HHS Public Access

Author manuscript

Nat Immunol. Author manuscript; available in PMC 2010 July 23.

Published in final edited form as:

Nat Immunol. 2009 August ; 10(8): 831–839. doi:10.1038/ni.1769.

Themis is a member of a new metazoan gene family and is required for completion of thymocyte positive selection

Andy L. Johnson[1,2], **L. Aravind**[3], **Natalia Shulzhenko**[2], **Andriy Morgun**[2], **See-Young Choi**[2], **Tanya L. Crockford**[1], **Teresa Lambe**[1], **Heather Domaschenz**[4], **Edyta Kucharska**[4], **Lixin Zheng**[5], **Carola C. Vinuesa**[4], **Michael J. Lenardo**[5], **Christopher C. Goodnow**[4], **Richard J. Cornall**[1,6], and **Ronald H. Schwartz**[2,6]

[1]Nuffield Department of Clinical Medicine, Oxford University, UK

[2]Laboratory of Cellular and Molecular Immunology, National Institutes of Health, Bethesda, MD USA 20892

[3]National Library of Medicine, National Institutes of Health, Bethesda, MD USA 20892

[4]Australian Cancer Research Foundation Genetics Laboratory, John Curtin School of Medical Research, Australian National University, Australia & Australian Phenomics Facility, Australia

[5]Laboratory of Immunology, National Institutes of Health, Bethesda, MD USA 20892

Abstract

T-cell receptor signaling in $CD4^+CD8^+$ double positive thymocytes determines cell survival and lineage commitment, but the genetic and molecular basis of this process is poorly defined. To address this issue, we used ethylnitrosourea mutagenesis to identify a novel T-lineage-specific gene, *Themis*, which is critical for the completion of positive selection. Themis contains a tandem repeat of a unique globular domain (CABIT), which includes a cysteine motif that defines a family of 5 uncharacterized vertebrate proteins with orthologs in most animal species. Themis-deficient thymocytes showed no major impairment in early T-cell receptor signaling, but exhibited altered expression of cell cycle and survival genes before and during positive selection. These data suggest a unique role for Themis in sustaining positive selection.

Correspondence should be addressed to R.J.C. (cornall@ndm.ox.ac.uk) or R.H.S. (rschwartz@niaid.nih.gov). .

[6]These authors contributed equally to this work

AUTHOR CONTRIBUTIONS

A.L.J. designed and conducted experiments, analyzed and interpreted results, and wrote the manuscript; L.A. conducted the bioinformatics analysis and designed related figures and text; N.S. and A.M. assisted with microarray experimental design and analysis; S-Y.C. performed the GST pull-down and some of the biochemical analyses by Western blotting; T.L.C. and T.L. contributed to *ex vivo* cellular analyses, tissue culture, and gene mapping; H.D. and E.K. performed the immunization and ANA screens; L.Z. contributed to cloning, confocal microscopy, and immunoprecipitation studies; C.C.V., M.J.L. and C.C.G. helped with supervision of A.L.J., designing experiments and writing the manuscript; and R.J.C. and R.H.S. directed the study, analyzed and interpreted results, wrote the manuscript, and were D.Phil. co-mentors for A.L.J.

Data Base Accession Number

GRP files containing raw array data are available under GEO15246 on the Gene Expression Omnibus data repository.

Keywords

CABIT domain; cell survival; phylogeny; microarrays

INTRODUCTION

Studies of T cell differentiation and inherited mutations that selectively disrupt this process have provided a rich source of new insights into biochemical pathways that have importance beyond the field of immunology1-6. The stages of T cell differentiation are well defined and serve as a valuable paradigm for understanding the molecular regulation of cell lineages and how individual proteins and biochemical pathways function to coordinate cell growth, survival and differentiation. Immature $CD4^-CD8^-$ double negative (DN) T cells that successfully rearrange and express a T cell receptor-β (TCRβ) chain, proliferate and upregulate CD4 and CD8 co-receptors to become $CD4^+CD8^+$ double positive (DP) cells. Positive selection, negative selection, and lineage commitment decisions are then based in large part on the duration and strength of TCR signaling7, which is heavily influenced by the extracellular milieu and nature of the peptide-MHC complexes presented by thymic stromal cells. Only cells with intermediate activation of TCR signaling pathways are positively selected; cells that are activated with too strong or too weak a signal die by negative selection or neglect, respectively. Lineage commitment to the mature single positive (SP) stages is also thought to depend on signal strength and duration, as stronger, sustained signaling leads to CD4 SP commitment while weaker and shorter signals result in the development of CD8 SP cells8, 9.

Here we describe the discovery of a previously unidentified family of animal proteins containing a novel domain with a conserved cysteine residue (CABIT domain). An ethylnitrosourea (ENU)-induced nonsense mutation in the T cell-specific founding member, Themis, selectively crippled the differentiation of DP T cells into SP cells. Themis did not play a major role in early TCR signaling. Instead, Themis-deficient DP cells failed to maintain normal expression of cell cycle and survival genes and to appropriately regulate metabolic pathways. As a consequence of this deficiency there was a failure to complete positive selection in these mice.

RESULTS

ENU mutation impairing thymocyte differentiation

By immunological screening of a library of pedigrees segregating thousands of ENU-induced nucleotide substitutions10, we identified one, 5AT161, with an abnormal response to immunization, anti-nuclear autoantibodies, low anti-CD3ε+CD28-induced proliferation and a decreased percentage of naive $CD44^{low}$ $CD4^+$ T cells in the blood (Supplementary Fig. 1a and data not shown). The latter two phenotypes proved to be inherited as a fully penetrant, recessive trait that was readily scored and fixed in subsequent generations. In homozygous 5AT161 mutants, flow cytometry of blood and lymph nodes (LN) showed a reversal of the normal CD4:CD8 ratio and an 8-fold decrease in the absolute number of naïve $CD4^+$ cells and a 3-fold decrease in the naïve $CD8^+$ cells (Fig. 1a). By contrast, the

number of antigen-experienced CD44^{hi} cells in both subsets was normal (Fig. 1a). Peripheral naïve CD44^{lo} cells sorted from 5AT161 mice proliferated and upregulated CD25 equivalently to control counterparts following stimulation with anti-CD3ε and anti-CD28 (Supplementary Fig. 1b), confirming that the decreased proliferation in the original screen was due to decreased cellularity. Absolute numbers of other non-T cell hematopoietic lineages (B cells, natural killer (NK) cells, dendritic cells (DCs), and granulocytes) were similar in the central and peripheral lymphoid organs of 5AT161 and control mice, indicating a selective defect in T cell development (Supplementary Fig. 2 and data not shown).

While DN and DP thymocyte numbers, including the positively selected CD69^{+} population, were normal in the thymus of 5AT161 mice, mature CD4SP and CD8SP populations were significantly reduced (Fig. 1b). The cellular loss increased in progressive developmental stages and was most severe in the CD4 lineage. Little to no decline was observed in the earliest uncommitted CD4^{+} subset (TCRβ^{hi}CD24^{hi}MHC-I^{lo}), while a significant loss was seen in the immature CD4^{+} subset (TCRβ^{hi}CD24^{hi}MHC-I^{hi}) (Fig. 1b). CD4^{+}Foxp3^{+} numbers were decreased to a slightly greater extent than the conventional CD4 SP cells, whereas other thymocyte subsets such as NKT and γδ T cells were unaffected (Supplementary Fig. 3).

T cell development requires a complex interaction between maturing thymocytes and surrounding epithelial cells through both cell contact and soluble factors. To address the cellular cause of the developmental defect in 5AT161 thymocytes, mixed bone marrow (BM) chimeras were generated by injecting lethally irradiated wild-type recipients with an equal mix of 5AT161 (CD45.2) and wild-type (WT, CD45.1) BM cells. The final CD45.2:CD45.1 ratio in the recipients was approximately 2.5:1 in all non-T cell hematopoietic compartments as well as in DN and DP thymocyte populations (Fig. 1c). In contrast, poor development of mutant SP thymocytes relative to wild-type cells was evident based on the reversal of this initial CD45.2:CD45.1 ratio in the mature CD4 and CD8 SP populations. The presence of wild-type epithelial cells and neighboring wild-type thymocytes was unable to rescue the defect in 5AT161-derived thymocytes, indicating that the developmental defect is T cell intrinsic.

