of these cells, and directly determining the sequence of the PCR product by high throughput sequencing at the Centre National de Sequençage (Evry, France). This strategy avoided possible artefacts from chimeric PCR and preferential amplification or cloning. Compilations of the resulting sequences are illustrated graphically here, and the complete sequence data are available as Supplemental Data http://research.joslin.harvard.edu/cbdm/introduction.html.

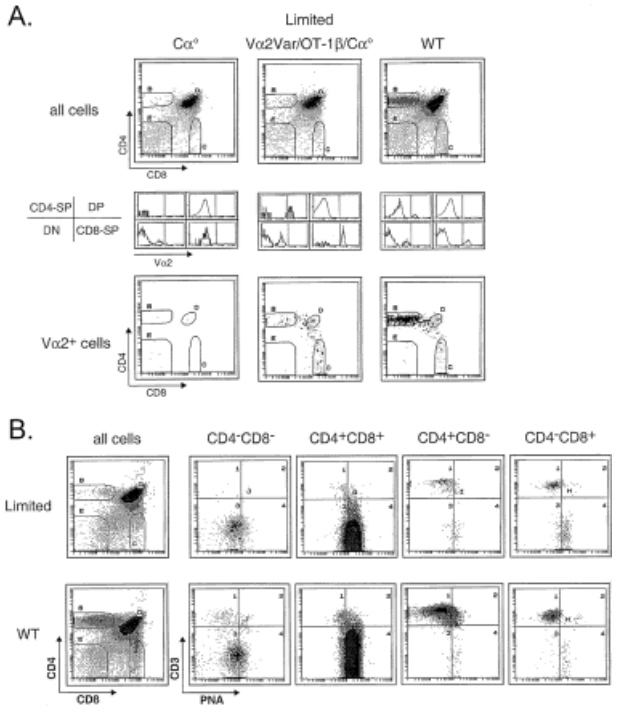

Fig. 1. Proper maturation of thymocytes from Limited mice. A. The transgene-encoded TCR is expressed in a timely fashion. Thymocytes from a TCRα-deficient mouse (Cαo), a Limited mouse, and a matched B6 control were stained with anti-CD4, -CD8 and -Vα2 (the Vα chain that corresponds to the Va2Var transgene). The top panels display the CD4/CD8 profiles of all thymocytes, the middle panels the Vα2 profiles of gated $CD4^-$ $CD8^-$ (DN), $CD4^+CD8^+$ (DP) and mature SP; note that a population of DN stains dully with the anti-Vα2 mAb, even in control Cαo mice, and must be caused by an artefactual cross-reactivity of the reagent or by reactivity to Vα2-like regions on TCRδ. The bottom panels display CD4/CD8 profiles gated on Vα2-high cells (gates as indicated in the middle panels). B. Normal maturation profiles. Phenotypic maturation was tested by staining with anti-CD4, -CD8, -CD3 and PNA-FITC. From the regions defined on the CD4/CD8 profiles of the left panels, gates were set on the different subpopulations and their PNAr/CD3 profiles displayed.

2.1 MHC class I- and class II-restricted TCR are found in the data set from transitional $CD4^+CD8^{int}$ thymocytes

The Shortman and Singer groups first showed that a transitional $CD4^+CD8^{int}$ population, expressing CD69 and high surface levels of TCR, harbored cells destined to become either CD4 and CD8 SP cells [2, 26]. Surprisingly, individual cells in the $CD4^+8^{int}$ population could express different transcriptional programs despite having similar surface phenotypes [2, 18]. Thus, it was of interest to see, from the perspective offered by the Limited mouse, whether sequences of both MHC class I- and class II-restricted TCR derived from this population, and whether all of the repeated sequences were found. To this end, thymocytes from three independent Limited mice were stained for CD4, CD8 and Vβ5; cells expressing a high level of Vβ5 were gated, and their CD4/CD8 profiles displayed (Fig. 2 A). From such stains, single $CD4^+CD8^{int}$ cells were individually sorted, and the sequences of their TCR determined after PCR amplification. As in previous analyses, the efficiency of this procedure was high (> 50 % of wells successfully amplified), such that the results should be a fair representation of the sequences actually present in the population.

Results on 196 cells appear in Fig. 2 B. The pattern is typical of a positively selected population, clearly different from that of pre-selection DP [20]: a proportion of the cells (31 %) expressed TCRα whose CDR3 motifs were found only once, but TCRs from the majority of cells (69 %) reflected a recurrent set of rearrangements of the minilocus, some of them at very high frequencies (up to 28 times). In contrast, these repeated sequences accounted for only 8 % of cells in the TCR^{lo} DP compartment (supplemental data; see also Fig. 6 A). Our previous analyses indicated that these repeated sequences could be divided into those consistently associated with CD4 lineage and those characteristic of CD8-lineage cells, in the thymus and the peripheral lymphoid organs. These are compiled in the coded graph of Fig. 2 C (white and striped, sequences restricted to the CD4 and the CD8 lineages, respectively; dark grey, sequences common to the two lineages; for clarity the CDR3 motifs are abbreviated, so that SD-Jα2 denotes SDDNYQL). It is clear that the $CD4^+CD8^{int}$ population harbors cells representative of both lineages. Furthermore, most of the frequently reiterated post-selection sequences were found in this population, although a few were missing, such as the LMD-Jα26 and STNT-Jα2 sequences. These results confirm the notion that the $CD4^+CD8^{int}$ population includes cells restricted by either MHC class and destined to either lineage.

