Impaired Activated/Memory Regulatory T Cell Clonal Expansion Instigates Diabetes in NOD Mice

Vanessa Mhanna, Gwladys Fourcade, Pierre Barennes, Valentin Quiniou, Hang P Pham, Paul-Gydeon Ritvo, Faustine Brimaud, Bruno Gouritin, Guillaume Churlaud, Adrien Six, et al.

Impaired activated/memory regulatory T cell clonal expansion instigates diabetes in NOD mice

Vanessa Mhanna[a], Gwladys Fourcade[a], Pierre Barennes[a], Valentin Quiniou[a,b], Hang P. Pham[c], Paul-Gydeon Ritvo[a], Faustine Brimaud[a], Bruno Gouritin[a], Guillaume Churlaud[a,b], Adrien Six[a,1], Encarnita Mariotti-Ferrandiz[a,1], David Klatzmann[a,b,1*]

[a]Sorbonne Université, INSERM, Immunology-Immunopathology-Immunotherapy (i3), F-75005 Paris, France.
[b]AP-HP, Hôpital Pitié-Salpêtrière, Biotherapy (CIC-BTi) and Inflammation-Immunopathology-Biotherapy Department (i2B), F-75013, Paris, France.
[c]ILTOO Pharma, Statistics Department, Paris, France.

[1]co-senior authors

*corresponding author
Email: david.klatzmann@sorbonne-universite.fr
Phone number: +33 1 42 17 74 60

Word count: 3994
Number of figures: 5
Keywords
Tolerance, immune regulation, immunopathology, immunotherapy, systems biology

Abstract

Regulatory T cell (Treg) insufficiency licenses the destruction of insulin-producing pancreatic β cells by auto-reactive effector T cells (Teffs), causing spontaneous autoimmune diabetes in non-obese diabetic (NOD) mice. We investigated the contribution to diabetes of the TCR repertoires of naive regulatory T cells (nTregs), activated/memory Tregs (amTregs), and $CD4^+$ Teffs from prediabetic NOD mice and normal C57BL/6 (B6) mice. NOD mice amTreg and Teff repertoire diversity was unexpectedly higher than that of B6 mice. This was due to the presence of highly expanded clonotypes in B6 amTregs and Teffs that were largely lost in their NOD counterparts. IL-2 administration to NOD mice restored such amTreg clonotype expansions and prevented diabetes development. In contrast, IL-2 administration only led to few or no clonotype expansions in nTregs and Teffs, respectively. Noteworthily, IL-2 expanded amTreg and nTreg clonotypes were markedly enriched in islet-antigen specific TCRs. Altogether, our results highlight the link between a reduced clonotype expansion within the activated Treg repertoire and the development of an autoimmune disease. They also indicate that the repertoire of amTregs is amenable to rejuvenation by IL-2.

Non-obese diabetic (NOD) mice spontaneously develop autoimmune diabetes due to the destruction of insulin-producing β cells by auto-reactive CD4+ and CD8+ T cells (*1*). Diabetes susceptibility in NOD mice is attributed to genetic defects encoded by the insulin-dependent diabetes (idd) loci including over 40 genes, among which the MHC class II and interleukin-2 (IL-2) genes have major contributions to the pathophysiological immune response leading to diabetes (*2*). The unique I-A^{g7} MHC-II molecule expressed by NOD mice has been reported to bind certain peptides with low affinity, thus impairing the thymocyte negative selection process (*3*,*4*). Reduced IL-2 production by activated effector T cells (Teffs) was found in NOD mice and linked to an impairment of Treg numbers and functionality (*5*-*7*). An increased diabetes incidence was observed in NOD mice bearing an IL-2 alpha receptor subunit with reduced affinity for IL-2, which was associated with a Treg but not a Teff defect (*8*). Likewise, the administration of low-dose IL-2 to NOD mice promotes Treg expansion and activation, prevents diabetes onset and even cures diabetic mice (*9*, *10*). Thus, the altered tolerance promoting diabetes in NOD mice could be the result of (i) a defect in thymic selection of Treg and Teff T-cell receptor (TCR) repertoires and/or (ii) an altered Treg fitness due to IL-2 deprivation.

Previous studies of the TCR repertoires of NOD mice and type 1 diabetes patients yielded inconsistent results. Some studies showed a restricted TRBV and TRBJ chain usage by $CD4^+$ T cells from pancreatic islets and pancreatic lymph nodes (PLN), indicating that overrepresented islet-specific TCRs preferentially express certain TRBVJ genes (*11*–*13*). Conversely, others showed a rather diverse gene usage among islet-infiltrating $CD4^+$ T cells in NOD mice (*14*), including islet-infiltrating memory $CD4^+$ T cells (*15*, *16*), and no common highly used TRBVBJ combinations in T cells from PLN of type 1 diabetes patients (*17*). Moreover, NOD Tregs were found to have a restricted TRA repertoire and limited overlap with thymic $CD4^+$ T cells (*18*, *19*), while we found similar TRB diversity between NOD peripheral Tregs and Teffs (*20*). These somehow conflicting results can be explained by the fact that (i) most of these studies were performed on unsorted $CD4^+$ T cells and thus did not distinguish between regulatory and effector populations and/or (ii) most Treg repertoire analyses were done on a limited fraction of the repertoire by using TCR transgenic NOD mice or by focusing on few rearrangements (*18*, *19*). More importantly, most studies have overlooked the existence of regulatory T cell subsets.

Indeed, we previously showed that Tregs could be phenotypically divided into $CD4^+Foxp3^+CD44^{low}CD62L^{high}$ naïve Tregs (nTregs) that have a low turnover, and $CD4^+Foxp3^+CD44^{high}CD62L^{low}$ activated/memory Tregs (amTregs) that have a very rapid turnover due to their interaction with their cognate antigens (*21*). In the healthy B6 mouse background, we showed that amTregs are enriched in deep tissue-draining lymph nodes (LNs) and are characterized by unique clonotype expansions (*22*). Moreover, we showed little overlap between amTreg and nTreg TRB repertoires (*22*). Thus, studying amTreg and nTreg TCR repertoires without separating them may be confounding.

Here we compared the TCR repertoire of splenic amTregs, nTregs and Teffs of prediabetic NOD and normal B6 mice to investigate whether and how an altered TCR repertoire contributes to diabetes. We found that NOD amTregs have an unexpectedly more diverse repertoire than that of B6 mice, due to the absence of expanded clonotypes. Treatment of prediabetic NOD mice with IL-2, which prevents diabetes occurrence, restored the expansion of clonotypes that are preferentially diabetes-reactive. Thus, the

TCR repertoire of amTregs is altered by an IL-2 insufficiency, which instigates diabetes development in NOD mice, and is amenable to therapeutic recovery.

Research Design and Methods

Study design

Given the rarity of some of the studied cell subsets, we pooled the spleens of 6 to 8 young mice matched for genetic background, age and sex. Six (three female and three male) non-treated C57BL/6 or NOD pools of mice were used to study the repertoire at homeostasis. Six young female NOD mice were treated with IL-2. Sample size was determined for statistical significance and experimental feasibility. Investigators were not blinded to the allocations during experiments and analyses.

