

Expression of an interleukin-2 partial agonist enhances regulatory T cell persistence and efficacy in mouse autoimmune models

Corresponding Author: Professor David Klatzmann

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This manuscript by Robert et al. reports on experiments to increase the efficiency of adoptive regulatory Foxp3+ T (Treg) cell therapy and convincingly demonstrates that ectopic expression of an IL-2 partial agonist is a viable approach to achieve an impressive increase in Treg cell efficacy in vivo. These results seem particularly relevant considering that adoptive Treg cell therapy has long been considered one of the most innovative approaches to interfere or even prevent undesired immune responses (e.g. autoimmune diseases, transplanted organ rejection, etc.). However, despite considerable efforts in basic and clinical research, including numerous encouraging results in animal models of human disease, the clinical translation of Treg cell-based therapies is still far from clinical practice, with GMP production of sufficiently large numbers of Treg cells and the relatively rapid loss of Treg cells in adoptively transferred patients being the major obstacles. This is probably best illustrated by the notion that Treg cell transfer (e.g. during HSC transplantation to prevent graft-versus-host disease) can significantly improve clinical outcome, but the beneficial effect was rather transient (several months), while other recent clinical trials did not show evidence of clinical improvement (e.g., Treg cell administration at onset of type 1 diabetes).

In initial experiments, the authors focused on recapitulating the rapid loss of Treg cells after injection into recipient mice, which could only be moderately improved by concomitant administration of soluble IL-2. In a series of elegantly designed experiments using both murine and human Treg cells (in conjunction with single cell RNAseq and sophisticated disease models, e.g. xenogeneic graft-versus-host disease following transfer of human T cells into NSG mice), the authors were then able to conclusively demonstrate that the specific IL-2 partial agonist they selected for their studies can overcome the current hurdles in Treg cell-based therapies described above when endogenously expressed by the adoptively transferred Treg cell population. This included improved proliferative expansion of Treg cells in vitro prior to injection, stability of a Foxp3+ phenotype in vivo (TSDR demethylation, see below), and most importantly, long-term maintenance of Treg cells expressing the IL-2 partial agonist in vivo. Overall, in my opinion, the present manuscript represents a significant advancement in the context of Treg cell-based therapy with high clinical relevance, the publication of which I enthusiastically support. Minor comments:

1. In line 42 of the Abstract, the authors use the term "suicide gene" which seems somewhat misleading given that they are referring to the expression of a congenic surface marker which is then targeted by depleting antibodies for Treg cell ablation in vivo.
2. It appears that no information on the IL-2 partial agonists used in the present study is provided, such as the nucleotide and amino acid sequence - please clarify.
3. The authors demonstrate that genetically engineered Treg cells produce an excess IL-2 partial agonist that can promote endogenous Treg cells in a bystander manner, and they also show that this does not appear to be the case for NK and CD4 T effector cells – how do the authors explain this apparent discrepancy? Could it be that IL-2 partial agonist Treg cells cause bystander activation of pathogenic CD8+ T effector cells? In my view, this is a particularly important question given that many adverse immune responses are mediated by the CD8 T effector compartment.
4. Lines 238 - 239: please check "... disease control, with a some degree of splenomegaly....".
5. In the Discussion section, the authors state that they "never observed a loss of Foxp3". Although this statement addresses an important question, I am not aware of any data in the present manuscript that would directly answer this question. Besides efficient TSDR demethylation in murine Treg cells with IL-2 partial agonist (shown in Figure 3G; an experimental

setting in which none of the Treg cell populations shows loss of stability), can the authors provide experimental evidence that expression of IL-2 partial agonist increases the stability of Treg cells at the level of Foxp3 protein expression? How does partial IL-2 agonist expression affect TSDR demethylation and Foxp3 protein expression in human Foxp3+ Treg cells?

6. Given the robust evidence in the literature that Treg cells of thymic and peripheral developmental origin exert specialized effector functions, which likely has implications for the rational design of Treg cell-based therapies, it seems important to clarify whether the IL-2 partial agonist used in the present study exerts similar effects on both developmental subsets.

7. It would be particularly helpful if the authors could comment on, and provide experimental evidence for, the mechanisms by which the IL-2 partial agonist enhances Treg cell efficacy, ranging from promotion of proliferative expansion to impairment of apoptotic cell death.