Reduced survival of 5AT161 thymocytes

To explore positive and negative selection in more detail, we crossed the 5AT161 mutation onto a variety of TCR transgenic backgrounds. Most *Rag2*$^{-/-}$5C.C7-TCR transgenic thymocytes are directed to the CD4 lineage. 5AT161 *Rag2*$^{-/-}$5C.C7-TCR transgenic mice on this background had a complete block in CD4 SP production, and the minor CD8 SP subset was also significantly affected (Fig. 2a). In contrast, the DP subpopulation was significantly increased, suggesting a complete developmental block occurring at that stage. Similar results were observed in crosses onto the MHC Class II-restricted 3A9 and AND TCR-transgenic backgrounds (data not shown).

The H-Y TCR transgene generates both DN and CD8^{+} TCR^{hi} cells in female *Rag2*$^{-/-}$B6 mice. 5AT161 *Rag2*$^{-/-}$H-Y-TCR transgenic mice had severely decreased CD8 SP thymocyte numbers (Fig. 2a). Interestingly, the TCR^{hi} DN subset was present in normal numbers. This

is consistent with previous studies suggesting that mature DN H-Y T cells are latent TCR γδ cells11, which are unaffected in 5AT161 mice (Supplementary Fig. 3). Overall, these TCR-transgenic data indicate a major role for the affected gene in 5AT161 mice in the ability of αβ T cells to undergo positive selection.

Normal negative selection in 5AT161 mice

The developmental failure of SP thymocytes could be caused by increased activation of negative selection pathways in cells that would otherwise undergo positive selection12. Male H-Y TCR-transgenic mice rapidly delete developing thymocytes due to expression of a male-specific peptide recognized by the transgenic TCR13. Deletion of H-Y TCR transgenic thymocytes was similar in 5AT161 and wild-type mice (Supplementary Fig. 4a). Negative selection was also examined in a mouse mammary tumor virus (MMTV) superantigen driven model. Superantigens specific for the $V_\beta5$ TCR and the I-E^k molecule are produced by an endogenous MMTV provirus14. Wild-type and 5AT161 mice had similar fold decreases in the abundance of $V_\beta5^+$ thymocytes and splenocytes when comparing the B6 (non-deleting) and the B10.BR (deleting) backgrounds in both CD4 and CD8 lineages (Fig. 2b). Finally, proteins important for negative selection, including Bim, Nur77, ERK5 and Ras, were expressed normally upon TCR cross-linking of 5AT161 thymocytes (Supplementary Fig. 4b, 4c and data not shown). These data suggest that enhanced negative selection does not cause the decrease in mature T cells observed in 5AT161 mice.

CABIT-2 domain *of* Themis contains a stop codon

To identify the mutation responsible for the T cell defect in 5AT161 mice, we screened F2 mice from (5AT161 x 129) and (5AT161 x CBA) intercrosses for a peripheral blood CD4:CD8 ratio of less than 1. As expected for a recessive trait, approximately 25% (34/130) of the F2 mice were affected (Supplementary Fig. 5). Using strain-specific genetic markers, the 5AT161 locus was mapped to a 2.4 Mb region between 27.5 and 29.9 Mb on chromosome 10, which contains10 genes (Fig. 3a). ENU randomly induces a point mutation every 1-2 Mb suggesting that very few mutations would be expected in an interval of this size15. A previously uncharacterized gene in the interval, E430004N04Rik, was a likely candidate based on its T cell-specific expression profile (http://symatlas.gnf.org). We identified a T→A transversion in affected 5AT161 mice that is predicted to generate a stop codon in place of the tyrosine residue at position 489 in the predicted open reading frame (ORF) of E430004N04Rik (Fig. 3b). In a consensus reached among the three laboratories that have recently uncovered the function of this gene, we decided to call it *Themis* for Thymus Expressed Molecule Involved in Selection16,17.

To better understand the function and evolutionary origin of the *Themis* gene, we analyzed the putative protein sequence of the most commonly identified transcript in EST databases using a program that employs empirically determined entropy thresholds to identify potential globular domains in proteins. The predicted protein, Themis, contains a tandem repeat of two homologous globular domains and a C-terminal non-globular tail enriched in proline residues (Fig. 4). The *Themis(Y489X)* mutation is near the end of the second of these provisional globular domains. Further analysis using the PSI-BLAST program against a

custom database of genome sequences revealed several other metazoan ORFs, containing the same globular domain. Since profile comparisons revealed no significant relationship to any previously characterized globular domain, we have named this structure the Cysteine-containing All Beta in Themis (CABIT) domain. The new gene family contains three mammalian homologues (*Themis*, *ICB-1*, and 9130404H23Rik), which each possess two copies of the CABIT domain and a highly conserved proline-rich region (Fig. 4). In contrast, Fam59A, Fam59B and related proteins from mammals to cnidarians, including the insect Serrano proteins, possess a single copy of the CABIT domain, a proline-rich region, and often a C-terminal SAM (sterile alpha motif) domain (Fig. 4).

Multiple sequence alignment predicts that the CABIT domain adopts an all β-strand structure with at least 12 strands, suggesting either an extended β-sandwich-like fold or a dyad of 6-stranded β-barrel units (Supplementary Fig. 6). Importantly, the alignment shows that the CABIT domains contain a nearly absolutely conserved cysteine residue, which is likely to be exposed since it occurs at the end of a predicted β-strand in the middle of the domain (Fig. 3c). All other conserved residues, including a highly conserved motif surrounding the cysteine residue ($\varphi XCX_{7\text{-}26}\varphi XLP\varphi X_3GXF$ with X = any amino acid and φ = any hydrophobic residue), are either hydrophobic residues forming cores of strands or polar residues marking the turns between strands. Local homology within subsets of the Themis family, such as a putative nuclear localization signal conserved only in the second domain of the three vertebrate Themis proteins, allow for potential variability in the function and/or regulation of these proteins; yet the conservation of domain architecture and of the cysteine residue suggests a conserved biochemical role for the CABIT domain in metazoans.

Themis interacts with Grb2

To characterize *Themis* expression, we conducted an RT PCR analysis on amplified cDNA generated from FACS-sorted lymphocyte populations of B6 mice. *Themis* expression was barely detectable in uncommitted DN1 thymocytes, but was highly expressed in the succeeding DN2 to DN4 populations. Expression peaked in the DP population before decreasing more than 10-fold in CD4 and CD8 SP cells in both the thymus and periphery (Fig. 5a). Published microarray studies show that *Themis* expression is down-regulated even further in $CD4^+$ regulatory T cells18-20. Immunoblotting with Themis-specific rabbit polyclonal antibodies (Supplementary Fig. 7) detected a single protein species near the predicted size of 73 kDa in thymus, LN and spleen samples from wild-type mice that was absent in *Themis(Y489X)* thymocytes and in non-lymphoid tissues (Fig. 5b). Relative to wild-type cells, *Themis(Y489X)* thymocytes had a 3-fold decrease in Themis mRNA, and only a faint band near the size of the predicted truncated protein was observed in mutant thymocyte lysates (data not shown). Thus, the mutant message and protein appear to be unstable.

A Themis-EGFP fusion protein localized to the cytoplasm when expressed in the HEK293 (human endothelium) and AKR1 (mouse DP thymocyte) cell lines (Fig. 5c and data not shown). To identify binding partners in an unbiased way, a GST-Themis fusion protein was used to pull-down other proteins from thymus and LN lysates, and these were then identified by mass spectrometry on bands from SDS-PAGE. From this analysis, Grb2 was discovered

as a binding partner. Its association was then verified in co-immunoprecipitation studies using both HEK293 and Jurkat cell lines overexpressing Themis-EGFP (Fig. 5d).

Normal early TCR signaling in *Themis(Y489X)* thymocytes

Grb2 is an SH3-SH2-SH3 adaptor protein involved in tyrosine kinase signaling cascades downstream of numerous cell surface receptors including the TCR21-23. To determine if Themis functioned downstream of the TCR, we activated wild-type and mutant thymocytes by cross-linking the TCR with antibodies, with or without anti-CD4. We noted no alterations in early TCR signaling in *Themis(Y489X)* cells. Total phosphotyrosine induction, Erk 1/2 phosphorylation, and calcium flux were all similar in mutant and wild-type thymocytes (Supplementary Fig. 8a and c). Degradation of IκBα was slightly enhanced in *Themis(Y489X)* thymocytes (Supplementary Fig. 8b), but no significant differences were observed in IKKα phosphorylation or the translocation of RelA into the nucleus (data not shown). Prolonged stimulation in overnight cell culture also showed similar induction of CD5 and CD69 expression on DP thymocytes (Fig. 6a and data not shown). Positive selection, however, depends on relatively weak interactions between the TCR and the peptide-MHC complex, and, even at low concentrations, antibody cross-linking could be too strong a stimulus to reveal the consequences of this event in immature thymocytes. To address this issue, we used pigeon cytochrome *c* (PCC)-specific 5C.C7 and AND TCR transgenic cells, which could be stimulated to express CD69 using a range of altered peptide ligands of varying affinities for the TCR24. Three PCC peptides, WT 88-104 and its K99P and K99Q variants, were selected for their strong, weak and extremely weak agonistic properties, respectively. *Themis(Y489X)* thymocytes responded normally to all of these ligands, even the barely stimulatory K99Q peptide (Fig. 6b). Thus, we conclude that initial TCR signaling in *Themis(Y489X)* thymocytes is mostly normal.