Mice

Eight- to ten-week-old male and female C57BL/6- and NOD-Foxp3-EGFP transgenic mice expressing green fluorescent protein (GFP) under the control of the Foxp3 gene promoter were respectively provided by B. Malissen (Luminy, Marseille) and V. Kuchroo (Brigham and Women’s Hospital, Boston, MA). All animals were maintained at the Sorbonne Université Centre d’Expérimentation Fonctionnelle animal facility under specific pathogen-free conditions in agreement with the current European legislation on animal care, housing and scientific experimentation (agreement number A751315). All procedures were approved by the local animal ethics committee.

IL-2 treatment

Recombinant AAV8 vectors were produced as described previously (*9*). Six female NOD-Foxp3-EGFP mice were injected once intraperitoneally with 10^{11} of rAAVs diluted in PBS at 6 weeks of age. Serum was collected two weeks post-injection and IL-2 levels were measured using a mouse IL-2 ELISA (eBioscience) according to the manufacturer's recommendations. Mice were sacrificed at nine weeks of age.

Spleen retrieval and cell sorting

Spleens were collected from euthanized mice and splenocytes were stained with anti-Ter-119-biotin (BD Biosciences), -B220-biotin (eBioscience) and -CD11c-biotin (eBioscience) antibodies for 20 min at 4°C and labeled with anti-biotin magnetic beads (Miltenyi Biotec) for 15 min at 4°C. B cells, erythrocytes and monocytes were depleted on an AutoMACS separator (Miltenyi Biotec) following the manufacturer's procedure. Enriched T cells were stained for 20 min at 4°C with the following monoclonal antibodies at predetermined optimal dilutions: CD3-APC (BD Biosciences), CD4-Horizon-V500 (BD Biosciences), CD8-Alexa-700 (BD Biosciences), CD44-PE (BD Biosciences) and CD62L-eFluor-450 (eBioscience) and sorted on a FACSAria II cytometer (BD Biosciences) with a purity > 95% into the following subsets: $CD4^{+}$ $FoxP3^{+}$ $CD62^{low}$ $CD44^{high}$ (activated/memory Tregs (amTregs)), $CD4^{+}$ $FoxP3^{+}$ $CD62^{high}$ $CD44^{low}$ (naive Tregs (nTregs)) and $CD4^{+}$ $FoxP3^{-}$ (T effector cells (Teffs)). 1.10^{5} to 5.10^{6} sorted cells were stored in RNAqueous kit lysis buffer (Invitrogen) at -80 °C.

cDNA library preparation for TR sequencing

RNA was extracted using the RNAqueous Total RNA Isolation Kit (Invitrogen). TRB libraries were prepared on 100 ng of RNA with the SMARTer Mouse TCRa/b Profiling Kit (Takarabio) and sequenced with HiSeq 2500 single read (300 bp) (Illumina) + 10% PhiX at the "LIGAN Genomics platform" (23).

Raw read processing

Raw data in FASTQ format were processed for TRB sequence annotation using MiXCR software (v.3.0.3) (*24*), which ensures PCR and sequencing error corrections and provides for each sample a list of unique amino acid TRBs and their corresponding counts. Annotated sequences were then analyzed using an in-house workflow to eliminate long amino acid CDR3 sequences (length higher than the mean + 8). The resulting datasets are summarized in Table S2.

Data analysis

Statistical comparison

Sample overlap was tested by computing the Jaccard distance (*25*) using the "vegan" R package and plotted into a heatmap using the "pheatmap" package (*26*). Clustering analysis was performed using the 'Ward' method. Principal component analysis was performed using the "stats" R package. Renyi entropy (*27*) was calculated using "vegan" and represents the distribution of clonal expansions as a function of the parameter *alpha*. Probability density plots were plotted using the "ggplot2" R package *(28)*.

Network analysis

For each cell population, the following method was adopted: The top most frequent 1000 CDR3 amino acid sequences from each of the six B6 or NOD mice were pooled and the mean frequency was calculated for every sequence. The list was then ordered decreasingly and the first 1000 sequences were selected. A matrix of pairwise Levenshtein distances (LD) between the CDR3s was computed and used to construct the networks. In these networks, vertices (CDR3 sequences) are connected by edges (LD=1) only if they differ by one amino acid (insertion/deletion/substitution). A connected cluster is defined as a set of two or more vertices connected by edges. The clustering coefficient of a node in a network quantifies how close its neighbors are to being a clique (a closed cluster). The maximum value is 1 if every neighbor connected to the node A is also connected to every other node within the neighborhood. The node size represents the mean frequency of each CDR3 in the network.

Analyses were performed using R packages: "stringdist" *(29)* was used to calculate LDs and "Igraph" *(30)* to calculate network properties. Graphics visualization were done using Cytoscape *(31)* and "ggplot2" for figures 3 and 5 respectively. In the latter, only clustered nodes are represented and edges are not shown.

Identification of expanded clonotypes

In each sample, the first (Q1) and third (Q3) quartiles and the interquartile range (IQR) were computed according to a negative binomial distribution fitted in the sequence read counts, while excluding private clonotypes with a count of 1. Expanded clonotypes are defined as those with counts greater than Q3 + 1.5 x IQR.

TCR Database

Mouse CDR3βs associated with diabetes (n=52), EAE (n=50), SLE (n=49) and cancer (n=52) were collected from Friedman's published database (*32*). These sequences were derived from $CD4^+$ T cells and identified by different methods including reactivity assays, tetramer staining, selection of diabetes-specific TCRs following diabetes induction and by sequencing of tumor T cell infiltrates in the context of cancer studies.

Statistical analyses

Statistical analyses were performed using the nonparametric Wilcoxon test with Holm multiple testing correction using "ggpubr" R package *(33)*. A p value of ≤ 0.05 was considered statistically significant; $*$ $p \leq 0.05$, $**$ $p \leq 0.01$, $***$ $p \leq 0.001$, $****$ $p \leq 0.0001$ and n.s. denotes not significant ($p > 0.05$). Mean percentage comparisons of diabetes, EAE, SLE and cancer CDR3s were done using the Wilcoxon paired test with diabetes as a reference group.

Data availability

Source data are provided in the manuscript and the Supplementary materials. Fastq data were deposited in NCBI Sequence Read Archive repository under the BioProject ID PRJNA635928.

Results

Differences in repertoire composition between B6 and NOD

We analyzed and compared the TCR beta chain (TRB) repertoire of splenic CD4$^+$Foxp3$^+$CD44^{low}CD62L^{high} nTregs, CD4$^+$Foxp3$^+$CD44^{high} CD62L^{low} amTregs and Foxp3$^-$CD4$^+$ Teffs in B6 and NOD mice.

Compared to B6 mice, flow cytometry showed a significant decrease in the NOD Treg population ($p \leq 0.05$) associated with a decrease of amTreg proportions ($p \leq 0.01$) (Table S 1). Cell numbers for each sorted population and their corresponding number of TRB sequences and unique amino acid clonotypes are summarized in Table S2.

Principal component analysis of TRBVBJ usage, i.e. the frequency of each combination in the repertoire irrespective of the clonotype abundance, clearly separated the genetic backgrounds and the cell populations (Figure 1A). The first two components showed different TRBVBJ usage between B6 and NOD mice, with the first component (PC1) that separates the two genetic backgrounds explaining more than 68% of the variability. There was also a clear difference between the cell populations from the same background, with the two Treg subsets being close, but quite apart from Teffs. This reflects differences in the combined generation frequency and survival through the thymic selection and peripheral expansion of each generated TRBVBJ recombination.