Reviewer #2

(Remarks to the Author)

This manuscript by Robert et al., entitled "Enhancing regulatory T cell efficacy through self-sufficiency for an interleukin 2 partial agonist" aims to improve the long-term engraftment of adoptively transferred T regulatory cell by engineering Tregs that produce their own source of IL-2. Its well established that infused Tregs have a limited lifespan in vivo and the IL2/IL2R interaction is essential for the home stasis of thymically derived Tregs. Currently protocols for the clinical setting depend on ex vivo expansion in the presence of high concentration of IL-2 and TCR signaling, possibly making these expanded highly depending on high concentration of IL-2 to survive in vivo. The study was undertaken to test the hypothesis that transducing Tregs with an IL-2 gene might overcome some the limitations associated with adoptive Treg therapy. The authors investigate 2 different IL-2 approaches. One approach utilized the wild-type IL-2 gene that shows limited efficacy and survivability. The second approach utilized a partial agonist IL-2 which showed much more efficacy. Although this is an important area of research and the field can greatly benefit from improved strategies for Treg-based therapies, there are several weaknesses that dampen enthusiasm. Many figures lack statistical analysis, there is a low number of mice examined, and there is limited analysis of the tissues making interpretation of the results very difficult to determine the strength of the data. Please see specific comments below.

Major Comments:

- Figure 1 lacks quantitative analysis of the results, including statistical analysis and is not adding much value to the overall conclusions. Much of this data is gleaned from the results presented in Figure 2. This figure should be moved to the supplemental data and graphs and statistical analysis included. Moreover, the reader is left pondering why these in vitro expanded Tregs isolated from Thy1.1 B6 mice are being transduced on day 5 of culture with Thy1.2 reporter gene given that Thy1.1 serves as the congenic marker to track the engraftment and persistence of the infused Tregs into Thy1.2 B6 recipients. Rationale should be provided that this strategy serves as a control for the later experiments presented in Figure 2 where IL-2wt and IL-2pa are transduced along with the Thy1.2 reporter for readout of transduction efficiency. This figure could be strengthened by including data on the FoxP3 and CD25 expression on the donor and endogenous Tregs, including the MFI. Furthermore, how does engraftment and persistence of these 9×10^5 control-Tregs compare to in vitro expanded cells that were not transduced on day 5? Is this strategy significantly impairing the in vitro expanded Tregs? This is an important comparison and would strengthen the data along with more rigorous quantitative analysis of the results and increase in the number of mice examined as Figure 1D has only $n=2$.
- Figure 2 is missing statistical analysis. It seems to appear based on the figure legend that " $n=2$ per group" was performed. However, there are several statements in the figure legend that are confusing. For example, data in (B) is represented of 2 animals per group as stated in the figure legend. However, the figure legend also states "Data were extracted from one of 2 independent experiments conducted." This is confusing, was it 2 animals per group in 2 independent experiments, so 12 mice in total per group? Why only show results from 1 experiment and not all 2? Data in panel L also states "data were extracted from one of three independent experiments". Why only show results from 1 out of 3? Moreover, several timepoints are given for the analysis of the blood, but there is only limited analysis of the tissues with only a single timepoint examined, day 14 post-Treg infusion. Without any statistical analysis, the low number of mice examined, and the limited analysis of the tissues, it is very difficult to determine the strength of the data.
- In Figure 3, how many mice were examined? Figure 3D should include the MFI on the histograms and a graph showing all the MFI from all the mice examined and statistical analysis performed on the 4 T-cell populations examined.
- Again, in Figure 4 many graphs lack statistical analysis, only panels E and H have statistics.
- In figure 5, examination of expansion and transduction was performed from sorted human PBMCs CD4+CD25+ CD127-low from $n=2$ per group in panel B-D. What was the transduction efficiency in these human cells? Statistics performed in panel B-D? In panel G, data is small due to the inclusion of IL-2 in the upper right corner. An insert that zooms in on the lower portion (not include the IL-2 in the upper right) to improve presentation of the data would be helpful.
- Figure 6, Again statistics missing for some of the data. Panels C and D need to include MFI, quantitative analysis through presentation of graphs and statistics of all the data points.
- Although the efficacy data in vivo in figure 4 and 6 shows great promise, there was no discussion on the impact that the IL-2pa-Tregs could have on tumor development or chronic infections. The results shown show high level engraftment and survivability of the donor Tregs cells but also a substantial increase in endogenous Tregs as well. This is a real concern as too many Tregs pose challenges as well and the authors do not even discuss this or a strategy to pursue in the future to

explore the safety of their engineering approach.