Altered gene expression in *Themis(Y489X)* thymocytes

Despite the lack of any major TCR-dependent signaling abnormalities and the presence of normal absolute numbers of cells, freshly isolated *Themis(Y489X)* DP thymocytes expressed slightly lower amounts of the developmental markers CD2, CD27, and CD5 (Fig. 6c). Given that Themis is first expressed at the DN1 stage of thymocyte development, these observations suggest that the loss of *Themis(Y489X)* DP thymocytes following positive selection might be due to a functional impairment established at a stage prior to positive selection. To search for alternative pathways affected by Themis deficiency at all stages of development, we compared gene expression profiles in pre-selection DP cells ($CD5^{lo}CD69^{lo}$), post-selection DP cells ($CD5^{hi}CD69^{+}$) and uncommitted $CD24^{+}$MHC-I^{lo} cells ($CD4^{+}CD8^{lo}TCR\beta^{hi}CD24^{+}$MHC-$I^{lo}$) sorted from wild-type and *Themis(Y489X)* mice by flow cytometry. We identified 325 genes that were either differentially expressed in at least one of the three populations in mutant mice or had altered expression changes during development of these thymocyte populations when compared to wild-type expression patterns (false discovery rate of 10%; Supplementary Table 1). Importantly, the transcription pattern did not demonstrate any changes in downstream gene targets of the two principal pathways known to be involved in positive selection, those involving Erk and NF-AT (Fig. 7a)25. The normal induction of these target genes bolsters the conclusion that initial TCR signaling in DP thymocytes is unaffected by the *Themis* mutation.

K-medians clustering was then used to separate the gene list into 5 clusters based upon expression differences between wild-type and mutant samples as well as developmental regulation within wild-type samples (Fig. 7b). This analysis revealed two broad categories of transcriptional differences: genes differentially expressed before and during positive selection (clusters 1, 2 and 4), and those with altered expression that was specific to the uncommitted $CD24^{+}MHC\text{-}I^{lo}$ population (clusters 3 and 5). Since cell loss was first observed in the immature SP cells, we reasoned that gene expression differences in the immediately preceding $CD24^{+}MHC\text{-}I^{lo}$ population could provide insight into the cell death process. In addition to examining clusters 3 and 5, we also identified all of the genes that were significantly up- or down-regulated in the $CD24^{+}MHC\text{-}I^{lo}$ population independent of their cluster designation (57 and 69 genes respectively). From this analysis, apoptosis-related genes emerged as the most highly enriched functional category in the down-regulated genes in this population ($P<0.001$) together with carbohydrate and vitamin metabolic processes that are essential for cell survival. Among the affected apoptosis-related genes was B-Raf, an Erk-signaling mediator that is required for development beyond the DP stage26. Affected differentiation and survival factor genes also included genes important in glucose uptake *(Slc2a3/Glut3)* inflammatory cytokine production (LITAF), and direct regulators of apoptosis (*Zmat1 and Bcl2alc*). Finally, *Notch1* expression was decreased despite the normal activation of Notch target genes in *Themis(Y489X)* thymocytes (data not shown). Thus, despite normal numbers of cells in the $CD24^{+}MHC\text{-}I^{lo}$ subset in mutant mice, it appears that these cells are pre-disposed to cell death and show signs of metabolic stress before completing positive selection.

In addition to the expression changes in the $CD24^{+}MHC\text{-}I^{lo}$ population, the most striking signature identified by functional annotation was for genes with decreased expression in the stages before and during selection, most notably in cluster 2. Genes in this cluster, including *cyclin B1, cdc2*, the transcription factor E2F1, and many E2F1 target genes, are almost uniformly associated with elements of cell survival and cell cycle progression ($P<10^{-20}$). Importantly, analysis of DNA content in these pre-selection cells did not indicate an accumulation of mutant thymocytes in any particular stage of the cell cycle (Supplementary Fig. 8). Therefore, the expression differences seem instead to set the stage for decreased survival in the subsequent developmental populations of the *Themis(Y489X)* mice. Some of these genes, including *Survivin* (also called *Birc5*) and *Lymphocyte-specific helicase* (also called *HELLS*), are important for the survival of DP thymocytes27-29. Interestingly, *Themis* is included in cluster 2, suggesting that it may also have a direct function in cell viability relating to cell cycle progression and/or differentiation.

Mutant thymocytes also exhibited aberrant expression of lipid biosynthesis genes (cluster 4, $P<0.01$), and consistently decreased expression of genes such as LDLR, StarD4, and Fads1 (cluster 1, $P<0.01$) that are involved in sterol transport and are transcriptionally regulated by sterol response element (SRE) transcription factors. Recently, the balance of cholesterol transport and cholesterol synthesis pathways by liver X receptor (LXR) and SREBF pathways was shown to regulate the proliferation of T cells after stimulation30, suggesting that changes to lipid metabolism and cell cycle pathways might play a coordinated role. Functional relationships between cell cycle and lipid metabolism pathways within our data

set of differentially expressed genes revealed extensive alterations to these pathways in *Themis(Y489X)* thymocytes (Supplementary Fig. 10). These data suggest that differences in the interpretation of TCR signals in the absence of Themis might not be at the level of conventional TCR-dependent signaling pathways, but rather a consequence of the combined loss of cell cycle proteins and pro-survival and metabolic factors.

DISCUSSION

By using a systematic screening strategy on ENU-induced mutant pedigrees, we have identified the uncharacterized gene *Themis* and showed that its product plays a critical role in the positive selection of αβ T cells. *Themis(Y489X)* thymocytes failed to generate normal numbers of mature $CD4^+$ and $CD8^+$ T cells even though the proximal and several downstream functions of TCR signaling were not greatly affected in mutant thymocytes. Genome-wide expression analysis revealed a clear pattern of decreased expression of genes involved in cell cycle, survival, cholesterol transport and metabolism in *Themis(Y489X)* thymocytes. By the $CD24^+$MHC-I^{lo} stage, these defects culminate in a pro-apoptotic environment that leads to cell death just before the cells commit to either SP lineage. These findings show that Themis supports normal thymocyte development and is required to prevent cell death during positive selection.

DP thymocytes are programmed to enter a physiological state in which TCR-dependent signals are accurately translated to determine death or survival. For example, calcineurin activity is required for early DP thymocytes to acquire highly sensitive Erk activation in response to TCR triggering31. If calcineurin is blocked by cyclosporine pretreatment, DP cells arrest with elevated expression of multiple cell cycle-regulating genes such as those encoding Cyclin B1 and Cdc2 and show impaired positive selection. In contrast, the *themis* mutant thymus contains a normal number of these Erk-sensitive cells (data not shown) and shows the opposite phenotype, namely, down-regulation of cell cycle and survival genes. This is an exaggerated version of the process that normally occurs as thymocytes undergo positive selection32. Mixed BM chimeras showed neither accumulation nor loss of mutant cells at any point before the SP stage and analysis of DNA content did not indicate an accumulation of mutant thymocytes in any particular stage of the cell cycle. Therefore, the enhanced down-regulation of cell cycle-related genes is not due to pre-mature cell cycle arrest or to a delay in the DP stage before positive selection. Instead, the altered expression of all these genes influences the outcome of TCR engagement during positive selection and predisposes the selected cells to apoptosis prior to CD4 versus CD8 lineage commitment.

During positive selection, TCR activation has the capacity to induce both pro-survival signals such as Bcl-2 and pro-apoptotic signals such as Bim and Nur77. The extent to which these pathways are activated depends on both the affinity of the TCR for the selecting peptide-MHC ligand and the internal milieu of the cell. In every measure of both positive and negative selection capacity tested, *Themis(Y489X)* thymocytes responded similarly to wild-type controls, yet mutant thymocytes failed to survive positive selection, and showed a more severe defect in CD4 SP production. The strength of signal model predicts that CD4 lineage commitment requires stronger and more sustained T cell activation relative to the signals that drive CD8 commitment7. The elevated activation requirement for $CD4^+$

development might more strongly impact the pro-apoptotic milieu present in mutant DP thymocytes resulting in the enhanced death seen in this lineage. TCR transgenic strains carrying the *Themis(Y489X)* mutation had an earlier and nearly complete block in $CD69^+$ DP thymocyte development. DP cell numbers were also slightly increased. These differences could be due to premature and elevated expression of the αβ TCR in the transgenic thymocytes, which may lead to enhanced TCR signaling during, or even prior to, the normal onset of the milieu required for positive selection. Thus, the timing and strength of TCR signals in relation to the expression of cell cycle and survival proteins in DP thymocytes could ultimately determine their developmental fate.