2.2 TCR sequences from thymocytes of MHC class II-deficient mice

It seemed important at this stage to verify that the assignments of MHC restriction made on the basis of sequence distributions in mature cells were indeed correct. To this end, we introduced the MHC class II knockout mutation [27] into the Limited mouse line to generate $Va2Var^+Vβ5^+Ca^{o/o}Aβ^{o/o}$ individuals, and analyzed the sequences derived from single $CD4^+CD8^{int}$ and mature $CD4^-CD8^+$ cells. The results presented in Fig. 3 confirm

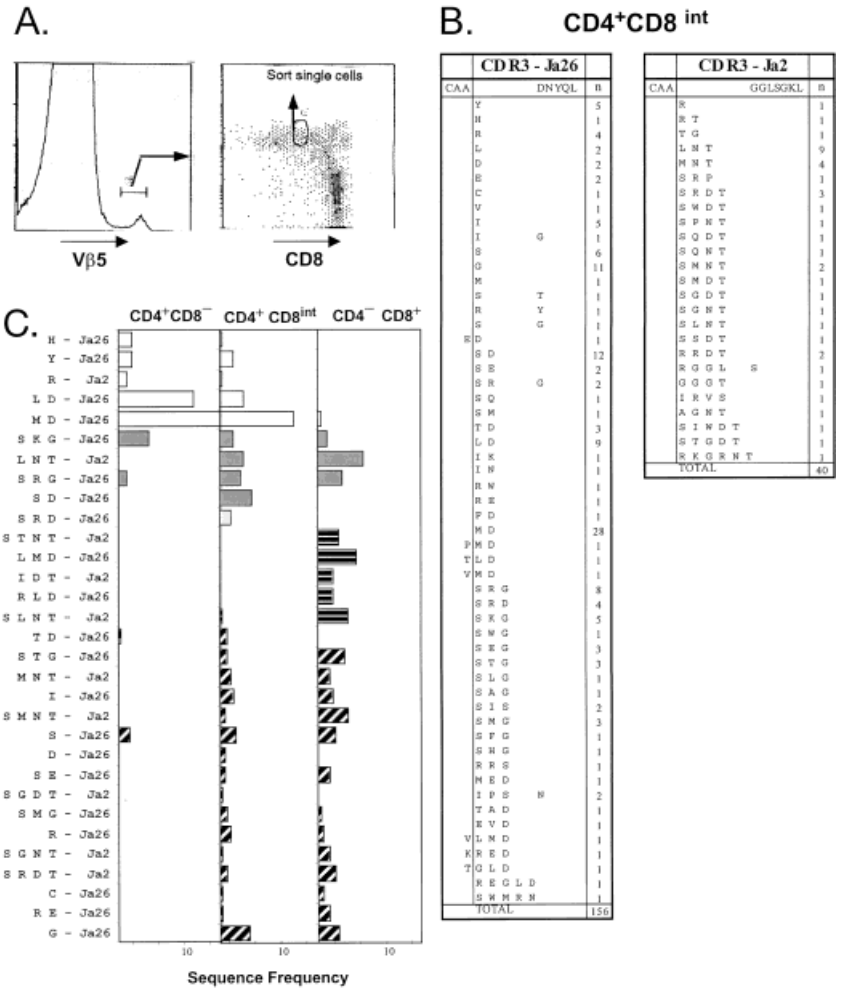

Fig. 2. CD4$^+$CD8^{int} cells harbor both MHC-I and -II restricted sequences, but not all of them. A: Cell sorting: Thymocytes from Limited mice were stained with anti-Vβ5, -CD4 and -CD8. Cells gated for high levels of transgene-encoded Vβ5 (left panel) were analyzed for CD4 and CD8 (right panel). Cells in the CD4$^+$CD8^{int} gate shown were sorted individually for RT-PCR and sequencing. B: CDR3 sequences of 196 individual CD4$^+$CD8^{int} cells. For clarity, the amino-acids common to all sequences are listed at the top, and only the distinguishing amino-acids shown for each; sequences utilizing Jα26 and Jα2 are listed in separate columns. The number of occurrences of the sequences in the data set is indicated on the right. C: Frequency histogram for the 31 oft-repeated sequences (defined as occurring 12 times or more in the combined data for all populations). For clarity, only the distinguishing amino-acids are shown before the indication of the Jα region used (eg. H-Ja26 stands for CAAHDNYQL, STNT-Ja2 for CAASTNTGGLSGKL). White and striped: sequences only observed in the CD4 and CD8 lineages, respectively. Dark grey: sequences found in both subsets. Light grey: the SRD-Ja26 sequence observed in CD4$^+$CD8^{int} cells from a number of mice, but never in the mature populations (blocked by negative selection?).

these assignments: the sequences thought to come from MHC class I-restricted TCR were present, as were most of the "shared" sequences. The MD-Jα26 and LD-Jα26 sequences, which account for many of the CD4-lineage and transitional cells in wild-type mice, were not found in the data set from the class II-deficient mice. However, the H-Jα26 and Y-Jα26 sequences normally confined to the CD4 lineage did persist in knockout thymocytes. Interestingly, these sequences were now found in mature CD8$^+$ cells, even though they were always absent from this compartment in normal mice. This suggests that differentiation of cells bearing these TCR along the CD8-lineage pathway is inhibited in the presence of MHC class II molecules, either as a result of negative selection or because competitive forces may route these cells along the CD4-lineage pathway.

Note that the LMD-Jα26/STNT-Jα2 group of sequences was again essentially absent from these CD4$^+$CD8^{int} cells.

	Transitional $CD4^+CD8^{int}$		Mature $CD4^-CD8^+$	
	$MHCII^+$	$MHCII^0$	$MHCII^+$	$MHCII^0$
H - Jα26	0,5	1,1	—	1,5
Y - Jα26	2,6	3,2	—	0,8
R - Jα2	0,5	—	—	—
L D - Jα26	4,6	—	—	—
M D - Jα26	14,3	—	—	—
S K G - Jα26	2,6	2,1	1,7	1,5
L N T - Jα2	4,6	1,1	8,7	1,5
S R G - Jα26	4,1	4,3	4,7	6,9
S D - Jα26	6,1	—	—	1,5
S R D - Jα26	2,0	1,1	—	1,5
T D - Jα26	1,5	10,6	—	0,8
S T N T - Jα2	—	—	4,1	3,1
L M D - Jα26	—	—	7,6	2,3
I D T - Jα2	—	—	2,9	1,5
R L D - Jα26	—	—	2,9	—
S L N T - Jα2	0,5	1,1	5,8	3,1
S T G - Jα26	1,5	1,1	5,2	1,5
M N T - Jα2	2,0	1,1	2,3	0,8
S M G - Jα26	1,5	2,1	0,6	—
R - Jα26	2,0	9,6	1,2	3,1
S G N T - Jα2	0,5	1,1	2,3	3,1
D - Jα26	1,0	1,1	—	—
S E - Jα26	1,0	1,1	2,3	—
S R D T - Jα2	1,5	5,3	3,5	4,6
S G D T - Jα2	0,5	1,1	—	1,5
I - Jα26	2,6	1,1	2,9	2,3
C - Jα26	0,5	—	1,2	2,3
S M N T - Jα2	1,0	5,3	5,8	2,3
R E - Jα26	0,5	5,3	2,3	0,8
G - Jα26	5,6	6,4	4,1	6,9
S - Jα26	3,1	2,1	3,5	10,0
n	196	94	172	130