Minor Comments

1. Figure 2 is missing labels A, B and C. Please add these labels.

Reviewer #3

(Remarks to the Author)

This is a timely and consequential manuscript, with the potential to greatly affect the clinical applicability of tTreg adoptive cell therapy, both for autoimmune disease and hematopoietic stem cell transplantation. There is one major point and a few minor points that need to be addressed.

Major point:

- Figure 6E: The clinical applicability of Tregs-IL2pa is paramount for this manuscript. While it is understood that using a sub-optimal ratio of Treg is an approach to help demonstrate increased XGVHD efficacy for Tregs-IL2pa, only a single assay is shown, and control Treg do not even show a hint of efficacy. The authors also do not actually list the p-value (and statistical method) for Tregs-IL2pa. This reviewer would like to see the xenoGVHD assay repeated. Important factors would be a significant amelioration of disease with Ctrl Tregs, and enhanced efficacy with Tregs-IL2pa. The weight figure was very effective and should be repeated. Many investigators have also assessed Treg survival by monitoring Treg persistence in blood (similar to murine assays these authors used). With their previous data, it seems very likely that Tregs-IL2pa will have increased persistence in the blood.

Minor points:

- I do not think the designation the authors have chosen for their cell product (e.Tregs) is appropriate since there are already >89 papers that use eTreg as an abbreviation for effector regulatory T cells..
- Figure 1: All the data presented appear to be from a single purification/expansion or from single animals receiving adoptive Treg transfer. Summary data should be presented for each panel (A-D) to show the relative variability of the initial set-up.
- Figure 2 (summary data is well represented here): Because there is a time-dependent change in both endogenous and Tregs-IL2pa, it is necessary to assess effects via cell# instead of % (of CD4+ or endogenous Tregs). This will also help to show whether there is an overall expansion in endogenous CD4+ conv. T cells as well. Finally, Figures 2L-M seem to be only n=1. Is this correct?
- While perhaps a bit of semantics, I question the authors statement that IL2pa increases the 'survival potential' of Tregs. Specifically, it appears that Tregs-IL2pa expand in number even over the first 14 days. Fig. 2H shows that when only 9×10^5 Tregs-IL2pa are injected on day 0, $>1 \times 10^6$ Tregs-IL2pa were present in just the spleen.
- Figure 3E: The authors anticipated questions on TCR clonality in Tregs-IL2pa, and I applaud the stringent conditions they chose to address the issue with (following 6 months of expansion in vivo). These data showed that Tregs-IL2pa had much lower diversity than any of the controls (control naïve and activated memory Tregs, or endogenous Treg present in Tregs-IL2pa injected animals). The authors statement in both the abstract and discussion is that Tregs-IL2pa have a diverse TCR repertoire. While technically true, I think this is somewhat misleading. The authors suggest that Tregs-IL2pa had the lowest diversity because 'they originate from a necessarily less diverse repertoire of cultured and expanded cells', but I don't think CD3/28 polyclonal stimulation has been shown to effect clonality. Also, Tregs-IL2pa had lower diversity than the endogenous Tregs with which they co-expanded in vivo.
- Figure 5G/H (and I): no indication is given as to whether the control Treg were given exogenous IL-2. This is somewhat confusing, especially given that the preceding figures (5B-F) looked at both possibilities. If exogenous IL-2 was included, please expressly denote. If not, I would question the comparison between an actively expanding population (Tregs-IL-2pa) and a dying population (Ctrl Tregs) (from 5C).
- Points that need to be addressed in discussion:
 - o In the NSG setting, did Tregs-IL-2pa result in increased expansion/persistence of co-transferred non-transduced cells (GFP- cells shown in 6D)?
 - o The authors document that both endogenous Tregs and Tregs-IL2pa significantly increase over time, to supra-physiological levels. While they discuss their own publication assessing immune suppression with low-dose IL-2 in NOD mice, they should also discuss publications showing that high Treg numbers can compromise immune control in mice (The Journal of Immunology, 2011, 186: 4862–4871). Generalized immunosuppression is a very important consideration in human adoptive tTreg therapy and needs to be discussed.