Interestingly, *Ptprk*, a receptor-type tyrosine phosphatase gene contiguous to *Themis* in all sequenced vertebrates, is also required for survival of SP thymocytes33. *Ptprk* is not mutated and is expressed normally in *Themis(Y489X)* mice (data not shown); however, a spontaneous deletion in Long Evans Cinnamon (LEC) rats disrupts the expression of both *Ptprk* and the *Themis* ortholog, causing an almost complete block of CD4 SP production, a more severe phenotype than that of *Themis(Y489X)* mice34. Retroviral expression of wild-type *Ptprk* alone in LEC thymocytes showed a partial rescue of CD4 SP production, leaving a phenotype very similar to that observed in *Themis(Y489X)* mice. This suggests that PTPRK and Themis are both required for thymocyte survival in LEC rats. It is not clear if these two proteins function in a shared pathway, but it is intriguing that the 3 *Themis-like* genes in vertebrate genomes are all closely linked to *Ptprk* paralogs. Ptprk is a tumor suppressor induced downstream of transforming growth factor-β signaling35-38, and its expression in the thymus peaks in the DN population33. Despite likely functioning in the proliferation of early thymocytes, cell loss does not occur until the DP to SP transition in PTPRK-deficient LEC rats. This pattern of early gene expression associated with a delayed loss of cells at the positive selection step is thus similar to the situation found for *Themis*-mutant mice.

The Themis protein contains a C-terminal proline-rich region and a tandem repeat of a previously unidentified globular domain, which features a fully conserved cysteine motif (CABIT domain). The proline-rich region likely interacts with SH3-domain-containing proteins such as Grb2 and could serve to localize Themis to particular intracellular compartments. The novel CABIT domain defines a protein family with 5 mammalian members and ancestral orthologs in nearly all animal lineages. Of the mammalian proteins, ICB-1 and 9130404H23Rik have similar architecture to Themis with a tandem repeat of the CABIT domain at the N-terminus. Limited in vitro studies of ICB-1 suggest that it also may have a role in regulating cell cycle progression and survival, since ICB-1 upregulation is found in several cancer cell lines39,40 and its inhibition by RNAi knockdown has anti-proliferative effects40. The more distantly related Fam59 proteins contain only a single CABIT domain, a less ordered proline-rich region, and a SAM domain, and are most similar to the majority of non-vertebrate ancestral forms. This basic structure is present in all metazoan lineages from cnidarians onwards with the exception of nematodes. Nematodes have lost elements of a number of ancestral animal-specific signaling pathways, such as NF-AT, NF-κB and SHH. Given that gene-loss often follows a correlated pattern along

functional lines, it is possible that the loss of CABIT-domain proteins in nematodes is correlated with the loss of one of these signaling pathways.

Conserved cysteine residues without associated conserved histidine or acidic residues such as those observed in the CABIT domains are found in a relatively small number of catalytic domains: namely, the E2 ubiquitin ligases and sulfur carrier or thiol redox proteins, such as those with OsmC, MOSC, thioredoxin, or ERV1 folds. Therefore, it is conceivable that the Themis CABIT domain could perform a similar biochemical function. Conservation of the cysteine and of the overall architecture of the domain, but increased variability in other regions of the protein, including the loss of the SAM domain and the presence of an NLS in the three vertebrate Themis proteins, suggests that this family of proteins is performing a similar biochemical role in alternative pathways or complexes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

A.L.J. is an NIH/Oxford Scholar. The Division of Intramural Research, National Institute of Allergy and Infectious Diseases and National Library of Medicine, National Institutes of Health, the Medical Research Council (RJC), the Wellcome Trust (RJC, CCG), an ARC Federation Fellowship (CCG), and the NIHR Biomedical Research Centre Programme (RJC) supported this work. The authors wish to thank K. Holmes, E. Stregevsky, and B. Hague for flow sorting; D.E. Anderson for mass spectrometry; T. Myers and Q. Su for assistance with microarrays; and the staffs of the Australian Phenomics Facility, BMS Oxford, and the NIH Comparative Medicine Branch for excellent animal husbandry.

Appendix

ON-LINE METHODS

Mice

Ethylnitrosourea treatment of C57BL/6 mice was as described previously42. TCR transgenic mice were 3A9 H-2^k 43 and $Rag2^{-/-}$5C.C7-H-2^k 44, AND-H-2^b 11, and H-Y H-2^b 45. Genetic mapping involved intercrosses of 5AT161 with 129/SvEv or CBA. Microsatellite markers and SNPs from both crosses were then used to create a fine map of the locus. Details of *Themis(Y489X)* specific primers are available on request. Mixed BM chimeras were prepared as previously described.46

Antibodies, Cell stimulations and Confocal Microscopy

Themis-specific rabbit polyclonal antiserum was generated against amino-acids 239-546 (Primm Biotech) and affinity purified by incubation with the immunogen bound to nitrocellulose membranes. Details of commercial antibodies are available on request. Pigeon cytochrome *c* (PCC) peptide stimulations of AND transgenic thymocytes were carried out as previously described24. Peripheral blood cells were stimulated with soluble purified anti-CD3ε (145-2C11, eBiosciences) and anti-CD28 ascites (37.51) for 48h and labelled with ^{3}H-thymidine. Sorted $CD44^{lo}$ $CD4^+$ and $CD8^+$ lymphocytes were stimulated with plate-bound anti-CD3ε. CD25 staining was examined at 24h, and cytokines at 48h (R&D

Biosystems ELISA kits). Full-length Themis cDNA (636 aa) cloned into pEGFP-N1 was transfected into HEK293 cells for 24 hours before staining with Hoechst 33342 nuclear dye and imaging of live cells (BIO-RAD MRC1024 confocal scanning laser microscope).

Biochemical Analysis

Thymocytes were stimulated with 1-10 μg/mL of biotinylated anti-CD3ε or anti-TCRβ (H57, BD) with biotinylated anti-CD28 (37.51, BD) or biotinylated anti-CD4 (GK1.5, BD). These were pre-bound for 30 seconds before crosslinking with 5μg/mL streptavidin (Southern Biotech). Calcium flux analysis and immunoblotting were performed as previously described47. Full-length Themis with an N-terminal GST fusion or a GST only control were transiently expressed in HEK293 cells overnight and bound to a column containing GST-coated beads. Mixed wild-type B6 thymus and LN lysates were passed slowly over the GST-only column to remove non-specific interactions and then over the GST-Themis column. After washing, the proteins were eluted with free GST, separated by SDS-PAGE, and Coomassie Blue stained Themis-specific bands were analysed by mass spectrometry48.

RNA sample preparation and amplification

For microarrays, *Themis(Y489X)* and wild-type thymocytes were sorted by flow cytometry into three populations: $CD4^+CD8^+CD5^{lo}CD69^{lo}$, $CD4^+CD8^+CD5^{hi}CD69^{hi}$, and $CD4^+CD8^-TCR\beta^{hi}CD24^+MHC\text{-}I^{lo}$. Three separate thymic samples were collected daily over three separate days into ice cold RNALater (Ambion) resulting in 8-9 samples per genotype and subpopulation. RNA was isolated using TRIZOL reagent (Invitrogen). For RT-PCR, DN thymocyte populations were prepared by depleting DP and SP populations using rat anti-CD4 and anti-CD8 Dynal beads (Invitrogen) and then by flow cytometry. Antibodies to CD4, CD8, B220, Gr-1, NK1.1, Mac1, CD11b, CD11c, and TCRβ were used to gate out non-DN cells. DN1 was gated as $CD25^-CD44^{hi}CD117^+$, DN2 as $CD25^+CD44^{mid/hi}$, DN3 as $CD25^+CD44^{lo/-}$, and DN4 as $CD25^-CD44^-CD24$. *Themis* primers were exon-3 forward TGAAATCCAAGGTGTGCTGA and exon-4 reverse 5′CGTCCGTAGACAGCAACTGA. For the arrays, 50 ng of RNA was amplified using the Ovation Aminoallyl-RNA amplification and labeling system (NuGen) and 2 μg of amplified cDNA was these labeled with Cy5 using the CyDye Post-Labeling Reactive Dye packs (GE Healthcare). 1-3 μg of the remaining cDNA from each sample was pooled to create an internal Cy3-labeled reference for each array. 2 μg each of labeled samples and reference were mixed and hybridized as described49 to Mmca arrays manufactured by the NIAID Microarray Facility (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL6806). Images were scanned by GenePix4000B Scanner (Axon Instruments/Molecular Devices) and analyzed using the mAdb program (http://madb.niaid.nih.gov/).