Fig. 3. Repeat sequences in MHC class II-deficient Limited mice. The MHC class II deficiency was bred into Limited mice, and the sequence of rearranged CDR3α from the TCRα minilocus determined on single cells sorted from $CD4^+CD8^{int}$ and mature $CD4^-CD8^+$ cells. The frequencies of the repeated set of sequences is shown (for complete data, see Supplemental Data), arranged as in Figure 1. The frequencies in corresponding populations from class II-proficient mice (data from Fig. 1 and [20]) are shown for reference.

2.3 Disjunction between CD4/CD8-lineage commitment and MHC restriction in early transitional cells

As discussed above, the $CD4^+CD8^{int}$ population contains precursors committed to both the CD4 and CD8 lineages, evidenced by protease stripping and knock-in reporter assays. We found both MHC class I- and class II-restricted TCR sequences in these cells (Figs. 2, 3). The key question becomes, then, whether the two overlap: are all class II-restricted TCR sequences found strictly in CD4 lineage-committed cells, and, likewise, are class I-restricted TCR sequences limited to cells committed to the CD8 lineage? In other words, does initial commitment match MHC restriction in early selection intermediates as it does in mature cells? To address this question, we crossed into the Limited mouse genome a CD4-βgal knock-in mutation, which provides a readout of transcriptional activity at the CD4 locus, responding faster than surface levels of CD4 to differentiative changes. When present in the heterozygous state ($CD4^{+/L}$ mice), this mutation does not interfere with CD4-dependent events [18]. Cells with active transcription of the CD4 locus can be detected by flow cytometry after loading with the β-galactosidase (βgal) substrate FDG. In particular, we have shown that the $CD4^+CD8^{int}$ population can be split into CD4 lineage-committed and CD8 lineage-committed cells on the basis of β-gal activity [18].

Single cells were sorted from the $CD4^+CD8^{int}$ population of $Va2Var^+Vβ5^+Ca^{o/o}CD4\text{-}βGal^{+/L}$ mice (three independent animals), either as $β\text{-}gal^{hi}$ or $βgal^{lo}$. The former was sorted at the highest level of TCR expression, to avoid selection intermediates having yet to shut-off the CD4 locus (Fig. 4 A). TCRα sequences were determined from 234 $βgal^+$ and 245 $βgal^-$ cells. Representative data for one mouse are shown in Fig. 4 B, and composite data from the three mice, for the thirty repeat sequences, in Fig. 4 C. Clearly there was a strongly biased representation, in that MHC class I-restricted sequences were more frequent in the CD8 lineage-committed βgal-negative group than in CD4 lineage-committed βgal-positive cells; conversely, MHC class II-restricted sequences were more frequent in the βgal-positive group. On the other hand, this bias was not absolute. Eighteen βgal-negative cells had the prototype class II-restricted LD-Jα26 or MD-Jα26 sequences, while class I-restricted TCR such as SGM-Jα26 and I-Jα26 were found in CD4-lineage-committed cells. The implication is that there is a preferential correlation of early commitment with MHC restriction, but that a significant proportion of cells appear committed to a lineage that does not match their receptor specificity.

It is also noteworthy that the LMD-Jα26 and STNT-Jα2 sequences, virtually absent from $CD4^+CD8^{int}$ cells (Fig. 2, 3) were again missing in the βgal-positive and -negative $CD4^+CD8^{int}$ data set.

2.4 Are TCR^{hi} DP cells the precursors of all lineages?

In most models of positive selection, it is held that the increase in surface expression of the TCR, together with CD69, is one of the earliest signs that a cell is engaged in positive selection, prior to any modifications of coreceptor levels [4, 21, 28, 29]. Thus, the TCR^{hi} DP population is thought to represent the earliest intermediate in selection [30]. One would predict, then, that all of the repeated sequences later selected into mature populations would be found in DP TCR^{hi} cells. This hypothesis was tested by sorting and sequencing the rearranged Va2Var transgene in 127 $Vβ5^{hi}CD4^+CD8^+$ cells from three independent