To further investigate these differences, we assessed the clonotype (i.e. unique amino acid TRBV-CDR3-TRBJ sequence) overlap between samples using the Jaccard distance in which the higher the value, the more dissimilar the pair of samples compared (Figure 1B). Hierarchical clustering on the Jaccard scores accurately separated B6 from NOD mouse samples. Surprisingly, within the latter cluster, amTregs are intermingled with nTregs and Teffs, while, as expected from our previous study (*22*), amTregs clustered separately in the B6 background. Overall, NOD mice display a peculiar repertoire structure compared to B6 mice both at the TRB gene and clonotype level.

Increased TCR diversity of Teffs and amTregs in NOD versus B6 mice

We next explored the overall repertoire diversity at the clonotype level. We compared cell populations and mouse backgrounds by calculating the Renyi entropy, which evaluates diversity as a function of an *alpha* parameter (Figure 2A). The higher the *alpha*, the higher is the weight put on frequent clonotypes. The higher the Renyi entropy, the more diverse is the repertoire.

In B6 mice, nTregs have the most diverse repertoire while amTregs have the least diverse one, and Teffs fall in between. In contrast, while nTregs in NOD mice have a repertoire with a diversity similar to that in B6 mice, the diversities of the Teff and amTreg repertoires are increased compared to B6 mice (Figure 2A). At *alpha*=0, for which the Renyi index represents clonotype richness, no differences were observed between the B6 and NOD backgrounds (Figure 2B, left). This indicates that the overall richness of each cell subset is not affected in NOD mice. At *alpha*=8, which gives more weight to abundant clonotypes, Renyi indices were significantly higher in NOD amTregs and Teffs than in the corresponding subsets in B6 mice (Figure 2B, right).

In view of these differences, we looked at the representativeness of the 100 most predominant clonotypes within each repertoire. The cumulative percentages of the 100 most predominant clonotypes in B6 amTregs and Teffs were significantly higher than their NOD counterparts (Figure 2C). This is due to the presence, in the B6 cell populations, of highly frequent clonotypes depicted by blue lines which thickness is proportional to their percentage in the repertoire (Figure 2D). In contrast, there were no such clonotypes in B6 nor in NOD nTregs.

Collectively, these results reveal that a healthy naïve Treg repertoire is diverse with no clonotype expansions. Conversely, the healthy amTreg repertoire, and to a lesser extent the Teff repertoire, are enriched in clonotypes with high frequency, likely resulting from antigen-driven expansions. In diabetes-prone NOD mice, such expansions are severely reduced in amTregs and Teffs.

Altered similarity structure of NOD amTreg and Teff repertoires

Next, we examined the architecture of the repertoires. To this end, the 1000 most frequent CDR3s per population were used to build similarity networks in which each CDR3 (node) has a size proportional to its frequency and is connected by edges to other CDR3s that differ to it by only one amino acid, i.e. having a Levenshtein distance of 1 (LD=1). This distance was shown to connect CDR3s that are likely to bind the same peptide (*34*).

The structure of the nTreg repertoire appeared similar for B6 and NOD mice, made of highly clustered CDR3s with low frequency (Figure 3A). In contrast, B6 amTreg repertoire was mostly composed of individual or lightly connected with high frequencies, reflecting polyclonal self-antigen-driven expansions (*22,35*). In accordance with the previous observations (Figures 1 and 2), these expansions are much reduced in amTregs from NOD mice. Similarly, node sizes of their Teffs appeared smaller than in B6 mice (Figure 3A, middle).

Node frequency showed no differences for nTregs, a modest decrease for Teffs and a marked decrease for amTregs in NOD mice compared to their counterparts in B6 mice (Figure 3B). The clustering coefficient of each CDR3 within each network, which gives an indication of the importance of these nodes in the global repertoire architecture, showed no significant difference between NOD and B6 nTregs (Figure 3C). However, significantly higher clustering coefficients were observed in Teffs and amTregs from NOD compared to B6 mice.

Overall, the structural analysis of these networks highlights a higher sequence similarity of Teff and amTreg CDR3s in NOD mice. As IL-2 production is notoriously low in NOD mice and Tregs rely primarily on it for proliferation, we hypothesized that the reduced amTreg expansions may result from an IL-2 shortage.

Low dose IL-2 normalizes NOD amTreg repertoires

As administration of IL-2 to NOD mice specifically enables the expansion of Tregs (*9*), we–investigated whether the abnormal repertoire diversity and structure of prediabetic NOD mice could be restored by an IL-2 treatment. NOD mice were injected with an adeno-associated virus coding for IL-2 (AAV-IL-2) that allows the continuous production of IL-2 at low dose and prevents the occurrence of diabetes (*9*). As expected, total Treg proportions increased significantly after IL-2 treatment (Figure 4A). Importantly, within the Treg subsets, amTreg proportions showed a significant increase, at the expense of nTregs (Figure 4A, Table S 3).

We next analyzed the expansions at the clonotype level after IL-2 treatment. To this end, we identified expanded clonotypes as the ones showing counts greater than the Q3 + 1.5 x IQR value, where Q3 is the third quartile and IQR the interquartile range of sequence count distribution (see the Methods section and Table S 4). The abundance of the expanded clonotypes from each sample was plotted in Figure 4B (left plots). Each circle represents a clonotype whose size is proportional to its frequency in the repertoire. amTregs and nTregs from AAV-IL-2-treated NOD mice were significantly enriched in expansions compared to amTregs and nTregs from B6 or NOD mice. This is confirmed by the density plots, which show that IL-2 treatment induces expansions of nTregs and amTregs compared to their B6 and NOD counterparts, but not of Teffs (Figure 4B, right plot). Thus, IL-2 administration restores amTreg clonotype expansions that were absent in NOD mice.

IL-2-expanded Tregs are enriched in diabetes-specific clonotypes

We next examined whether the expanded clonotypes could be linked to diabetes. We queried a published catalogue of TCRs with defined specificity *(32)*. We identified a total of 203 mouse CDR3ß amino acid sequences from CD4+ T cells associated with diabetes (n=52), experimental autoimmune encephalomyelitis (EAE; n=50), systemic lupus erythematosus (SLE; n=49) and cancer (n=52). We then searched for these sequences within the expanded clonotypes and calculated their mean enrichment in each cell subset (Figure 5A). Interestingly, diabetes-specific sequences showed higher expansions compared to non-diabetes-specific sequences in nTregs and amTregs from IL-2–treated NOD mice. There was also a modest enrichment of these sequences in B6 amTregs. To further analyze the specificity pattern towards diabetes, we connected disease-specific CDR3s by LD=1 to the pool of expanded amTreg and nTreg CDR3s. The number of neighbors with LD=1 to diabetes-related sequences showed a significant increase compared to other disease-specific CDR3s (Figure 5B)**.**

Noteworthily, clonotypes identified as specific for islet antigens such as insulin and glutamic acid decarboxylase by tetramer staining were detected in large clusters of IL-2 expanded clonotypes (Figure 5C, and Table S5)**.**

Discussion

In this study, we report an altered amTreg TCR repertoire in NOD mice that can be restored to normal by an IL-2 treatment. IL-2 expanded Tregs were enriched in islet-antigen-specific TCRs and were associated with protection from diabetes. These results have both heuristic and applied implications.