Microarray analysis

To identify genes that were differentially expressed, we used random variance t-test and regression analysis implemented in BRB Array Tools (http://linus.nci.nih.gov/BRB-ArrayTools.html). We also identified genes with differential changes in expression between mutant *Themis(Y489X)* and wild-type populations during developmental progression from

the DP to the $CD24^{+}MHC\text{-}I^{lo}$ stage just prior to lineage commitment. For all analyses, genes with a false discovery rate (Benjamini-Hochberg method) of <10% were accepted. Differentially expressed genes were initially separated based on elevated or reduced expression in mutant samples. A K-median clustering algorithm was then used to further cluster (MeV: MultiExperiment Viewer, TIGR50) the genes based on expression differences (increasing, decreasing, or mostly steady) between populations in wild-type samples. The "Figure of Merit" method was used to estimate the number of clusters. We identified overrepresented Gene Ontologies categories (Biological Process) in each of the five clusters using Explain (Biobase) software accepting those with $P<0.01$ adjusted for multiple comparisons. Biochemical and functional networks were assembled using Ingenuity Pathways Analysis and Pathway Studio programs as well as individual literature searches.

Evolutionary analysis

Sequence profile searches were performed against the NCBI non-redundant (NR) protein database (NIH), and a locally compiled database of proteins from completely or near-completely sequenced genomes. PSI-BLAST searches were performed using an expectation value (E-value) of 0.01 used as the threshold for inclusion into the profile 64; searches were iterated until convergence. Alignment-derived HMM searches were performed using the HMMer package. Multiple alignments were constructed using the Kalign 67 program. Protein secondary structure was predicted using a multiple alignment as the input for the JPRED2 program, which uses information extracted from a PSSM, HMM, and residue frequencies in alignment columns. Pair-wise comparisons of HMMs, using a single sequence or multiple alignments as query, against profiles of proteins in the PDB database were performed with the HHPRED program. Similarity-based clustering was performed using the BLASTCLUST program (ftp://ftp.ncbi.nih.gov/blast/documents/blastclust.html) with empirically determined length and score threshold parameters. Phylogenetic analysis was carried out using neighborhood-joining and minimum evolution-based methods with gamma distributed rates and a JTT substitution matrix as implemented in the MEGA4 program 70. The shape parameter α was estimated empirically through a series of trials. Maximum likelihood trees were also obtained by first generating the least-squares tree (FITCH program of the PHYLIP package 71) with subsequent local rearrangement using the PROTML program (MOLPHY package 72). All large-scale procedures were carried out using the TASS software package (L. Aravind, unpublished).

Statistical Analysis

Statistical analysis was performed using GraphPad Prism Version 4.0. Statistical analysis was performed using unpaired, two-tailed, t-test, with a 95% C.I.

REFERENCES

1. He X, et al. The zinc finger transcription factor Th-POK regulates CD4 versus CD8 T-cell lineage commitment. Nature. 2005; 433:826–833. [PubMed: 15729333]
2. Hughes P, Bouillet P, Strasser A. Role of Bim and other Bcl-2 family members in autoimmune and degenerative diseases. Curr Dir Autoimmun. 2006; 9:74–94. [PubMed: 16394656]
3. Mitchell BS, Kelley WN. Purinogenic immunodeficiency diseases: clinical features and molecular mechanisms. Ann Intern Med. 1980; 92:826–831. [PubMed: 6247948]

4. Siggs OM, et al. Opposing functions of the T cell receptor kinase ZAP-70 in immunity and tolerance differentially titrate in response to nucleotide substitutions. Immunity. 2007; 27:912–926. [PubMed: 18093540]
5. Taniuchi I, Littman DR. Epigenetic gene silencing by Runx proteins. Oncogene. 2004; 23:4341–4345. [PubMed: 15156191]
6. Weerkamp F, van Dongen JJ, Staal FJ. Notch and Wnt signaling in T-lymphocyte development and acute lymphoblastic leukemia. Leukemia. 2006; 20:1197–1205. [PubMed: 16688226]
7. Germain RN. T-cell development and the CD4-CD8 lineage decision. Nat Rev Immunol. 2002; 2:309–322. [PubMed: 12033737]
8. Hernandez-Hoyos G, Sohn SJ, Rothenberg EV, Alberola-Ila J. Lck activity controls CD4/CD8 T cell lineage commitment. Immunity. 2000; 12:313–322. [PubMed: 10755618]
9. Kappes DJ, He X. CD4-CD8 lineage commitment: an inside view. Nat Immunol. 2005; 6:761–766. [PubMed: 16034433]
10. Vinuesa CG, Goodnow CC. Illuminating autoimmune regulators through controlled variation of the mouse genome sequence. Immunity. 2004; 20:669–679. [PubMed: 15189733]
11. Terrence K, Pavlovich CP, Matechak EO, Fowlkes BJ. Premature expression of T cell receptor (TCR)alphabeta suppresses TCRgammadelta gene rearrangement but permits development of gammadelta lineage T cells. J Exp Med. 2000; 192:537–548. [PubMed: 10952723]
12. Jimi E, Strickland I, Voll RE, Long M, Ghosh S. Differential role of the transcription factor NF-kappaB in selection and survival of CD4+ and CD8+ thymocytes. Immunity. 2008; 29:523–537. [PubMed: 18957265]
13. Kisielow P, Bluthmann H, Staerz UD, Steinmetz M, von Boehmer H. Tolerance in T-cell-receptor transgenic mice involves deletion of nonmature CD4+8+ thymocytes. Nature. 1988; 333:742–746. [PubMed: 3260350]
14. Woodland D, Happ MP, Bill J, Palmer E. Requirement for cotolerogenic gene products in the clonal deletion of I-E reactive T cells. Science. 1990; 247:964–967. [PubMed: 1968289]
15. Cook MC, Vinuesa CG, Goodnow CC. ENU-mutagenesis: insight into immune function and pathology. Curr Opin Immunol. 2006; 18:627–633. [PubMed: 16889948]
16. Accompanying paper of Dr. Paul
17. Accompanying paper of Dr, Nick Gascoigne
18. Fontenot JD, Rasmussen JP, Gavin MA, Rudensky AY. A function for interleukin 2 in Foxp3-expressing regulatory T cells. Nat Immunol. 2005; 6:1142–1151. [PubMed: 16227984]
19. Hill JA, et al. Foxp3 transcription-factor-dependent and -independent regulation of the regulatory T cell transcriptional signature. Immunity. 2007; 27:786–800. [PubMed: 18024188]
20. Marson A, et al. Foxp3 occupancy and regulation of key target genes during T-cell stimulation. Nature. 2007; 445:931–935. [PubMed: 17237765]
21. Burack WR, Cheng AM, Shaw AS. Scaffolds, adaptors and linkers of TCR signaling: theory and practice. Curr Opin Immunol. 2002; 14:312–316. [PubMed: 11973128]
22. Cary LA, Guan JL. Focal adhesion kinase in integrin-mediated signaling. Front Biosci. 1999; 4:D102–113. [PubMed: 9889179]
23. Margolis B, Skolnik EY. Activation of Ras by receptor tyrosine kinases. J Am Soc Nephrol. 1994; 5:1288–1299. [PubMed: 7893993]
24. Lucas B, Germain RN. Unexpectedly complex regulation of CD4/CD8 coreceptor expression supports a revised model for CD4+CD8+ thymocyte differentiation. Immunity. 1996; 5:461–477. [PubMed: 8934573]
25. Liston A, Lesage S, Gray DH, Boyd RL, Goodnow CC. Genetic lesions in T-cell tolerance and thresholds for autoimmunity. Immunol Rev. 2005; 204:87–101. [PubMed: 15790352]
26. Tsukamoto H, et al. B-Raf-mediated signaling pathway regulates T cell development. Eur J Immunol. 2008; 38:518–527. [PubMed: 18228248]
27. Geiman TM, Muegge K. Lsh, an SNF2/helicase family member, is required for proliferation of mature T lymphocytes. Proc Natl Acad Sci U S A. 2000; 97:4772–4777. [PubMed: 10781083]
28. Okada H, et al. Survivin loss in thymocytes triggers p53-mediated growth arrest and p53-independent cell death. J Exp Med. 2004; 199:399–410. [PubMed: 14757745]