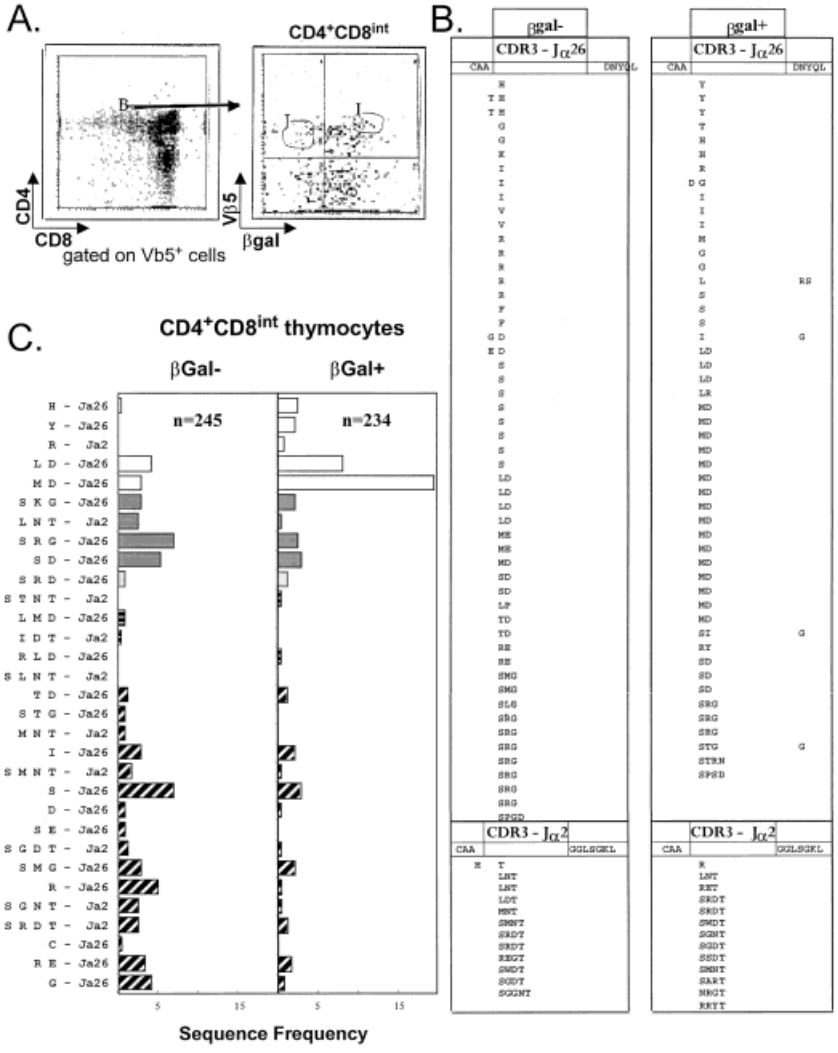

Fig. 4. Early lineage commitment does not strictly follow MHC restriction. The CD4-βgal knock-in mutation was crossed into Limited mice (as heterozygotes, such that the wild-type CD4 locus provides CD4 function). A: Thymocytes were analyzed by multicolor flow cytometry. From the CD4/CD8 plot of the left panel (shown here as gated on Vβ5$^+$ cells for clarity), a gate was set for CD4$^+$CD8^{int} cells and applied to the FDG (fluorescent substrate released by β-Gal) and Vβ5 profile of the right panel; sorting gates for β-gal$^-$ and βgal^{hi} cells are shown. B: Sequences from representative single cells from one such mouse. C: Frequency histogram of sequences from the repeated set among β-gal$^-$ and β-gal$^+$ CD4$^+$CD8^{int} cells (3 mice), coded as in Fig. 2. Sequences of this repeated set represented 64 % and 70 % in gal$^-$ and gal$^+$ sets, respectively.

Limited mice. The CD4 and CD8 sorting gates were strictly positioned at the core of the DP population to avoid any contamination with other populations (Fig. 5 A). The results are presented in Fig. 5 B, aligned with those from the CD4$^+$CD8^{int} cells for comparison. Again, this sequence set had the hallmark of a selected population, 55 % of the sequences belonging to the "repeated group". Strikingly, and contrary to expectation, a number of sequences were almost completely missing from this data set. Several of the class I-restricted sequences were also not observed, such as SRDT-Jα2 or G-Jα26, which is normally a major component of CD8 lineage population, and virtually all class II-restricted sequences. Those that were detected included the LMD-Jα26 and STNT-Jα2 sequences that were consistently rare or absent from CD4$^+$CD8^{int} cells (Figs. 2–4). It appears, then, that the CD4$^+$CD8$^+$CD3^{hi} stage is not a general one through which all maturing cells transit.

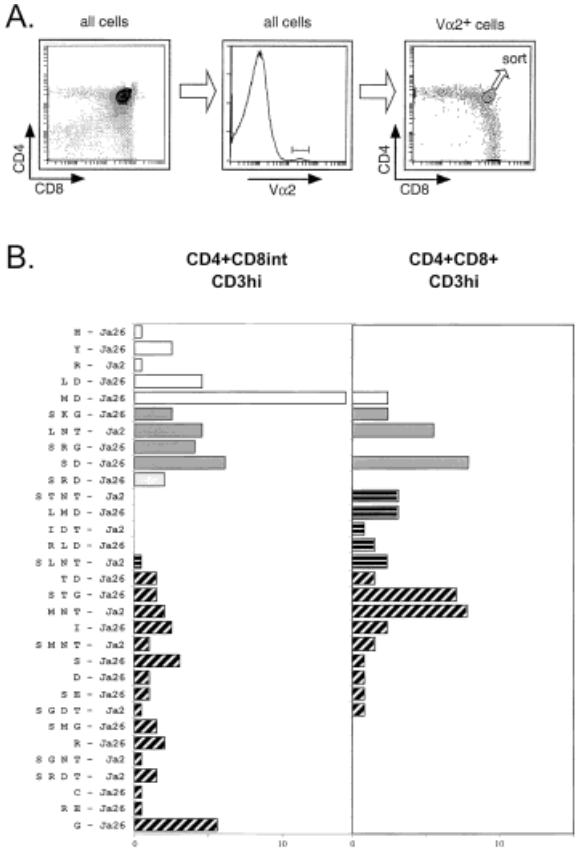

Fig. 5. TCR^{hi} DP represent a distinct pathway. Frequency distribution histogram of sequences from CD3^{hi} DP, arranged as for Fig. 2. The repeated sequences represented 55 % of the total sequences in this population. The corresponding plot for sequences from CD4^{+}CD8^{int} cells is shown for reference on the left.

3 Discussion

We have made use of the unique possibilities offered by the Limited mouse to track the differentiation pathways followed by cohorts of cells carrying subtly dissimilar TCR. As discussed in our previous report [20], single amino-acid differences can lead to reversals in the relative affinity for MHC class I and II molecules. It is likely that the array of variants of the OT-1 CDR3α encompasses receptors with a spectrum of avidities for the two classes of MHC molecule. Our approach extends that of SantAngelo et al., who traced positive selection events by following the relative proportion of a frequent TCRα in a TCRβ transgenic mouse [21]. In conjuction with the proportion of out-of-frame sequences, the evolution of this "parental" sequence led the authors to deduce a map of positive selection pathways. The different experimental setup here and the reliance on single-cell PCR and high-throughput sequencing make for a much greater number of traceable sequences (30 *vs.* 1), and thus a more general view. Some of the conclusions differ, in particular since it now appears that not all class-I restricted TCR follow the same maturation pathway as the "parental" sequence of SantAngelo et al.