Reviewer #4

(Remarks to the Author)

In this manuscript, Robert et al provide intriguing results on the genetic modification of regulatory T cells (Treg) to confer what they call "self-sufficiency" for IL-2 signaling and improve on current limitations (namely survival, expansion, and stability) incurred by Treg-based therapies for different immunological disease. The key novelty is the identification of the need to use a gene encoding a partial agonist form of IL-2 (IL-2pa) to enhance the properties of adoptively transferred cells. It is noteworthy that the study includes both murine and human Treg investigations, providing important proof that the proposed strategy has translational potential and is clinically relevant. Overall, the results presented are promising and could have a meaningful impact on the design of more effective Treg therapies. Unfortunately, there are also multiple major points that are not addressed by the authors, reducing significantly the enthusiasm for the manuscript in the current form.

Major Points:

1. The most important weakness of the manuscript is the absence of information on the partial agonist of IL-2 used in this work. The authors do not provide a rationale of how they arrived at the version implemented and this sequence is not provided (and others tested). Moreover, it is not clear if this IL-2pa sequence is the same or not between mouse and human Treg studies.
2. It is remarkable that the transferred Treg(IL-2pa) keep expanding in the recipient and, even more impressive, that they support the expansion of endogenous Treg to the point that they constitute more than 60% of total CD4 T cells. As pointed out by the Authors themselves, this progressive expansion could become an issue of global immunosuppression. The authors should test this concern (they have published a previous paper using established models). The simple mention in the Discussion that their previous study (using a viral vector for sustained IL-2 production in vivo) did not indicate there should be a concern is not a valid argument for the conditions of this very different model system.
3. The in vivo model of Treg(IL-2pa) suppression, based on the Foxp3-DTR mouse model, is not really a “therapeutic model”. The results clearly show that Treg(IL-2pa), injected 20 days before DT treatment, can replace endogenous Treg and prevent the inflammation induced by the depletion induced by DT. However, it does not prove that Treg(IL-2pa) are therapeutic. They should be tested in a more clinically relevant model (as done in the past by the group).
4. In many parts of this manuscript figures, there is no cumulative data presented (only representative) limiting our ability to draw robust conclusions (e.g. Figure 1B-D, Figure 2L-M, Figure 3D). As importantly, in many instances there is a lack of error bars and statistical analysis. The most prominent example is Figure 6E. This is probably the piece of data that provides the most important indication of the therapeutic potential of human Treg (IL-2pa). Unfortunately, there is no statistical analysis presented and we have to believe there is no statistically significant difference between groups.
5. The scRNAseq analysis of murine Treg groups is potentially very interesting, but the results presented only focus on showing a similarity between Treg(IL-2pa) and endog Treg. The UMAP shows that these two populations also have significant elements of distinction, and it would be really useful to know where these differences reside. In similar fashion, the scRNAseq of human Treg(IL-2pa) is presented to indicate an increase in the IL-2 signaling pathway and in glycolysis. What is the meaning of it? And could the authors draw any parallelism between the mouse and human analysis? Finally, could these dataset shed any light on why Treg with wt IL-2 disappear suddenly? The authors imply toxicity but no data is showed to support this interpretation.

Minor points:

- 1) Figure 1, panels B and C would be even more informative if, in addition to providing cumulative data (see above), would show the ratio between transduced and non-transduced injected cells. It would be nice to see Foxp3 expression as well.
- 2) Figure 2 is missing A, B, C labels
- 3) Figure 3, organize the markers in panel B and D in the same order to improve readability. Additionally, could the authors better explain in the text what the rationale is to wait 6 month post injection to analyze Treg?
- 4) Figure 4, after panel A, there is no inclusion of ctrl Treg in all other panels.
- 5) Figure 5G, it is hard to read the volcano plot, can IL-2 be excluded to better show the changes in gene expression?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In my view, the revised version of the manuscript addresses most of the reviewer's concerns and would be suitable for publication. However, one significant issue remains unresolved, as the revised manuscript still lacks sufficient information about the compound used. After careful consideration, I am willing to support the publication of the revised manuscript on the condition that the authors provide the relevant information about the reagents or make the reagents available to the scientific community under specific terms (such as an MTA, NDA, or formal collaboration), and that Nature Communications proceed with retraction of the article should the authors fail to comply.