29. Xing Z, Conway EM, Kang C, Winoto A. Essential role of survivin, an inhibitor of apoptosis protein, in T cell development, maturation, and homeostasis. J Exp Med. 2004; 199:69–80. [PubMed: 14699085]
30. Bensinger SJ, et al. LXR signaling couples sterol metabolism to proliferation in the acquired immune response. Cell. 2008; 134:97–111. [PubMed: 18614014]
31. Gallo EM, et al. Calcineurin sets the bandwidth for discrimination of signals during thymocyte development. Nature. 2007; 450:731–735. [PubMed: 18046413]
32. Lee MS, Hanspers K, Barker CS, Korn AP, McCune JM. Gene expression profiles during human CD4+ T cell differentiation. Int Immunol. 2004; 16:1109–1124. [PubMed: 15210650]
33. Kose H, et al. Maturational arrest of thymocyte development is caused by a deletion in the receptor-like protein tyrosine phosphatase kappa gene in LEC rats. Genomics. 2007; 89:673–677. [PubMed: 17434290]
34. Asano A, Tsubomatsu K, Jung CG, Sasaki N, Agui T. A deletion mutation of the protein tyrosine phosphatase kappa (Ptprk) gene is responsible for T-helper immunodeficiency (thid) in the LEC rat. Mamm Genome. 2007; 18:779–786. [PubMed: 17909891]
35. Flavell JR, et al. Down-regulation of the TGF-beta target gene, PTPRK, by the Epstein-Barr virus encoded EBNA1 contributes to the growth and survival of Hodgkin lymphoma cells. Blood. 2008; 111:292–301. [PubMed: 17720884]
36. Nakamura M, et al. Novel tumor suppressor loci on 6q22-23 in primary central nervous system lymphomas. Cancer Res. 2003; 63:737–741. [PubMed: 12591717]
37. Wang SE, Wu FY, Shin I, Qu S, Arteaga CL. Transforming growth factor {beta} (TGF-{beta})-Smad target gene protein tyrosine phosphatase receptor type kappa is required for TGF-{beta} function. Mol Cell Biol. 2005; 25:4703–4715. [PubMed: 15899872]
38. Yang Y, et al. Transforming growth factor-beta1 inhibits human keratinocyte proliferation by upregulation of a receptor-type tyrosine phosphatase R-PTP-kappa gene expression. Biochem Biophys Res Commun. 1996; 228:807–812. [PubMed: 8941358]
39. Treeck O, et al. Detection of increased icb-1 transcript levels in maturing HL-60 cells: a novel marker for granulocytic and monocytic in vitro differentiation. Leuk Res. 2002; 26:765–769. [PubMed: 12191572]
40. Hurgin V, Novick D, Werman A, Dinarello CA, Rubinstein M. Antiviral and immunoregulatory activities of IFN-gamma depend on constitutively expressed IL-1alpha. Proc Natl Acad Sci U S A. 2007; 104:5044–5049. [PubMed: 17360358]
41. Treeck O, Kindzorra I, Pauser K, Treeck L, Ortmann O. Expression of icb-1 gene is interferon-gamma inducible in breast and ovarian cancer cell lines and affects the IFN gamma-response of SK-OV-3 ovarian cancer cells. Cytokine. 2005; 32:137–142. [PubMed: 16219472]
42. Jun JE, et al. Identifying the MAGUK protein Carma-1 as a central regulator of humoral immune responses and atopy by genome-wide mouse mutagenesis. Immunity. 2003; 18:751–762. [PubMed: 12818157]
43. Ho WY, Cooke MP, Goodnow CC, Davis MM. Resting and anergic B cells are defective in CD28-dependent costimulation of naive CD4+ T cells. J Exp Med. 1994; 179:1539–1549. [PubMed: 7909325]
44. Tanchot C, Barber DL, Chiodetti L, Schwartz RH. Adaptive tolerance of CD4+ T cells in vivo: multiple thresholds in response to a constant level of antigen presentation. J Immunol. 2001; 167:2030–2039. [PubMed: 11489985]
45. Huesmann M, Scott B, Kisielow P, von Boehmer H. Kinetics and efficacy of positive selection in the thymus of normal and T cell receptor transgenic mice. Cell. 1991; 66:533–540. [PubMed: 1868548]
46. Silver KL, et al. MyD88-dependent autoimmune disease in Lyn-deficient mice. Eur J Immunol. 2007; 37:2734–2743. [PubMed: 17853409]
47. Cornall RJ, et al. Polygenic autoimmune traits: Lyn, CD22, and SHP-1 are limiting elements of a biochemical pathway regulating BCR signalling and selection. Immunity. 1998; 8:497–508. [PubMed: 9586639]
48. Wan F, et al. Ribosomal protein S3: a KH domain subunit in NF-kappaB complexes that mediates selective gene regulation. Cell. 2007; 131:927–939. [PubMed: 18045535]

49. Schaupp CJ, Jiang G, Myers TG, Wilson MA. Active mixing during hybridization improves the accuracy and reproducibility of microarray results. Biotechniques. 2005; 38:117–119. [PubMed: 15679093]

50. Saeed AI, et al. TM4: a free, open-source system for microarray data management and analysis. Biotechniques. 2003; 34:374–378. [PubMed: 12613259]

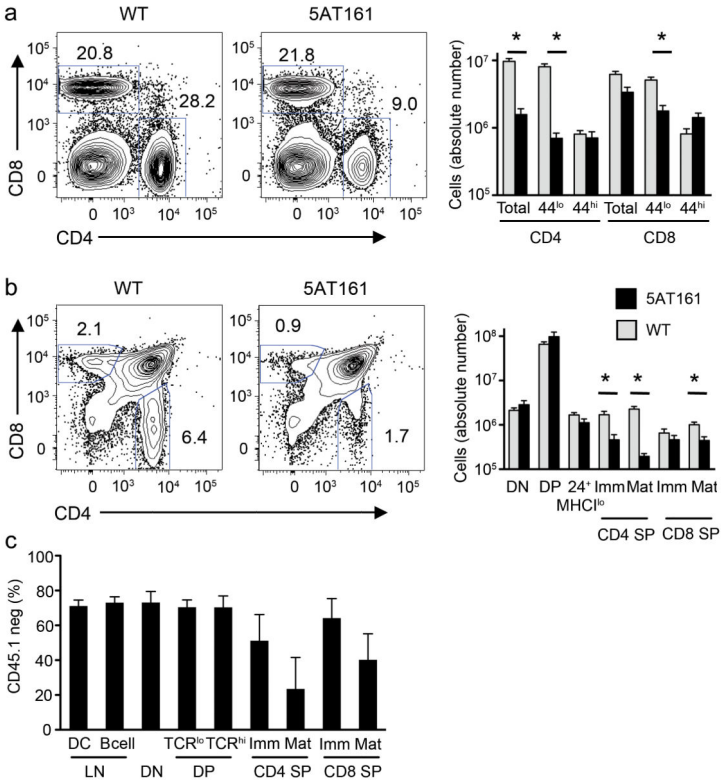

Figure 1. Decreased CD4$^+$ and CD8$^+$ T cell production in 5AT161 mutant mice

Representative (1 of 4 independent experiments) flow cytometry profiles of CD4 and CD8 expression (left) and cellularity (right, shown as mean with standard error) of (**a**) LN T cells and (**b**) thymocyte subsets from 7 week-old mice. * P<0.02. TCRβ^{hi} SP thymocyte subsets were gated as follows: 24$^+$MHC-I^{lo} = CD4$^+$CD8^{lo}CD24$^+$MHC-I^{lo}; immature (Imm) = MHC-I^{hi}CD24$^+$; and mature (Mat) = MHC-I^{hi}CD24^{lo}. (**c**) Percentage of 5AT161 mutant-derived CD45.1$^-$ cells in lymphocyte populations analyzed 8 weeks after CD45.1$^+$ wild-type mice were lethally irradiated and injected i.v. with an equal mix of CD45.1$^-$ 5AT161 mutant and CD45.1$^+$ wild-type BM. Data show geometric mean with 95% confidence interval error bars (n=5). * indicates P<0.02 compared to the TCR^{hi} DP population.