3.1 Different TCR sequences are associated with different maturation pathways

The complete sequence data from this and the preceding report are summarized in Fig. 6 A. These data argue for a multi-branch view of positive selection, schematized in Fig. 6 B. The overall conclusion is that the pathways of thymocyte differentiation are not monomorphic, and that the different TCR promote differentiation along different routes. In particular, not all developing thymocytes transit through the TCR^{hi} DP stage, nor through the CD4^{+}CD8^{int} stage. One could argue that the absence or rarity of a given sequence in a particular subpopulation reflects an accelerated transit time, rather than a true bypass; this caveat would appear unlikely to explain the reciprocal distribution we observe in the TCR^{hi}DP and CD4^{+}CD9^{int} populations. This is not only true of class I- *vs.* class II-restricted TCR, but also distinguishes between various class II-restricted receptors (indicated by the different markings in Figs. 2–5). Since these TCR are all expressed from the same transgene, and since they are identical except for small differences in CDR3α, the various pathways they promote must be instructed by the affinity/avidity for self MHC:peptide ligands. The commonly accepted sequence of TCR^{lo} DP > TCR^{hi} DP > CD4^{+}CD8^{int} > SP thus represents an oversimplification:

Some cells shunt the TCR^{hi} DP stage altogether. This is most obviously true for those whose TCR exhibit MHC class II-restricted repeated sequences, but also for a subset of class I-restricted sequences: G-Jα26, SRDT-Jα2 are representative of a group of sequences that are rarely or never found in TCR^{hi} DP cells (Figs. 5, 6 A). This observation is reminiscent of old results from Marodon and Rocha [31]: comparative analysis of thymi from H-Y and 2B4 TCR transgenic mice (class I- and II-restricted receptors, respectively) indicated that only the former showed a sizeable TCR^{hi} DP population. These authors concluded that TCR^{hi} DP are the precursors of mature CD8^{+} cells but not of CD4 SP. The present data confirm the absence of CD4^{+}-lineage precursors from TCR^{hi} DP, but also show that only a subset of CD8^{+} lineage precursors transit through this intermediate stage. Molecularly, this may represent an earlier shut-off of CD4 gene activity, even before the TCR is fully up-regulated.

Conversely, several MHC class I-restricted sequences were not or were only exceptionally found in the data set

A.

	Tot	LN CD4+ WT	CD4+CD8- WT	CD4+CD8int WT	CD4+CD8int MHC II+	CD4+CD8int CD4+/LacZ+	CD4+CD8int CD4+/LacZ-	DP CD3lo	DP CD3hi	CD4-CD8+ WT	CD4-CD8+ MHC II+	CD4-CD8+ CD4+/LacZ-	LN CD8+ WT
		total %	total %	total %	total %	total %	total %	total %	total %	total %	total %	total %	total %
S E G W N T - Ja2	26	26 12.3	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -
S V G W N T - Ja2	20	20 9.4	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -
S I G W N T - Ja2	13	13 6.1	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -	- -
H - Ja26	23	6 2.8	5 2.7	1 0.5	1 1.1	6 2.6	1 0.4	1 0.4	- -	- -	2 1.5	- -	- -
Y - Ja26	24	5 2.4	5 2.7	5 2.6	3 3.2	5 2.1	- -	- -	- -	- -	1 0.8	- -	- -
R - Ja2	21	15 7.1	3 1.6	1 0.5	- -	2 0.9	- -	- -	- -	- -	- -	- -	- -
L D - Ja26	71	3 1.4	27 14.4	9 4.6	- -	19 8.1	10 4.1	3 1.1	- -	- -	- -	- -	- -
M D - Ja26	102	1 0.5	45 24.1	28 14.3	- -	74 31.6	7 2.9	3 1.1	3 2.4	1 0.6	- -	- -	- -
S K G - Ja26	46	5 2.4	11 5.9	5 2.6	2 2.1	5 2.1	7 2.9	- -	3 2.4	3 1.7	2 1.5	2 1.7	1 0.5
L N T - Ja2	53	4 1.9	- -	9 4.6	1 1.1	1 0.4	6 2.4	2 0.7	7 5.5	15 8.7	2 1.5	3 2.6	3 1.4
S R G - Ja26	71	- -	3 1.6	8 4.1	4 4.3	6 2.6	17 6.9	3 1.1	- -	8 4.7	9 6.9	8 6.8	5 2.3
S D - Ja26	54	1 0.5	- -	12 6.1	- -	7 3.0	13 5.3	1 0.4	13 10.2	- -	2 1.5	4 3.4	1 0.5
S R D - Ja26	13	- -	- -	4 2.0	1 1.1	3 1.3	2 0.8	1 0.4	- -	- -	2 1.5	- -	- -
T D - Ja26	27	- -	1 0.5	3 1.5	10 10.6	3 1.3	3 1.2	- -	2 1.6	- -	1 0.8	2 1.7	2 0.9
S T N T - Ja2	22	- -	- -	- -	- -	1 0.4	- -	2 0.7	4 3.1	7 4.1	4 3.1	4 3.4	- -
L M D - Ja26	29	1 0.5	- -	- -	- -	- -	2 0.8	- -	4 3.1	13 7.6	3 2.3	3 2.6	3 1.4
I D T - Ja2	12	- -	- -	- -	- -	- -	1 0.4	- -	1 0.8	5 2.9	2 1.5	1 0.9	2 0.9
R L D - Ja26	12	- -	- -	- -	- -	1 0.4	- -	- -	2 1.6	5 2.9	- -	4 3.4	- -
S L N T - Ja2	25	- -	- -	1 0.5	1 1.1	- -	- -	- -	3 2.4	10 5.8	4 3.1	4 3.4	- -
S T G - Ja26	30	- -	- -	3 1.5	1 1.1	- -	2 0.8	2 0.7	9 7.1	9 5.2	2 1.5	2 1.7	- -
M N T - Ja2	23	- -	- -	4 2.0	1 1.1	- -	2 0.8	- -	10 7.9	4 2.3	1 0.8	1 0.9	- -
S M G - Ja26	21	- -	- -	3 1.5	2 2.1	5 2.1	7 2.9	- -	- -	1 0.6	- -	3 2.6	- -
R - Ja26	38	- -	- -	4 2.0	9 9.6	1 0.4	12 4.9	- -	- -	2 1.2	4 3.1	6 5.1	- -
S G N T - Ja2	20	- -	- -	1 0.5	1 1.1	1 0.4	6 2.4	1 0.4	- -	4 2.3	4 3.1	- -	2 0.9
D - Ja26	12	- -	- -	2 1.0	1 1.1	1 0.4	2 0.8	2 0.7	1 0.8	- -	- -	1 0.9	2 0.9
S E - Ja26	14	- -	- -	2 1.0	1 1.1	- -	2 0.8	- -	1 0.8	4 2.3	- -	1 0.9	3 1.4
S R D T - Ja2	34	- -	- -	3 1.5	5 5.3	3 1.3	6 2.4	- -	- -	6 3.5	6 4.6	1 0.9	4 1.9
S G D T - Ja2	13	- -	- -	1 0.5	1 1.1	1 0.4	3 1.2	- -	1 0.8	- -	2 1.5	- -	4 1.9
I - Ja26	37	- -	- -	5 2.6	1 1.1	5 2.1	7 2.9	- -	3 2.4	5 2.9	3 2.3	3 2.6	5 2.3
C - Ja26	12	- -	- -	1 0.5	- -	- -	1 0.4	- -	- -	2 1.2	3 2.3	- -	5 2.3
S M N T - Ja2	40	- -	- -	2 1.0	5 5.3	1 0.4	4 1.6	- -	2 1.6	10 5.8	3 2.3	2 1.7	11 5.1
R E - Ja26	39	- -	- -	1 0.5	5 5.3	4 1.7	8 3.3	- -	- -	4 2.3	1 0.8	4 3.4	12 5.6
G - Ja26	96	- -	- -	11 5.6	6 6.4	2 0.9	10 4.1	- -	- -	7 4.1	9 6.9	11 9.4	40 18.7
S - Ja26	148	- -	4 2.1	6 3.1	2 2.1	7 3.0	17 6.9	- -	1 0.8	6 3.5	13 10.0	17 14.5	75 35.0
Repeated seqs	1298	100 47.2	104 55.6	135 68.9	64 68.1	164 70.1	158 64.5	21 7.8	70 55.1	131 76.2	85 65.4	87 74.4	180 84.1
All seqs	2196	212	187	196	94	234	245	268	127	172	130	117	214