Reviewer #2

(Remarks to the Author)

This revised manuscript by Robert et al., entitled “Enhancing regulatory T cell efficacy through self-sufficiency for an interleukin 2 partial agonist” aims to improve the long-term engraftment of adoptively transferred T regulatory cell by engineering Tregs that produce their own source of IL-2. Overall, the authors have addressed many of the reviewers' comments and added to the discussion. A major comment and some minor comments are below.

Major comments:

- 1) Still have concerns about the safety of their Treg product due to the high-level engraftment and survivability of the donor Tregs cells but also a substantial increase in endogenous Tregs as well. This is a real concern as too many Tregs pose challenges as well and the authors do not even discuss this or a strategy to pursue in the future to explore the safety of their engineering approach. Although this group has previous data from other models that a substantial increase does not impair immune responses or increase susceptibility to tumors, future studies should be performed to specially address this potential concern and nor so easily dismissed because of their previous work with a different approach that resulted in an increase and sustained Treg population.

Minor comments:

- 1) Page 7 line 154: It states figure 1B but believe it should be Figure 2B. Please clarify.
- 2) Page 7, lines 155-157. Should add that this is data from Figure 2A, the observed 10% at later timepoints for Tregs-IL2pa Recipients. However, Figure 2B shows that day 14 and day 36 the %Thy1.1 among the CD4 T cells averages around 2% (with a couple of mice higher around 5%) which differs from the Figure 2A right panel significant which shows around 5% with no error bars indicating that the data is consistent. What is the difference between panel A and B. Authors state in their rebuttal that day 14 and day 36 were consistent time points over multiple experiments (Panel B, n=5-11 mice) and it seems that Panel A is only 2 mice per group. In panel B, which data points are the 2 mice per group from panel A? Thinking in the Tregs-IL2pa group that they are the 2 highest points in panel B, then where are the error bars in panel A. Please clarify,
- 3) Please clarify in the figure legend and text for Figure 4A that the in vitro Treg suppressor assay is using day 3 in vitro transduced Tregs and not Thy1.1+ transduced Tregs isolated from adoptively transferred mice. It is possible that the Tregs-IL2pa isolated from adoptively transferred mice could have higher suppressive capacity compared to the Tregs-CTRL if isolated early timepoint on a per cell basis. Although the in vivo experiment shows that the Tregs-IL2pa have suppressor function, it would also be interesting to know whether the Tregs-IL2pa or even the stimulated endogenous Tregs are more robust suppressor compared to Tregs from unmanipulated recipients.

Reviewer #3

(Remarks to the Author)

Thank you for addressing the comments point by point.

Reviewer #4

(Remarks to the Author)

In this revised version of the manuscript, Robert et al provide some additional data in response to the concerns raised by the reviewers. Here are comments specific to the points raised by this reviewer.

Major Points:

1. The authors indicate that the modifications introduced in the sequence of human IL-2 to generate a partial agonist are proprietary and cannot be disclosed. This is unfortunate but understandable. However, they do not provide enough information for other investigators to "identify their own suitable IL-2pa". Additional information that does not undermine the patentability of their invention could be provided to support the reproducibility of their work.
2. Thank you for adding to the Discussion a few points outlining the probable differences between the results presented in the murine model and the behavior that human Treg might have clinically. Such a point is important to define potential limitations of the work that have not been addressed yet as well as potential risks that will need to be accounted for.
3. In relation to the in vivo model of mouse Treg(IL-2pa) suppression, if the authors don't want to add any more convincing experimental model of therapeutic efficacy, they should change the definition from "therapeutic" to "prophylactic"
4. Thank you for addressing the issue of lack of cumulative data and statistics. The data can be better interpreted now.
5. The new figure (panels A, B, and C) presented in the rebuttal should be included as Supplementary Data in the manuscript.