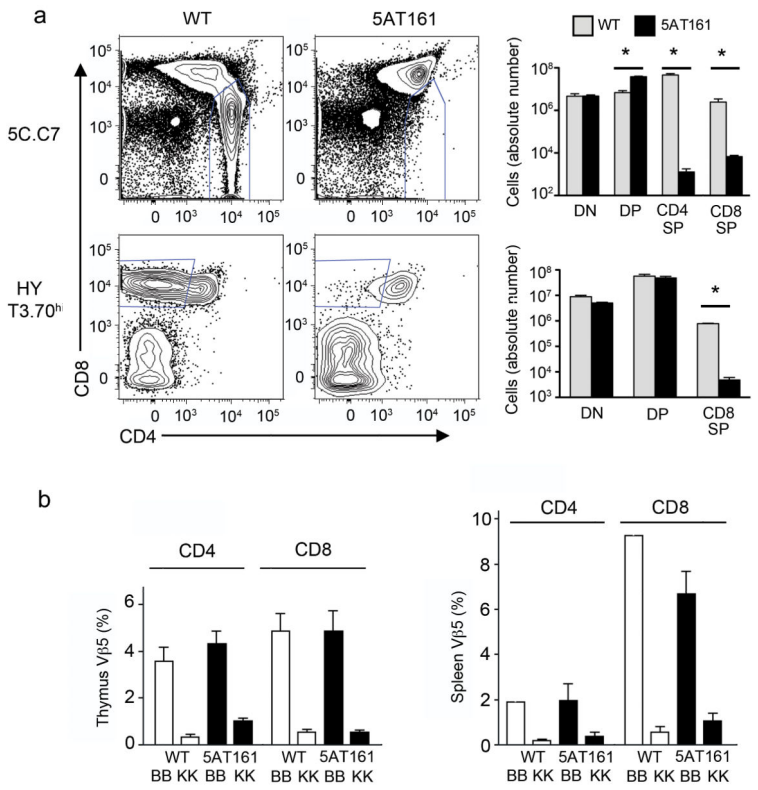

Figure 2. Impaired positive selection and normal negative selection

Representative (1 of 5 experiments for 5C.C7 and 2 experiments for H-Y) flow cytometry profiles of CD4 and CD8 expression (left) and cellularity (right, shown as mean with standard error) of thymocytes from 3 month-old 5C.C7 TCR transgenic mice or female H-Y TCR transgenic mice. * $P<0.01$. (b) The percentage of $V_\beta5^+$ cells within the $CD4^+$ or $CD8^+$ T cell populations from the thymus and spleen of 2-4 month old mutant (Y489X) or wild-type (WT) mice on B6 (BB: non-deleting) and B10.BR (KK: deleting) backgrounds is shown ($n \geq 5$ mice per group).

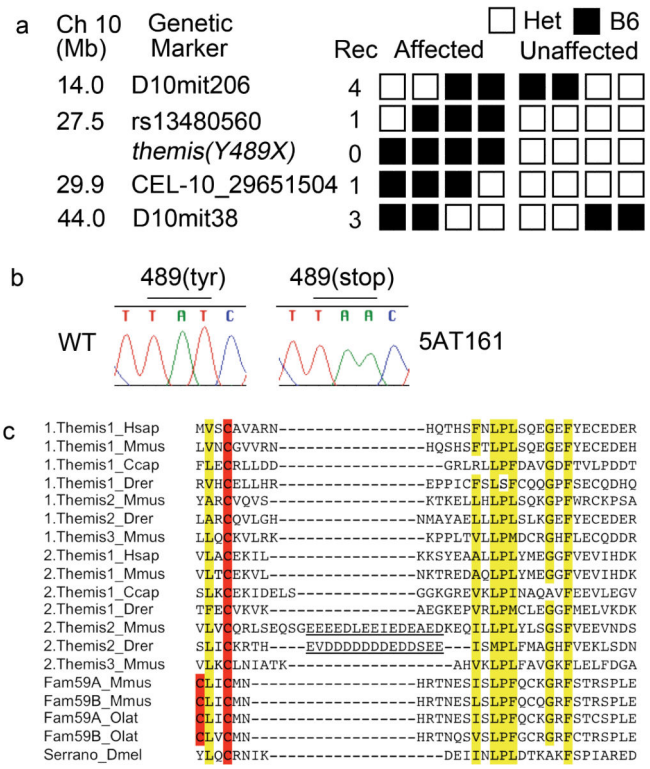

Figure 3. Induced premature stop codon in the *Themis* gene (E430004N04Rik) in 5AT161 mutant mice

(**a**) The critical mapping interval of the 5AT161 mutation with the location of key SNPs and microsatellite markers between B6 and 129 or CBA strains. B6 homozygotes (black squares) and heterozygotes (white squares) are shown for affected and unaffected F2 mice along with the number of recombinants (Rec) observed in affected mice for each genetic marker. (**b**) Sequence trace histograms of the Tyr-489 mutated codon in wild-type and 5AT161 mice. (**c**) Amino acid sequence alignment of the region surrounding the conserved cysteine residue in the CABIT domain(s) of multiple Themis protein family members. Highly conserved hydrophobic residues (yellow), conserved cysteine residue(s) (red), and putative coiled-coil motif (underlined) are highlighted. In the case of the two domain proteins like Themis itself, the two CABIT domains are indicated by the prefix 1 or 2. The organism abbreviations are: Ccap : *Capitella capitata*; Dmel : *Drosophila melanogaster*; Drer : *Danio rerio*; Hsap : *Homo sapiens*; Mmus : *Mus musculus*; Olat : *Oryzias latipes*.

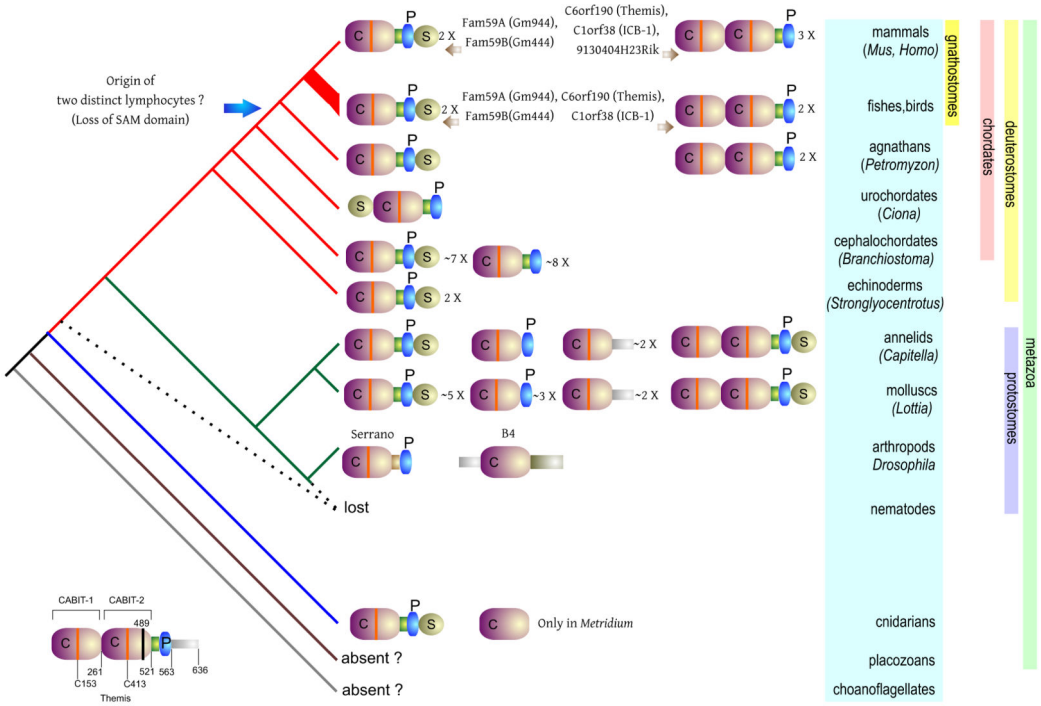

Figure 4. Domain architectures and evolution of the CABIT domain

A domain schematic of Themis showing two globular CABIT domains (purple), a proline-rich region (blue), a conserved helical segment (green), conserved cysteine motifs (orange bars), the site of the mutation (black bar), and the location of key amino acid residues is shown in the lower left corner of the figure. In the rest of the figure, similar domain architectures for all of the CABIT domain proteins have been superimposed on the animal phylogenetic tree. The major animal groups are indicated to the right of the figure. Proteins show the CABIT domain (C), the SAM domain (S), and the low complexity proline-rich stretch (P). The conserved helical segments preceding the proline-rich region are shown in green as distinct elements and extensions unique to particular proteins are shown as grey elements. In gnathostome vertebrates and insects, the names of the orthologs are also indicated. The location of the conserved cysteine in the CABIT domain is shown as an orange bar; those representatives lacking it are presumed to be inactive versions of the domain.