B.

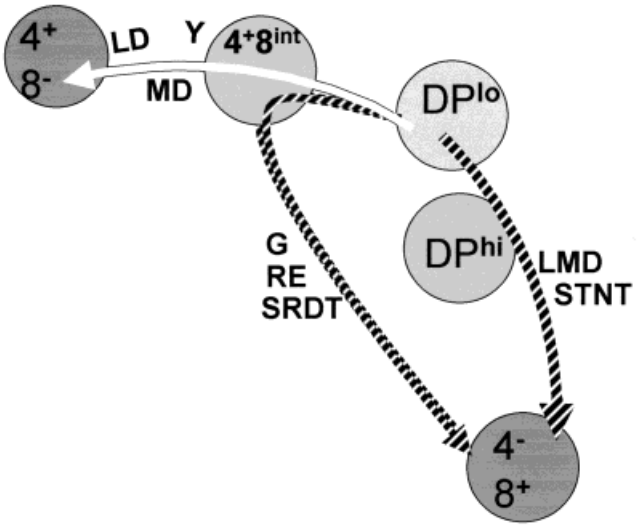

Fig. 6. Alternate pathways of thymocyte differentiation. A: A compilation of the entire sequence data for the 2196 cells from Limited mice and their variants analyzed in this report and in [20]. The number of occurrences and calculated frequency is shown for the repeated sequences (their number and total proportion are shown at the bottom). Coding white: MHC class II-restricted sequences; striped MHC class I-restricted sequences; grey: sequences found in both MHC class I- and MHC class II-restricted compartments. B: Schematic representation of the maturation pathways followed by cells expressing different TCR sequences in Limited mice. MHC class II-restricted cells (white arrows, LD-Jα26, MD-Jα26, Y-Jα26) proceed directly to the $CD4^+$ $CD8^{int}$ stage, without transition as TCR^{hi} DP. Some MHC class I-restricted cells follow the same pathway (G-Jα26, RE-Jα26, SRDT-Jα2), while others seem to proceed only via TCR^{hi} DP intermediates (LMD-Jα26, STNT-Jα2). Cells which can be found in both types of intermediates (STG-Jα26, MNT-Jα2, dotted lines) may follow a sequential transition (TCR^{hi} DP to $CD4^+CD8^{int}$) or alternatively follow either route in a stochastic fashion.

from $CD4^+CD8^{int}$ cells. Clear examples are STNT-Jα2 and LMD-Jα26, otherwise well represented in mature populations. Overall, the representation in $CD4^+CD8^{int}$ cells was inversely correlated with that in TCR^{hi} DP (Figs. 5, 6): those sequences that were missing from the $CD4^+CD8^{int}$ data set were well-represented in the TCR^{hi} DP set, and vice-versa. We and others have noted that different TCR transgenic mice vary in the abundance of their $CD4^+CD8^{int}$ thymocyte compartment (reviewed in [6]). The present observations extend this dichotomy to a situation where variations in transgene expression are not longer an issue, indicating that it is the fine structure of TCR, and by inteference the quality of the MHC : peptide contacts they engage in, that governs this capability.

There are a few exceptions to the mutually exclusive presence of repeated sequences in the $CD4^+CD8^{int}$ and TCR^{hi} DP compartments (Fig. 5). This is the case for the "dual-restricted" sequences, easily understandable if interactions with MHC class II molecules preferentially promote transition to $CD4^+CD8^{int}$ cells while engagement by class I molecules elicits a TCR^{hi} DP phenotype. Yet, some purely class I-restricted sequences are also in this group (MNT-Jα2, I-Jα26, STG-Jα26). In these cases, it is not possible to decide whether the classical TCR^{hi} DP $\rightarrow$ $CD4^+CD8^{int}$ precursor/product relationship applies, or whether these TCR have the capacity to independently promote differentiation along either route.