Minor Points

All properly addressed.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

"Enhancing regulatory T cell efficacy through self-sufficiency for an interleukin-2 partial agonist" - Nature Communications

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

This manuscript by Robert et al. reports on experiments to increase the efficiency of adoptive regulatory Foxp3+ T (Treg) cell therapy and convincingly demonstrates that ectopic expression of an IL-2 partial agonist is a viable approach to achieve an impressive increase in Treg cell efficacy in vivo. These results seem particularly relevant considering that adoptive Treg cell therapy has long been considered one of the most innovative approaches to interfere or even prevent undesired immune responses (e.g. autoimmune diseases, transplanted organ rejection, etc.). However, despite considerable efforts in basic and clinical research, including numerous encouraging results in animal models of human disease, the clinical translation of Treg cell-based therapies is still far from clinical practice, with GMP production of sufficiently large numbers of Treg cells and the relatively rapid loss of Treg cells in adoptively transferred patients being the major obstacles. This is probably best illustrated by the notion that Treg cell transfer (e.g. during HSC transplantation to prevent graft-versus-host disease) can significantly improve clinical outcome, but the beneficial effect was rather transient (several months), while other recent clinical trials did not show evidence of clinical improvement (e.g., Treg cell administration at onset of type 1 diabetes).

In initial experiments, the authors focused on recapitulating the rapid loss of Treg cells after injection into recipient mice, which could only be moderately improved by concomitant administration of soluble IL-2. In a series of elegantly designed experiments using both murine and human Treg cells (in conjunction with single cell RNAseq and sophisticated disease models, e.g. xenogeneic graft-versus-host disease following transfer of human T cells into NSG mice), the authors were then able to conclusively demonstrate that the specific IL-2 partial agonist they selected for their studies can overcome the current hurdles in Treg cell-based therapies described above when endogenously expressed by the adoptively transferred Treg cell population. This included improved proliferative expansion of Treg cells in vitro prior to injection, stability of a Foxp3+ phenotype in vivo (TSDR demethylation, see below), and most importantly, long-term maintenance of Treg cells expressing the IL-2 partial agonist in vivo. Overall, in my opinion, the present manuscript represents a significant advancement in the context of Treg cell-based therapy with high clinical relevance, the publication of which I enthusiastically support.

→ We thank the reviewer for these encouraging comments.

Minor comments:

1. In line 42 of the Abstract, the authors use the term "suicide gene" which seems somewhat misleading given that they are referring to the expression of a congenic surface marker which is then targeted by depleting antibodies for Treg cell ablation in vivo.

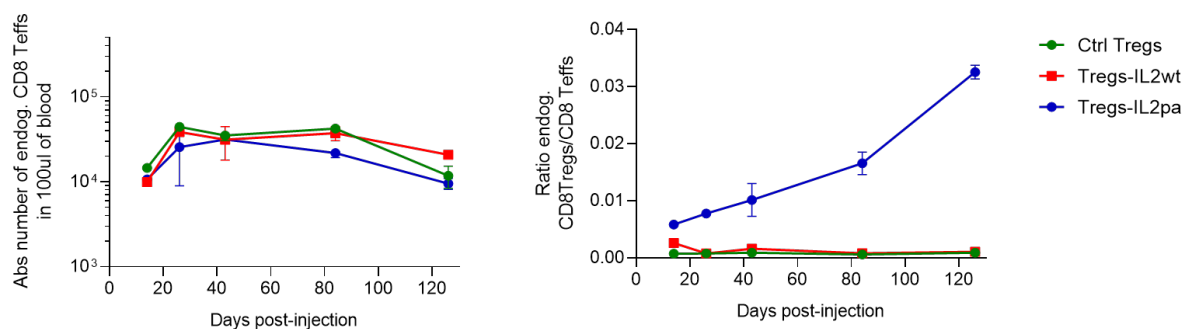
→ The Thy1.1 gene is indeed used as a surrogate for a "suicide gene" or a "kill switch" as a term more often used for a protein expression that allows cell depletion with an antibody. We have therefore replaced the term suicide gene with "kill switch for conditional cell killing" in the abstract.

2. It appears that no information on the IL-2 partial agonists used in the present study is provided, such as the nucleotide and amino acid sequence - please clarify.

➔ The specific sequence we selected is proprietary. In order to give this strategy a chance to be developed in the clinic, we cannot disclose it at this time. However, we would like to emphasize that (i) we are providing enough mechanistic information to understand our results and (ii) we are disclosing enough information for others to identify their own suitable IL-2pa for the purpose of making efficient IL-2 self-sufficient Tregs. In fact, we are fully disclosing the only important parameter, which is the level of loss of biological activity of IL-2pa (approximately 1000-fold based on pSTAT5).