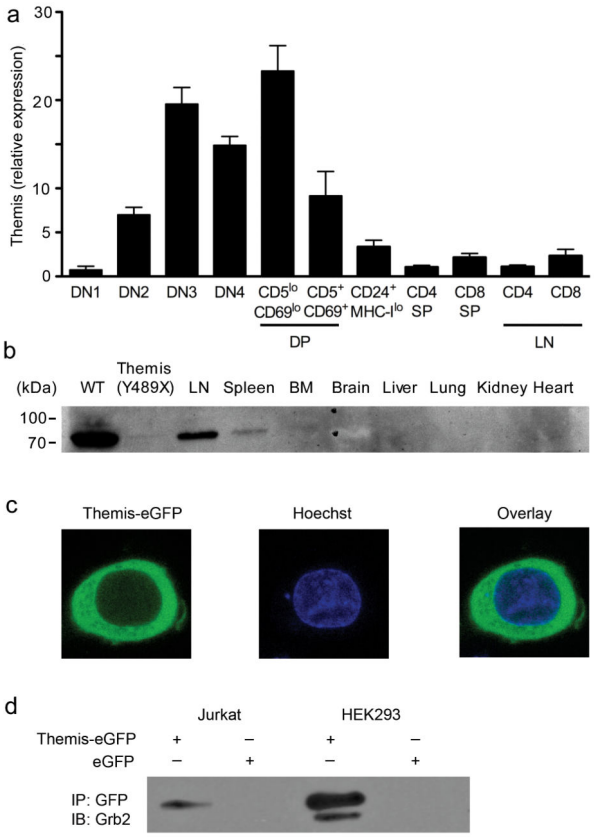

Figure 5. Themis expression

(**a**) RT-PCR expression analysis using *Themis* exon 3 forward and exon 4 reverse primers on sorted T cell populations. Expression was calculated using the $\Delta\Delta C_T$ method with β actin as the standard. Data are shown as the mean fold expression over B cell values for 4-12 biological replicates per population. (**b**) Rabbit polyclonal anti-Themis immunoblot on lysates from unfractionated homogenized tissues of a wild-type B6 and a *Themis(Y489X)* thymus, as well as other tissues from wild-type B6 mice. Data for thymus, LN and spleen are representative of ≥ 3 independent experiments. (**c**) Confocal imaging of unfixed HEK293 cells transiently transfected overnight with Themis-EGFP (green) and labeled with Hoechst 33342 nuclear dye (blue). Original magnification x63. Data are representative of 4 independent experiments. (**d**) Anti-Grb2 immunoblot showing specific co-immunoprecipitation with Themis-EGFP in both Jurkat and HEK293 cell lines that had been transiently transfected overnight with Themis-EGFP or EGFP control constructs. Data are representative of 3 independent experiments.

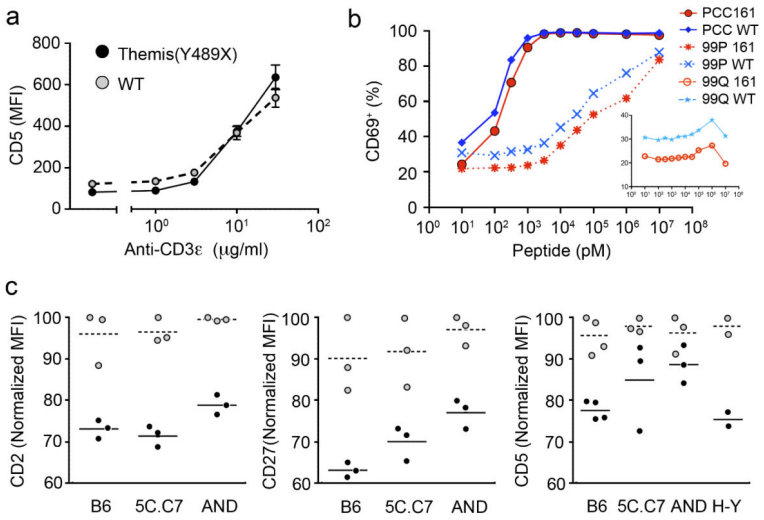

Figure 6. Normal TCR-driven upregulation of activation markers and decreased expression of development markers in *Themis* mutant DP thymocytes

(**a**) Mean fluorescence intensity (MFI) of CD5 expression on *Themis(Y489X)* (black with solid line) and wild-type (gray with dashed line) DP thymocytes after overnight stimulation with varying doses of plate-bound anti-CD3ε. The data points to the left of the break are for unstimulated samples. Live DP cells were gated based on forward scatter and the absence of 7AAD staining. (**b**) Percentage of $CD69^+$ cells after overnight stimulation of *Themis(Y489X)* (red) or wild-type (blue) AND thymocytes cultered 2:1 with P13.9 cells24 and varying doses of the agonist ($PCC_{88\text{-}104}$), the weak agonist (K99P), or the very weak agonist (K99Q, insert) peptide ligands. Live DP cells were gated as in (a). Data are representative of 3 independent experiments for both (a) and (b). (**c**) MFI expression of the indicated cell surface proteins on $TCR\beta^{lo/mid}$ DP thymocytes from non-transgenic (B6) and the indicated TCR transgenic mice. Values were normalized to the maximum MFI within each experiment. Horizontal bars are the mean for each genotype. Data are representative of 2 independent experiments for the H-Y mice and ≥5 experiments for all others.

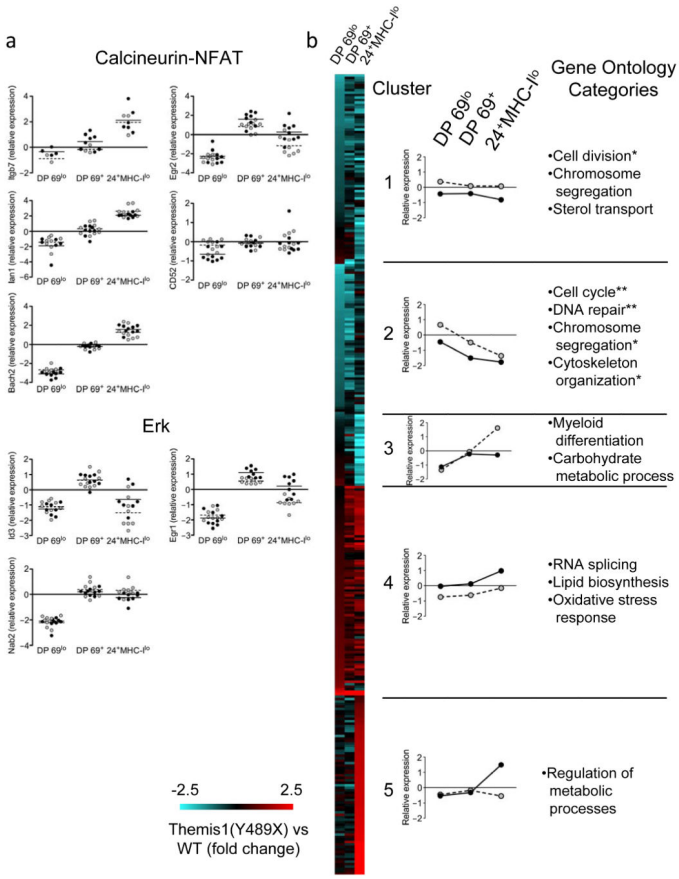

Figure 7. Altered expression of survival, cell cycle and lipid metabolism genes in *Themis(Y489X)* thymocytes

(**a**) Log_2 expression of Erk and NF-AT target genes in wild-type (gray circles with mean as dotted line) and *Themis(Y489X)* (black circles with mean as solid line) thymocyte populations. Data taken from gene-specific probes are relative to a reference of mixed wild-type and mutant samples from all populations. (**b**) Microarray gene expression analysis identified 325 differentially expressed transcripts between *Themis(Y489X)* and wild-type thymocytes that were separated into 5 clusters based on differential expression between mutant and wild-type cells as well as expression changes between populations in wild-type samples (see Supplementary Table 1). Gene Ontology categories enriched in each cluster (*P<0.001, **P<10^{-5}, and P<0.01 for all others), and median relative expression of genes in each cluster in mutant (black with solid line) and wild-type (gray with dotted line) samples are shown. The fold-change between averaged mutant and wild-type expression in each population (n = 8 or 9) is displayed in the heatmap for each transcript.