The structure of the am- and n-Treg repertoires in a healthy mouse background

In normal B6 mice, a control model that do not spontaneously develop any inflammatory or autoimmune diseases, nTreg repertoires are very diverse, even more than those of Teffs, with most clonotypes forming dense networks. These are the attributes of an unmodified and tightly connected post-thymic selection repertoire (*36*). In contrast, amTreg repertoires have a reduced diversity and are enriched in frequent clonotypes with few connections. This illustrates the responses to distinct self-antigens, as expected from a cell population that is constantly activated to control autoimmunity in numerous tissues. Such similarity structures have also been described in the context of immunization and aging (*34*).

In addition, amTreg repertoires show a higher overlap with each other than with nTregs or Teffs, indicating that the self-antigens to which amTregs respond are shared between individuals. This is also in agreement with our previous report of a low sequence overlap between the two Treg subsets in the peripheral lymph nodes (*22*). The comparison with Treg TRB repertoires from the NOD autoimmune-prone mouse strain strengthens this conclusion. Reduction of amTreg major expansions is associated with disease susceptibility, and their recovery with disease prevention.

Collectively, these results identify amTregs as the major subset involved in the pathogenesis of autoimmunity and warrant separate analysis of amTregs and nTregs when investigating their roles in health or disease.

amTreg activation depends on IL-2

The severe reduction in frequent clonotypes found in NOD mice, and their appearance under IL-2 provision, highlight a major role of IL-2 in supporting amTreg survival.

NOD mice have a genetically determined IL-2 insufficiency, which is also reported in human type 1 diabetes as well as in SLE and rheumatoid arthritis (*37*,*38*). It is thus likely that the observed defect in amTreg fitness contributes to the pathophysiology of these diseases. If so, the good news is that disease-relevant amTregs~~,~~ can be regenerated by IL-2 treatment. We show herein that one injection of rAAV-IL-2 in prediabetic mice was enough to expand NOD amTregs at the cell and repertoire levels. In our experiments, the repertoire was analyzed 21 days after the initiation of treatment. As the approximate duration of T cell differentiation in the thymus is of 28 days (*39*), newly generated thymocytes are unlikely to contribute to the expanded clonotypes in the amTreg population. Thus, the IL-2-induced amTregs most probably originate from peripheral amTregs and/or nTregs that received their activation signal in the presence of IL-2. The diabetes-related TCR expansions in both nTregs and amTregs further support that both populations contribute to the restoration of a more balanced and complete disease-relevant amTreg repertoire.

Treg specificity to diabetes-related antigens

The existence of public databases containing TCRs with known peptide specificities allowed us to investigate the presence of antigen-specific sequences in our repertoires. Diabetes-related CDR3s identified from the database were overrepresented in IL-2–treated NOD Tregs and most connected in LD=1 networks, as shown by their high number of neighbors. Linking the CDR3 similarity networks to annotated TCR datasets was previously described by Madi et al., and showed that CDR3s with similar specificity tend to cluster together (*34*). This indicates that diabetes-specific CDR3s could have core motifs engaged in pancreatic antigen recognition, and shows the potential of such a tool in identifying novel pathogenic TCRs. Indeed, repertoire analysis of islet-infiltrating T cells revealed an accumulation of TCRs that were reactive in vitro to uncharacterized autoantigens, distinct from the well-known glutamic acid decarboxylase and insulin antigens (*13*)**.**

These observations highlight a relationship between the specific expansion of pancreatic antigen-related TCRs and the prevention of diabetes post IL-2 treatment.

Contribution of I-A^{g7} to NOD pathophysiology

It was hypothesized that diabetes development in NOD mice is associated with the I-A^{g7} molecule that they express. Co-expression with a different type of MHC or mutations within the I-A^{g7} molecule were shown to prevent insulitis and diabetes (*4*, *40*). On the other hand, substitution of non-MHC regions from the C57BL/10 into the NOD strain was sufficient to prevent insulitis or diabetes, despite their expression of the I-A^{g7} molecule (*41*). Our studies highlight the importance of non-MHC genes in diabetes onset in the NOD mouse model. Indeed, we report here that IL-2 administration to pre-diabetic NOD mice, previously shown to prevent the disease outcome, improves amTreg fitness. This would argue that NOD Treg repertoires have the full potential to control autoimmune responses, but fail to do so because an IL-2 deficiency impairs the amTreg repertoire that can control islet autoreactivity.

However, our results do not exclude a potential enrichment of autoreactive T cells in the NOD mice repertoire that would take the lead in the periphery where Tregs fail to expand. In this regard, NOD mice have an overrepresentation of CDR3β sequences carrying hydrophobic amino acid doublets at positions 6 and 7 that have been shown to confer high self-reactivity (*42*). Further studies of the TCR repertoire selection process in the thymus of NOD mice should help address this important issue.

Altogether, our results reveal a link between a reduced expansion of clonotypes within the activated Treg TCR repertoire and the development of an autoimmune disease.–IL-2 treatment not only activates Tregs, but also restores a repertoire of TCRs with relevant specificities. Thus, our work indicates that fit amTregs are key to protection from autoimmune diseases and highlights the value of studying these cells separately from the whole Treg population. Importantly, amTreg repertoire abnormalities and restoration are detected in the spleen, which affords a readout that is more representative of the circulation than pancreatic lymph nodes and thus has a translational value. Our results should guide further studies to elucidate the dynamics of pancreatic antigen-specific amTregs in diabetic versus non-diabetic NOD mice, including in mice outside of a pathogen-free environment, and identify TCR-based biomarkers of diabetes.

Acknowledgments

We thank F. Jebbawi and L. Barientos for their help in sample collection and cell sorting, M. Mendez and N. Papaya for their help in RNA extraction and library preparations, and K. El Soufi for his help in network analyses. We thank the staff of the animal care facility for their excellent handling of mice. We also thank the UMR 8199 LIGAN-PM Genomics platform (Lille, France) which belongs to the 'Federation de Recherche' 3508 Labex EGID (European Genomics Institute for Diabetes; ANR-10-LABX-46) and was supported by the ANR Equipex 2010 session, for performing the sequencing of our libraries.

Funding

This work was funded by the TRiPoD ERC-Advanced EU (No. 322856) and RHU iMAP (ANR-16-RHUS-0001) grants to DK, with additional support from the Sorbonne Université, INSERM and AP-HP. ILTOO pharma supported HPP.

Author contributions

VM, GF, PB, VQ, GC, FB, BG, P-GR and EM-F performed the experiments. VM analyzed the data, with contributions from all authors. VM and HPP performed the statistical analyses. VM, EM-F, AS and DK conceived the experiments. VM, AS, EM-F and DK wrote the manuscript, with input from all authors. DK conceptualized and supervised the study and is the guarantor for the contents of the article.

Duality of interest

DK is an investor of a patent related to the use of IL-2 in the treatment of autoimmune diseases, owned by his public institution and licensed to ILTOO pharma, in which DK has interests. The other authors declare no competing financial interests.

References

1. J. Verdaguer, D. Schmidt, A. Amrani, B. Anderson, N. Averill, P. Santamaria, Spontaneous autoimmune diabetes in monoclonal T cell nonobese diabetic mice. *J. Exp. Med.* **186**, 1663–1676 (1997).

2. Y.-G. Chen, C. E. Mathews, J. P. Driver, The Role of NOD Mice in Type 1 Diabetes Research: Lessons from the Past and Recommendations for the Future. *Front. Endocrinol.* **9**, 51 (2018).

3. O. Kanagawa, B. A. Vaupel, G. Xu, E. R. Unanue, J. D. Katz, Thymic positive selection and peripheral activation of islet antigen-specific T cells: separation of two diabetogenic steps by an I-A(g7) class II MHC beta-chain mutant. *J. Immunol.* **161**, 4489–4492 (1998).