3. The authors demonstrate that genetically engineered Treg cells produce an excess IL-2 partial agonist that can promote endogenous Treg cells in a bystander manner, and they also show that this does not appear to be the case for NK and CD4 T effector cells – how do the authors explain this apparent discrepancy? Could it be that IL-2 partial agonist Treg cells cause bystander activation of pathogenic CD8+ T effector cells? In my view, this is a particularly important question given that many adverse immune responses are mediated by the CD8 T effector compartment.

➔ We previously showed that the IL-2R downstream activation of genes after IL-2 binding requires a 100-fold lower IL-2 concentration for Tregs vs Teff (Yu et al., Diabetes, 2015). Thus, a partial agonist of IL-2 is not sufficient to activate and expand these cells. This is also the case for CD8 T effector cells, as shown in the figures below. These figures have been generated with the same mice as in Figures 2A and 2C to G.

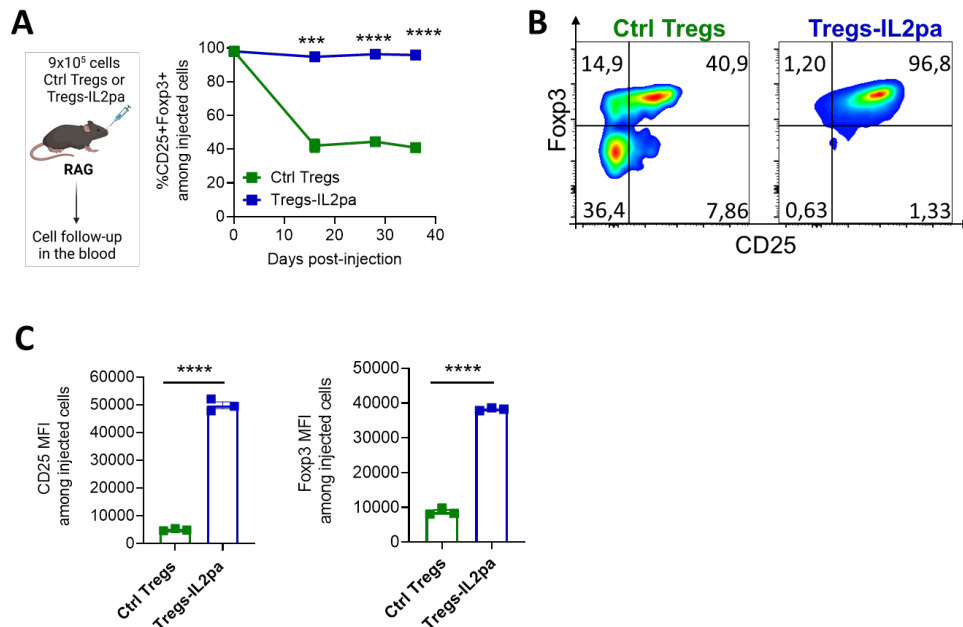


4. Lines 238 - 239: please check "... disease control, with a some degree of splenomegaly....".

➔ The mistake has been corrected.

5. In the Discussion section, the authors state that they "never observed a loss of Foxp3". Although this statement addresses an important question, I am not aware of any data in the present manuscript that would directly answer this question. Besides efficient TSDR demethylation in murine Treg cells with IL-2 partial agonist (shown in Figure 3G; an experimental setting in which none of the Treg cell populations shows loss of stability), can the authors provide experimental evidence that expression of IL-2 partial agonist increases the stability of Treg cells at the level of Foxp3 protein expression?

➔ We have injected several Tregs-IL2pa products into numerous mice, including in RAG2KO mice which are largely deprived of IL-2, and looked at the persistence of the Treg phenotype over time, sometimes more than a year after adoptive transfer. In all these experiments, the injected Tregs-IL2pa maintained the expression of Foxp3 and CD25. In contrast, control Tregs injected into RAG2KO mice progressively lost Foxp3 and CD25 expression as shown in the figures below.