3.2 In early selection intermediates, CD4/CD8 lineage commitment and MHC specificity do not completely match

The other main conclusion to emerge from our studies is that lineage commitment in early transitional intermediates is correlated with MHC class restriction, but not as tightly as it is in mature thymocytes. In CD4 lineage- and CD8 lineage-committed $CD4^+CD8^{int}$ cells, identified on the basis of CD4-βgal knock-in expression, the distribution of sequences was skewed, such that there was an over-representation of class I-restricted sequences in the βgal^- cells and conversely, more class II-restricted sequences in the β-gal^{hi} cells (Figs. 4, 6). But this dichotomy was far from absolute, a number of cells showing mismatched commitment, in either direction (19% and 11% mismatches among CD8 lineage- and CD4 lineage-committed cells, respectively). This is not equally true for all sequences: I-Jα26, SMG-Jα26 and LD-Jα26 were found in many more mismatched cells than R-Jα26 or MD-Jα26, for example. That mismatched $CD4^+CD8^{int}$ transitional cells exist could already be inferred from protease stripping and reporter expression experiments in MHC class II-deficient mice [2, 18]. The present results show that TCR vary in their propensity to induce mismatched intermediates. Finally, since such mismatches were absent from mature thymocyte pools, the implication is that a secondary selective step does not allow these mismatched intermediates to survive beyond the $CD4^+CD8^{int}$ stage. Brugnera et al. have proposed the interesting concept that early commitment may be reversible, though on the basis of controversial data [32]. This notion could provide an alternative out for mismatched intermediates which, rather than dying, might switch differentiative programs.

These conclusions are consistent with views of lineage commitment in which the choice is not instructed strictly by the class of MHC molecule recognized, but by the thermodynamic constants of the TCR/MHC:peptide interactions. The various TCR structures generated in the Limited mouse, with their fine variations in CDR3α, are each likely to engage different sets of MHC:self-peptide combinations. Depending on the amount of the ligands available on the surface of stromal cells and on the stability of the engagements, these engagements can translate into different intracellular instructions, akin to what has been accomplished by experimental manipulation of lck activation or the time of agonist engagement [14, 16]. This information then channels the cell into the TCR^{hi} DP or $CD4^+CD8^{int}$ pathway, and into CD4- or CD8-lineage commitment. How early the cells transiting through the $CD4^+CD8^{int}$ pathway become committed remains an open question. The TCR^{hi} DP option is not available for MHC class II-restricted cells, perhaps because the higher level of lck brought into play by CD4 co-engagement has an overriding influence, but class I-restricted cells have more leeway.

In conclusion, tracing the fate of TCR sequences in Limited mice has revealed more complexity than expected in lineage relationships during thymocyte differentiation. This complexity will need to be accommodated by any molecular models that attempt to relate lineage commitment and signal processing.

4 Materials and methods

4.1 Mice

The Va2Var minilocus transgene has been described [20. It was initially injected into (B6xSJL)F2 eggs, and subsequently backcrossed to B6. To generate "Limited" mice, the transgenic line was bred with a TCRβ chain transgene encoding the OT-1 β chain [25] and the TCRα KO mutation [33]. The MHC class II-deficiency was bred in from Aβ° mice backcrossed to the C57BL/6 background [27]. The βgal knock-in into the CD4 locus has been described [18]. All mice used in this work were homozygous for the $H\text{-}2^b$ haplotype. Animals were maintained in the conventional facility of the IGBMC under Ministère de l'Agriculture (Agrément 67227) and European Economic Community guidelines.

4.2 Cell sorting and TCR sequencing

All procedures for cell labeling, sorting, single-cell PCR and sequencing were performed as described in [20]. Detection of β-gal activity after hypotonic loading with the fluorogenic substrate FDG was performed as described [18].

Acknowledgements: We thank S. Chan for discussion, P. Brottier and J. Weissenbach (Génoscope-Centre National de Séquencage) for sequencing, and P. Gerber, P. Marchal, J. Hergueux and C. Ebel for assistance. This work was supported by institute funds from the INSERM, the CNRS, the

Hopital Universitaire de Strasbourg, and by a grant from the Association pour la Recherche contre le Cancer. MCN received fellowships from the Programa Gulbenkian de Doutoramento em Biologia e Medicina and the Fundacao para a ciencia e a tecnologia.

References

1 **Lundberg, K., Heath, W., Kontgen, F., Carbona, F. R. and Shortman, K.,** Intermediate steps in positive selection: differentiation of $CD4^+8^{int}$ thymocytes into $CD4^-8^+$ TCR^{hi} thymocytes. *J. Exp. Med.* 1995. **181:** 1643–1651.

2 **Suzuki, H., Punt, J. A., Granger, L. G. and Singer, A.,** Asymmetric signaling requirements for thymocyte commitment to the $CD4^+$ versus $CD8^+$ T cell lineages: a new perspective on thymic commitment and selection. *Immunity* 1995. **2:** 413–425.

3 **Ellmeier, W., Sawada, S. and Littman, D. R.,** The regulation of CD4 and CD8 coreceptor gene expression during T cell development. *Annu. Rev. Immunol.* 1999. **17:** 523–554.

4 **Swat, W., Dessing, M., von Boehmer, H. and Kisielow, P.,** CD69 expression during selection and maturation of CD4+8+ thymocytes. *Eur. J. Immunol.* 1993. **23:** 739–746.

5 **Jameson, S. C., Hogquist, K. A. and Bevan, M. J.,** Positive selection of thymocytes. *Annu. Rev. Immunol.* 1995. **13:** 93–126.

6 **Chan, S., Correia-Neves, M., Benoist, C. and Mathis, D.,** CD4/CD8 lineage commitment: matching fate with competence. *Immunol. Rev.* 1998. **165:** 195–207.

7 **Singer, A., Bosselut, R. and Bhandoola, A.,** Signals involved in CD4/CD8 lineage commitment: current concepts and potential mechanisms. *Semin. Immunol.* 1999. **11:** 273–281.

8 **Davis, C. B., Killeen, N., Casey Crooks, M. E., Raulet, D. and Littman, D. R.,** Evidence for a stochastic mechanism in the differentiation of mature subsets of T lymphocytes. *Cell* 1993. **73:** 237–247.

9 **Chan, S. H., Waltzinger, C., Baron, A., Benoist, C. and Mathis, D.,** Role of coreceptors in positive selection and lineage commitment. *EMBO J* 1994. **4482–4489.**

10 **Robey, E., Itano, A., Fanslow, W. C. and Fowlkes, B. J.,** Constitutive CD8 expression allows inefficient maturation of CD4+ helper T cells in class II major histocompatibility complex mutant mice. *J. Exp. Med.* 1994. **179:** 1997–2004.

11 **Goldrath, A. W., Hogquist, K. A. and Bevan, M. J.,** CD8 lineage commitment in the absence of CD8. *Immunity* 1997. **6:** 633–642.