4. T. Lund, L. O'Reilly, P. Hutchings, O. Kanagawa, E. Simpson, R. Gravely, P. Chandler, J. Dyson, J. K. Picard, A. Edwards, D. Kioussis, A. Cooke, Prevention of insulin-dependent diabetes mellitus in non-obese diabetic mice by transgenes encoding modified I-A β-chain or normal I-E α-chain. *Nature*. **345**, 727–729 (1990).

5. D. Klatzmann, A. K. Abbas, The promise of low-dose interleukin-2 therapy for autoimmune and inflammatory diseases. *Nature Reviews Immunology*. **15**, 283–294 (2015).

6. S. A. Long, K. Cerosaletti, P. L. Bollyky, M. Tatum, H. Shilling, S. Zhang, Z.-Y. Zhang, C. Pihoker, S. Sanda, C. Greenbaum, J. H. Buckner, Defects in IL-2R signaling contribute to diminished maintenance of FOXP3 expression in CD4(+)CD25(+) regulatory T-cells of type 1 diabetic subjects. *Diabetes*. **59**, 407–415 (2010).

7. E. Sgouroudis, A. Albanese, C. A. Piccirillo, Impact of protective IL-2 allelic variants on CD4+ Foxp3+ regulatory T cell function in situ and resistance to autoimmune diabetes in NOD mice. *J. Immunol.* **181**, 6283–6292 (2008).

8. C. J. Dwyer, A. L. Bayer, C. Fotino, L. Yu, C. Cabello-Kindelan, N. C. Ward, K. H. Toomer, Z. Chen, T. R. Malek, Altered homeostasis and development of regulatory T cell subsets represent an IL-2R-dependent risk for diabetes in NOD mice. *Sci Signal*. **10** (2017), doi:10.1126/scisignal.aam9563.

9. G. Churlaud, V. Jimenez, J. Ruberte, M. Amadoudji Zin, G. Fourcade, G. Gottrand, E. Casana, B. Lambrecht, B. Bellier, E. Piaggio, F. Bosch, D. Klatzmann, Sustained stimulation and expansion of Tregs by IL2 control autoimmunity without impairing immune responses to infection, vaccination and cancer. *Clinical Immunology*. **151**, 114–126 (2014).

10. M. Rosenzwajg, G. Churlaud, R. Mallone, A. Six, N. Dérian, W. Chaara, R. Lorenzon, S. A. Long, J. H. Buckner, G. Afonso, H.-P. Pham, A. Hartemann, A. Yu, A. Pugliese, T. R. Malek, D. Klatzmann, Low-dose interleukin-2 fosters a dose-dependent

regulatory T cell tuned milieu in T1D patients. *Journal of Autoimmunity*. **58**, 48–58 (2015).

11. I. Marrero, A. Vong, Y. Dai, J. D. Davies, T cell populations in the pancreatic lymph node naturally and consistently expand and contract in NOD mice as disease progresses. *Molecular Immunology*. **52**, 9–18 (2012).

12. L. Li, Q. He, A. Garland, Z. Yi, L. T. Aybar, T. B. Kepler, J. A. Frelinger, B. Wang, R. Tisch, Cell-Specific CD4+ T Cell Clonotypes in Peripheral Blood and the Pancreatic Islets Are Distinct. *The Journal of Immunology*. **183**, 7585–7591 (2009).

13. F. J. Baker, M. Lee, Y. -h. Chien, M. M. Davis, Restricted islet-cell reactive T cell repertoire of early pancreatic islet infiltrates in NOD mice. *Proceedings of the National Academy of Sciences*. **99**, 9374–9379 (2002).

14. Nakano, T cell receptor V gene usage of islet beta cell-reactive T cells is not restricted in non-obese diabetic mice. *Journal of Experimental Medicine*. **173**, 1091–1097 (1991).

15. I. Marrero, D. E. Hamm, J. D. Davies, High-Throughput Sequencing of Islet-Infiltrating Memory CD4+ T Cells Reveals a Similar Pattern of TCR Vβ Usage in Prediabetic and Diabetic NOD Mice. *PLoS ONE*. **8**, e76546 (2013).

16. I. Marrero, C. Aguilera, D. E. Hamm, A. Quinn, V. Kumar, High-throughput sequencing reveals restricted TCR Vβ usage and public TCRβ clonotypes among pancreatic lymph node memory CD4 + T cells and their involvement in autoimmune diabetes. *Molecular Immunology*. **74**, 82–95 (2016).

17. H. R. Seay, E. Yusko, S. J. Rothweiler, L. Zhang, A. L. Posgai, M. Campbell-Thompson, M. Vignali, R. O. Emerson, J. S. Kaddis, D. Ko, M. Nakayama, M. J. Smith, J. C. Cambier, A. Pugliese, M. A. Atkinson, H. S. Robins, T. M. Brusko, Tissue distribution and clonal diversity of the T and B cell repertoire in type 1 diabetes (2016),doi:10.1172/jci.insight.88242.

18. C. Ferreira, Y. Singh, A. L. Furmanski, F. S. Wong, O. A. Garden, J. Dyson, Non-obese diabetic mice select a low-diversity repertoire of natural regulatory T cells. *Proceedings of the National Academy of Sciences*. **106**, 8320–8325 (2009).

19. J. Kern, R. Drutel, S. Leanhart, M. Bogacz, R. Pacholczyk, Reduction of T Cell Receptor Diversity in NOD Mice Prevents Development of Type 1 Diabetes but Not Sjögren’s Syndrome. *PLoS ONE*. **9**, e112467 (2014).

20. P.-G. Ritvo, A. Saadawi, P. Barennes, V. Quiniou, W. Chaara, K. El Soufi, B. Bonnet, A. Six, M. Shugay, E. Mariotti-Ferrandiz, D. Klatzmann, High-resolution repertoire analysis reveals a major bystander activation of Tfh and Tfr cells. *Proc Natl Acad Sci USA*. **115**, 9604–9609 (2018).

21. S. Fisson, G. Darrasse-Jèze, E. Litvinova, F. Septier, D. Klatzmann, R. Liblau, B. L. Salomon, Continuous activation of autoreactive CD4+ CD25+ regulatory T cells in the steady state. *J. Exp. Med.* **198**, 737–746 (2003).

22. A.-S. Bergot, W. Chaara, E. Ruggiero, E. Mariotti-Ferrandiz, S. Dulauroy, M. Schmidt, C. von Kalle, A. Six, D. Klatzmann, TCR sequences and tissue distribution discriminate the subsets of naïve and activated/memory Treg cells in mice: Molecular immunology. *European Journal of Immunology*. **45**, 1524–1534 (2015).

23. P. Barennes, V. Quiniou, M. Shugay, E. S. Egorov, A. N. Davydov, D. M. Chudakov, I. Uddin, M. Ismail, T. Oakes, B. Chain, A. Eugster, K. Kashofer, P. P. Rainer, S. Darko, A. Ransier, D. C. Douek, D. Klatzmann, E. Mariotti-Ferrandiz, Benchmarking of T cell receptor repertoire profiling methods reveals large systematic biases. *Nat Biotechnol* (2020)

24. D. A. Bolotin, S. Poslavsky, I. Mitrophanov, M. Shugay, I. Z. Mamedov, E. V. Putintseva, D. M. Chudakov, MiXCR: software for comprehensive adaptive immunity profiling. *Nature Methods*. **12**, 380–381 (2015).

25. P. Jaccard, Étude comparative de la distribution florale dans une portion des Alpes et du Jura. *Imprimerie Corbaz & Comp.* (1901)

26. R. Kolde, pheatmap (2019).

27. R. Alfred, On measures of entropy and information. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (1961), pp. 547–61.

28. H. Wickham, D. Navarro, T. Lin Pedersen, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2nd edition).

29. P.J. van der Loo, Mark, The stringdist Package for ApproximateString Matching. *The R Journal*. Vol. 6/1, (2020)

30. G. Csardi, T. Nepusz, The igraph software package for complex network research (2006).

31. P. Shannon, Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*. **13**, 2498–2504 (2003).

32. N. Tickotsky, T. Sagiv, J. Prilusky, E. Shifrut, N. Friedman, McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics*. **33**, 2924–2929 (2017).

33. Alboukadel, ggpubr R Package: ggplot2-Based Publication Ready Plots (2020)

34. A. Madi, A. Poran, E. Shifrut, S. Reich-Zeliger, E. Greenstein, I. Zaretsky, T. Arnon, F. V. Laethem, A. Singer, J. Lu, P. D. Sun, I. R. Cohen, N. Friedman, T cell receptor repertoires of mice and humans are clustered in similarity networks around conserved public CDR3 sequences. *eLife*. **6** (2017), doi:10.7554/eLife.22057.

35. G. Darrasse-Jèze, A.-S. Bergot, A. Durgeau, F. Billiard, B. L. Salomon, J. L. Cohen, B. Bellier, K. Podsypanina, D. Klatzmann, Tumor emergence is sensed by self-specific CD44hi memory Tregs that create a dominant tolerogenic environment for tumors in mice. *Journal of Clinical Investigation* (2009)

36. Y. Elhanati, A. Murugan, C. G. Callan, T. Mora, A. M. Walczak, Quantifying selection in immune receptor repertoires. *Proceedings of the National Academy of Sciences*. **111**, 9875–9880 (2014).

37. G. D. Kitas, M. Salmon, M. Farr, J. S. Gaston, P. A. Bacon, Deficient interleukin 2 production in rheumatoid arthritis: association with active disease and systemic complications. *Clin. Exp. Immunol.* **73**, 242–249 (1988).

38. K. S. Zier, M. M. Leo, R. S. Spielman, L. Baker, Decreased Synthesis of Interleukin-2 (IL-2) in Insulin-dependent Diabetes Mellitus. *Diabetes*. **33**, 552–555 (1984).

39. V. Thomas-Vaslin, H. K. Altes, R. J. de Boer, D. Klatzmann, Comprehensive Assessment and Mathematical Modeling of T Cell Population Dynamics and Homeostasis. *J Immunol*. **180**, 2240–2250 (2008).

40. L. Gioia, M. Holt, A. Costanzo, S. Sharma, B. Abe, L. Kain, M. Nakayama, X. Wan, A. Su, C. Mathews, Y.-G. Chen, E. Unanue, L. Teyton, Position β57 of I-A ^{g7} controls early anti-insulin responses in NOD mice, linking an MHC susceptibility allele to type 1 diabetes onset. *Sci. Immunol.* **4**, eaaw6329 (2019).

41. D. V. Serreze, M. Bridgett, H. D. Chapman, E. Chen, S. D. Richard, E. H. Leiter, Subcongenic Analysis of the Idd13 Locus in NOD/Lt Mice: Evidence for Several Susceptibility Genes Including a Possible Diabetogenic Role for β2-Microglobulin. *The Journal of Immunology*. **160**, 1472–1478 (1998).

42. B. D. Stadinski, K. Shekhar, I. Gómez-Touriño, J. Jung, K. Sasaki, A. K. Sewell, M. Peakman, A. K. Chakraborty, E. S. Huseby, Hydrophobic CDR3 residues promote the development of self-reactive T cells. *Nature Immunology*. **17**, 946–955 (2016).

Figures

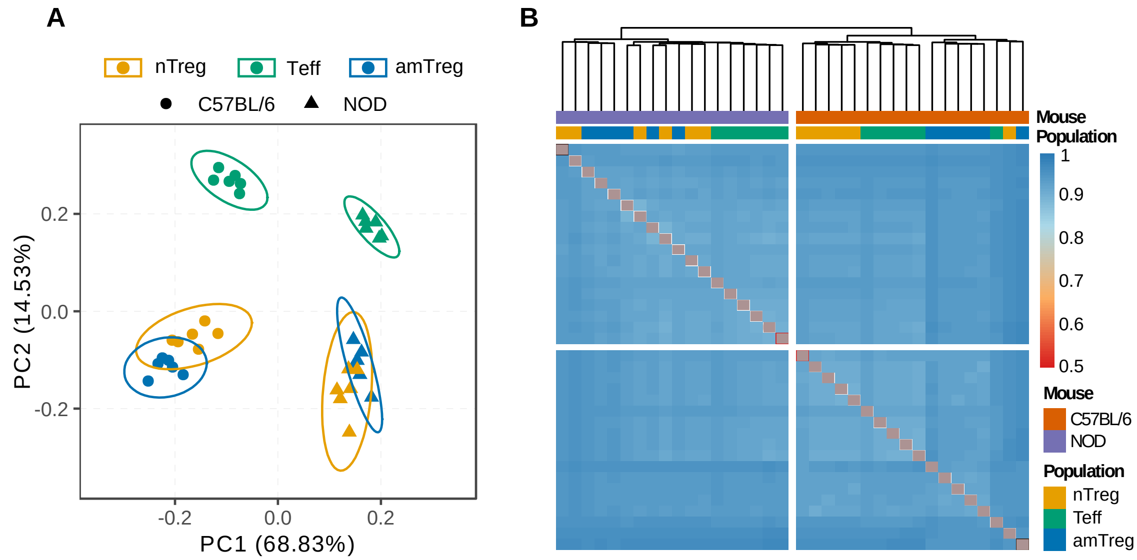

Figure 1. B6 and NOD mice express different TCR repertoires. (A) PCA biplot for TRBVBJ usage of all B6 (dots) and NOD (triangles) samples. Cell populations are identified based on their colors. **(B)** Heatmap of the pairwise Jaccard distances calculated at the clonotype level. Scores range between 0 (complete similarity) and 1 (no similarity). The hierarchical clustering was performed using Ward’s method.