12 **Matechak, E. O., Killeen, N., Hedrick, S. M. and Fowlkes, B. J.,** MHC class II-specific T cells can develop in the CD8 lineage when CD4 is absent. *Immunity* 1996. **4:** 337–347.

13 **Rooke, R., Waltzinger, C., Benoist, C. and Mathis, D.,** Positive selection of thymocytes induced by gene transfer: MHC class II-mediated selection of CD8-lineage cells. *Int. Immunol.* 1999. **11:** 1595–1600.

14 **Hernandez-Hoyos, G., Sohn, S. J., Rothenberg, E. V. and Alberola-Ila, J.,** Lck activity controls CD4/CD8 T cell lineage commitment. *Immunity* 2000. **12:** 313–322.

15 **Albert Basson, M., Bommhardt, U., Cole, M. S., Tso, J. Y. and Zamoyska, R.,** CD3 ligation on immature thymocytes generates antagonist-like signals appropriate for CD8 lineage commitment, independently of T cell receptor specificity. *J. Exp. Med.* 1998. **187:** 1249–1260.

16 **Yasutomo, K., Doyle, C., Miele, L. and Germain, R. N.,** The duration of antigen receptor signalling determines CD4+ versus CD8+ T-cell lineage fate. *Nature* 2000. **404:** 506–510.

17 **von Boehmer, H.,** Developmental biology of T cells in T-cell receptor transgenic mice. *Annu. Rev. Immunol.* 1990. **8:** 531–556.

18 **Chan, S., Correia-Neves, M., Dierich, A., Benoist, C. and Mathis, D.,** Visualization of CD4/CD8 T cell commitment. *J. Exp. Med.* 1998. **188:** 2321–2333.

19 **Pircher, H., Ohashi, P. S., Boyd, R. L., Hengartner, H. and Brduscha, K.,** Evidence for a selective and multi-step model of T cell differentiation: $CD4^+CD8^{low}$ thymocytes selected by a transgenic T cell receptor on major histocompatibility complex class II molecules. *Eur. J. Immunol.* 1994. **24:** 1982–1987.

20 **Correia-Neves, M., Waltzinger, C., Mathis, D. and Benoist, C.,** The shaping of the T cell repertoire. *Immunity* 2001. **14:** 21–32.

21 **Sant'Angelo, D. B., Lucas, B., Waterbury, P. G., Cohen, B., Brabb, T., Goverman, J., Germain, R. N. and Janeway, Jr. C. A.,** A molecular map of T cell development. *Immunity* 1998. **9:** 179–186.

22 **Kelly, J. M., Sterry, S. J., Cose, S., Turner, S. J., Fecondo, J., Rodda, S., Fink, P. J. and Carbone, F. R.,** Identification of conserved T cell receptor CDR3 residues contacting known exposed peptide side chains from a major histocompatibility complex class I-bound determinant. *Eur. J. Immunol.* 1993. **23:** 3318–3326.

23 **Hogquist, K. A., Jameson, S. C., Heath, W. R., Howard, J. L., Bevan, M. J. and Carbone, F. R.,** T cell receptor antagonist peptides induce positive selection. *Cell* 1994. **76:** 17–27.

24 **Kouskoff, V, Signorelli, K., Benoist, C. and Mathis, D.,** Cassette vectors directing expression of T cell receptor genes in transgenic mice. *J. Immunol. Methods* 1995. **180:** 273–280.

25 **Correia-Neves, M., Waltzinger, C., Wurtz, J. M., Benoist, C. and Mathis, D.,** Amino acids specifying MHC class preference in TCR V alpha 2 regions. *J. Immunol.* 1999. **163:** 5471–5477.

26 **Lundberg, K., Heath, W., Kontgen, F., Carbone, F. R. and Shortman, K.,** Intermediate steps in positive selection: differentiation of $CD4^+8^{int}$/thymocytes into $CD4^-8^+$ TCRhi thymocytes. *J. Exp. Med.* 1995. **181:** 1643–1651.

27 **Cosgrove, D., Gray, D., Dierich, A., Kaufman, J., Lemeur, M., Benoist, C. and Mathis, D.,** Mice lacking MHC class II molecules. *Cell* 1991. **66:** 1051–1066.

28 **Chan, S. H., Benoist, C. and Mathis, D.,** A challenge to the instructive model of positive selection. *Immunol. Rev.* 1993. **135:** 119–131.

29 **Lucas, B. and Germain, R. N.,** Unexpectedly complex regulation of CD4/CD8 coreceptor expression supports a revised model for CD4+CD8+ thymocyte differentiation. *Immunity* 1996. **5:** 461–477.

30 **Shortman, K., Vremec, D. and Egerton, M.,** The kinetics of T cell antigen receptor expression by subgroups of CD4+8+ thymocytes: delineation of CD4+8+3(2+) thymocytes as post-selection intermediates leading to mature T cells. *J. Exp. Med.* 1991. **173:** 323–332.

31 **Marodon, G. and Rocha, B.,** Generation of mature T cell populations in the thymus: CD4 or CD8 down-regulation occurs at different stages of thymocyte differentiation. *Eur. J. Immunol.* 1994. **24:** 196–204.

32 **Brugnera, E., Bhandoola, A., Cibotti, R., Yu, Q., Guinter, T. I., Yamashita, Y., Sharrow, S. O. and Singer, A.,** Coreceptor reversal in the thymus: signaled CD4+8+ thymocytes initially terminate CD8 transcription even when differentiating into CD8+ T cells. *Immunity* 2000. **13:** 59–71.

33 **Philpott, K. L., Viney, J. L., Kay, G., Rastan, S., Gardiner, E. M., Chae, S., Hayday, A. C. and Owen, M. J.,** Lymphoid development in mice congenitally lacking T cell receptor $\alpha\beta$-expressing cells. *Science* 1992. **256:** 1448–1452.

Correspondence: Diane Mathis and Christophe Benoist, Joslin Diabetes Center, One Joslin Place, Boston, MA 02215, USA
Fax: 617 264 2744
e-mail: cbdm@joslin.harvard.edu

M. Correia-Neves' present address: Laboratory of Microbiology and Immunology of Infection, Institute for Molecular and Cell Biology, University of Porto, Porto, Portugal