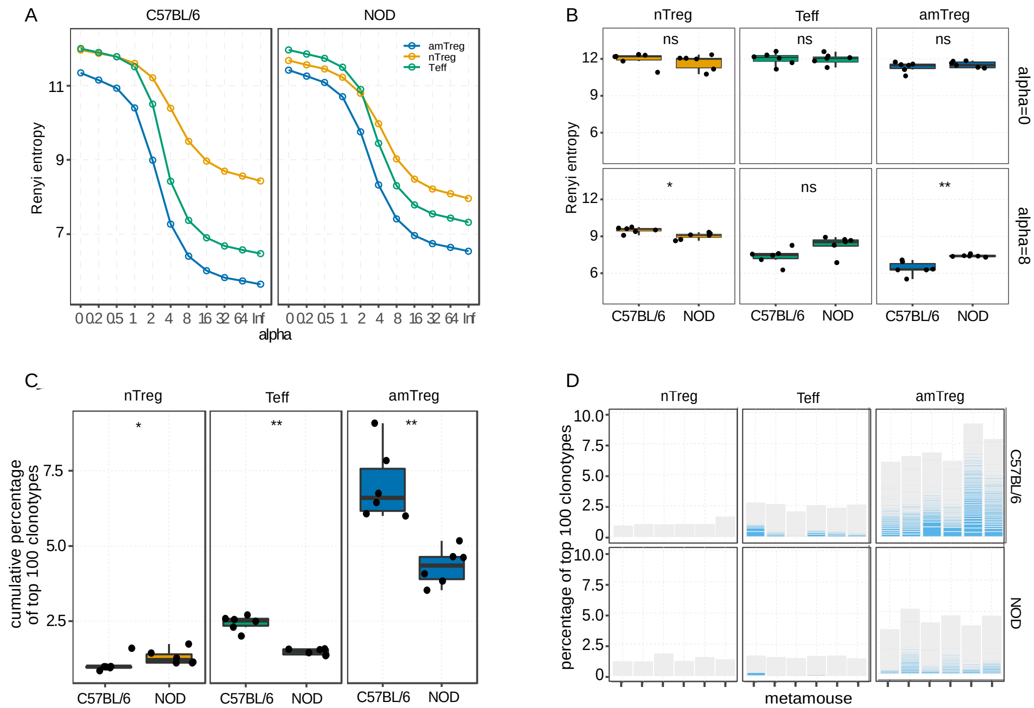

Figure 2. Increased diversity of NOD Teffs and amTregs compared to B6. (A) Renyi profiles for nTregs (yellow), CD4 Teffs (green) and amTregs (blue) in B6 and NOD mice. Points represent the mean of all six mice and shaded areas represent the standard error. **(B)** Renyi values were compared between B6 and NOD mice within each cell population at *alpha* = 0 and *alpha* = 8. **(C)** Mean cumulative percentage of the 100 most predominant clonotypes of the six mice for each cell population. **(D)** Percentage of the top 100 clonotypes within the three cell populations in B6 (upper plot) and in NOD (lower plot) backgrounds. Each blue line and its thickness represent a unique clonotype and its occurrence, grey "bars" are the stacking of slim grey lines representing rare clonotypes. (B, C) n=6, ns $P > 0.05$, * $P \leq 0.05$, ** $P \leq 0.01$, Wilcoxon test with Holm's correction.

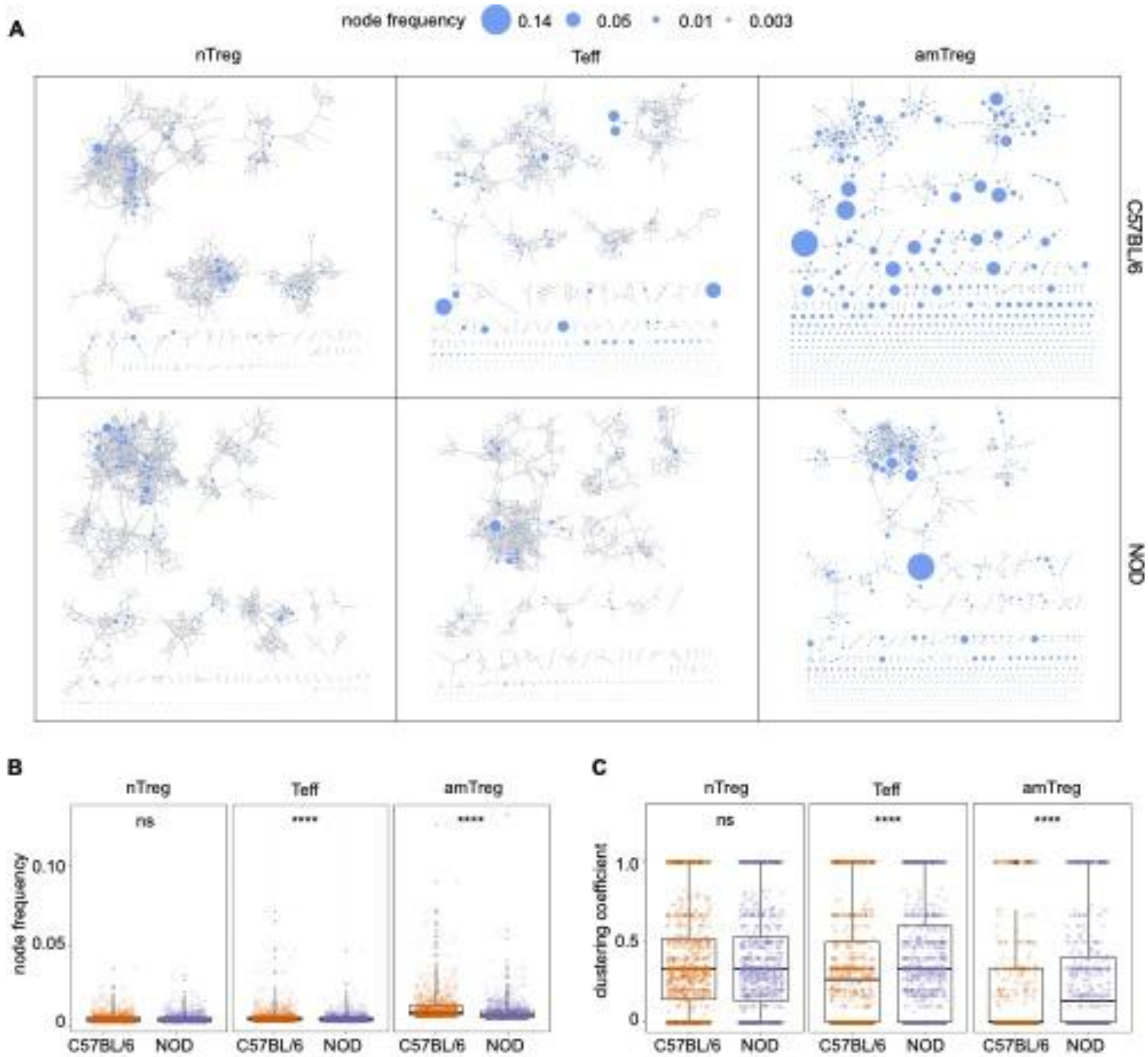

Figure 3. Similarity structure of nTreg, Teff and amTreg CDR3 repertoires. (A) For each cell population, the most frequent 1000 CDR3s across all mice were identified based on their mean frequency and used to construct networks based on the Levenshtein distance: CDR3 sequences (nodes) with one amino acid difference (insertion/deletion/substitution) are connected by an edge. One network was constructed per cell population and mouse background. **(B)** CDR3 frequency comparison between B6 and NOD mice for each cell subset. **(C)** Comparison of the-clustering coefficient, which quantifies the degree to which neighbor nodes tend to cluster together, between B6 and

NOD mice in each cell subset. CDR3s with no connections are excluded. **(B, C)** Each dot represents a CDR3 and boxplots represent the median, first and third quartiles. ns P > 0.05, *** P ≤ 0.001, **** P ≤ 0.0001, Wilcoxon test with Holm's correction.

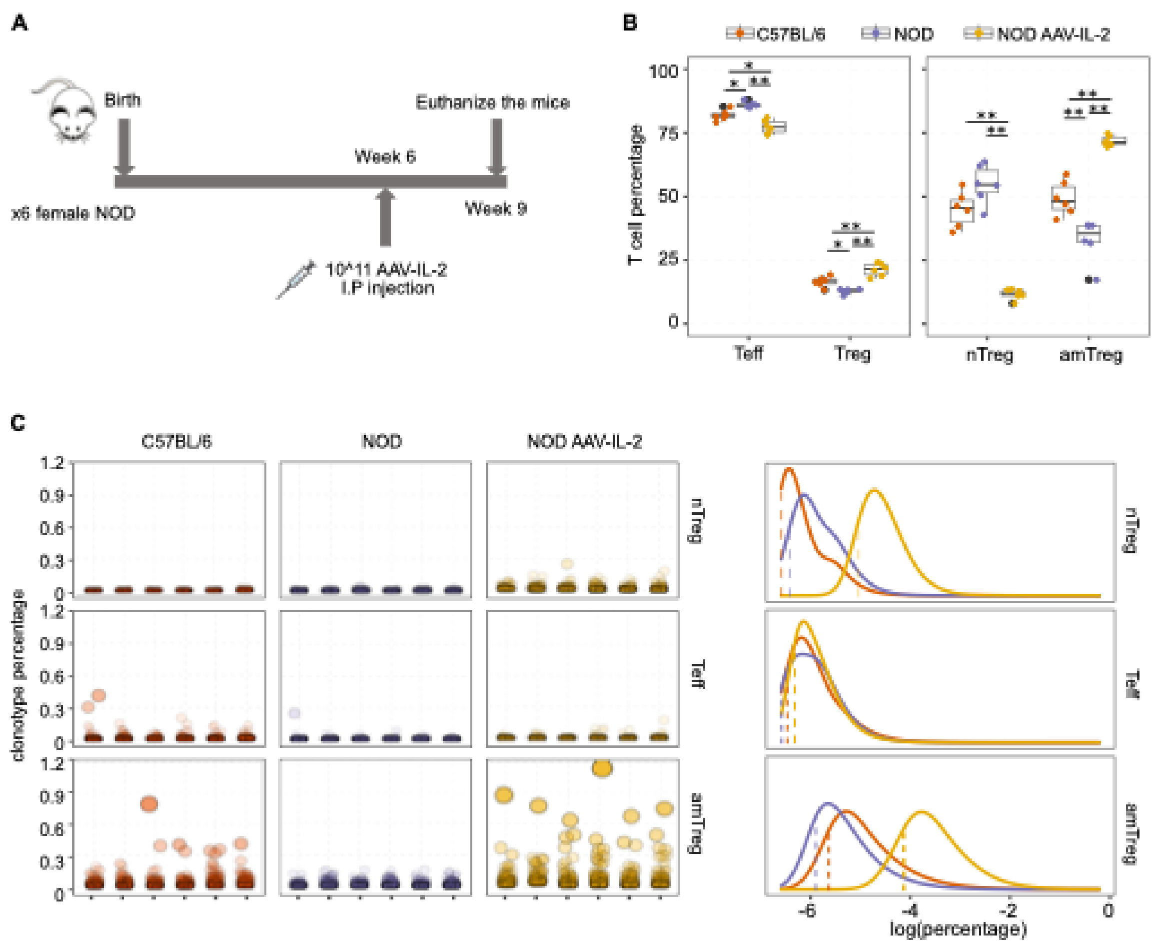

Figure 4. IL-2 induces Treg clonotype expansions. (A) Experimental scheme. **(B)** Box plots show the percentage of T cells in each B6 (orange), NOD (purple) and IL-2–treated NOD (yellow) mice (n=6 per background). **(C, left plot)** The abundance (in %) of expanded clonotypes for each sample (each column represents a single sample) within the three populations and mouse backgrounds. Expanded clonotypes were identified as the ones with counts greater than Q3+1.5xIQR according to a fitted negative binomial

distribution. **(C, right plot)** Probability density plot based on the mean expanded clonotype distribution between the three mouse conditions for each cell subset. Colored dashed lines represent the mode of each curve.

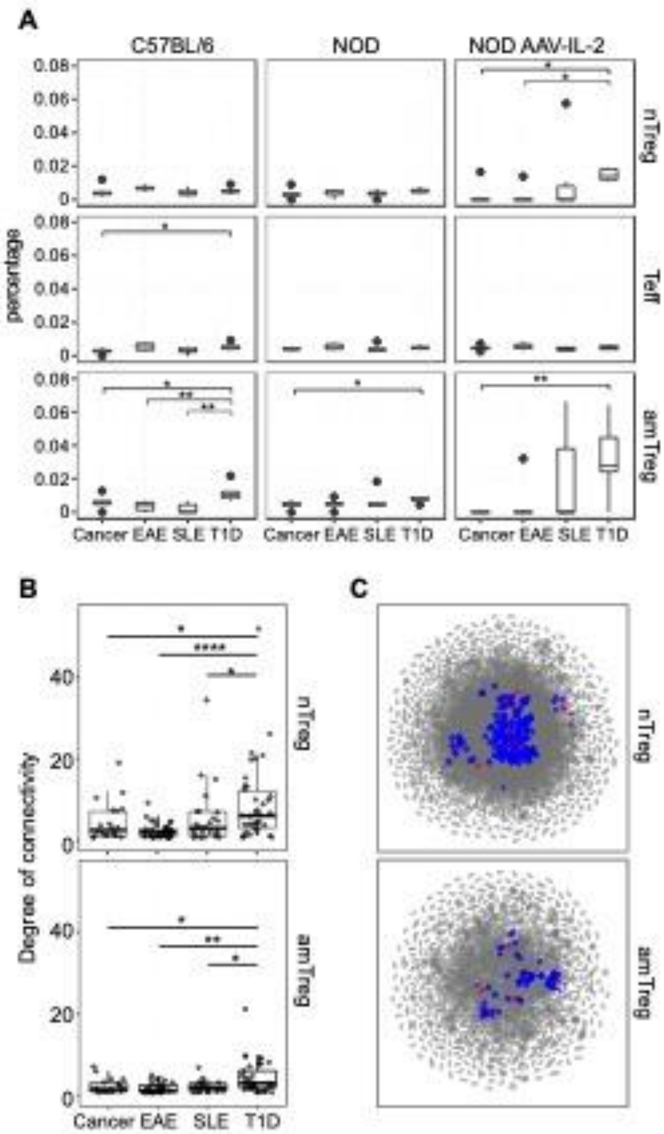

Figure 5. IL-2 preferentially expands clonotypes with a diabetes-related specificity. **(A)** The proportion of CDR3s specific for cancer, EAE, SLE and diabetes found within the expanded clonotypes of each population (n=6). Box-and-whisker plots show the median and outliers. **(B)** CDR3s with a disease-related specificity were connected by

LD=1 to the pool of amTreg and nTreg CDR3s identified across the six IL-2 treated NOD mice. The degree of connectivity, i.e the number of LD=1 edges connecting a given node, of disease-specific CDR3s is plotted (n=52 in cancer, n=50 in EAE, n=49 in SLE and n=52 in T1D). **(A, B)** Wilcoxon test was performed with the T1D group as a reference: ns $p > 0.05$, * $P \leq 0.05$, ** $P \leq 0.01$ and *** $P \leq 0.001$. **(C)** Graphical representation of the LD=1 networks in NOD AAV-IL-2 treated nTregs and amTregs. Unconnected sequences are not shown. Identical sequences found in the database and samples from IL-2 treated mice are shown in red, their neighbors in blue as well as diabetes-related sequences from the database.