*Legend: Tregs were isolated from the spleen and lymph nodes of B6 Thy1.1 mice. On day 5 of culture, the cells were transduced with the control vector MSCV-Thy1.2 for the expression of the reporter gene mouse Thy1.2 alone in the transduced cells (Ctrl Tregs) or the vector MSCV-IL2pa-Thy1.2 for the expression of the partial agonist form of IL-2 and Thy1.2 (Tregs-IL2pa). Two days after transduction, 9x10⁵ unsorted transduced cells per mouse were injected into RAG2KO mice. Injected cell survival was tracked in mouse blood (n=3 per group). **A.** Percentage of Foxp3+CD25+ cells among the injected cells. **B.** Representative pseudocolor plot of CD25 and Foxp3 expression among the injected cells. Data were obtained from staining at day 36 post-infusion. **C.** Mean Fluorescence Intensity (MFI) of CD25 and Foxp3 in the injected cells on day 36.*

We have now added these figures to the paper as new panel 3H and Supp. Figure 3C.

How does partial IL-2 agonist expression affect TSDR demethylation and Foxp3 protein expression in human Foxp3+ Treg cells?

➔ TSDR methylation has been directly analyzed only in murine cells as most of the article is focused on the murine Tregs. However, as shown in the Figure 5F and Figure 6C, Foxp3 expression is maintained and actually higher in human Tregs-IL2pa than in Ctrl Treg, in vitro and in vivo, respectively. This indirectly shows that the demethylation is maintained.

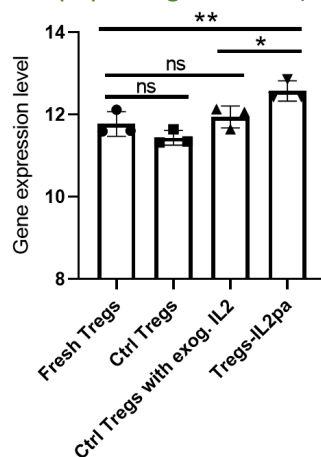
6. Given the robust evidence in the literature that Treg cells of thymic and peripheral developmental origin exert specialized effector functions, which likely has implications for the rational design of Treg

cell-based therapies, it seems important to clarify whether the IL-2 partial agonist used in the present study exerts similar effects on both developmental subsets.

- ➔ The conversion of Teff to pTregs in vivo occurs in inflammatory environments. First, as the Tregs-IL2pa are derived from highly purified Tregs from normal mice, there are very few contaminating effector T cells that could convert to pTregs. Second, the pharmacodynamic of Tregs is monitored in the absence of inflammation. For these two reasons, the expanded Treg-IL2pa must represent the expansion of tTregs.
- ➔ Additionally, in our deep phenotyping analysis by flow cytometry (Figure 2D), Tregs-IL2pa, the IL-2pa-stimulated endogenous Tregs, and control fresh Tregs from unmanipulated mice were all found to be Helios⁺, a marker consistent with the known phenotype of tTregs.
- ➔ However, to directly address the reviewer's comment, we conducted an experiment to assess the development of pTregs in B6 Thy1.2 mice. In this experiment, we co-injected 4.5×10^5 sorted transduced Thy1.1⁺Thy1.2⁺Foxp3-GFP⁺ Tregs-IL2pa with 2×10^6 fresh Thy1.1 CD4⁺Foxp3⁻ Tconv cells originating from the same mouse strain, and thus harboring a non-expressed GFP gene. The conversion of such cells to pTregs would have led to a population of Foxp3-GFP⁺Thy1.1⁺Thy1.2⁻ in the blood. We did not observe such cells, indicating that the generation of pTregs is at most too low to be detected.

7. It would be particularly helpful if the authors could comment on, and provide experimental evidence for, the mechanisms by which the IL-2 partial agonist enhances Treg cell efficacy, ranging from promotion of proliferative expansion to impairment of apoptotic cell death.

- ➔ We do not explicitly state that each Treg expressing IL2pa has a better efficacy than a regular Treg. In classical in vitro suppression experiments, which are known to only partially assess the Treg suppressive activity and in fact mostly the deprivation of IL-2 from Teff, Tregs-IL2pa perform as well as Ctrl Tregs (Figure 4A with murine cells and figure 5I with human cells).
- ➔ However, they have a high expression of multiple markers linked to Treg suppressive function such as CD25, CTLA-4, GITR, LAP, CD39... and as such could have on an individual basis a better suppressive activity. As a population, their superior activity is mostly due to their expansion leading to a much higher Treg/Teff ratio.
- ➔ The fact that Treg-IL2pa survive so well indicates that they are less prone to undergo apoptosis. In line with this, Tregs-IL2pa have a high expression of ICOS (a molecule known to promote T-cell survival by reducing apoptotic signaling) and have the highest expression of the anti-apoptotic gene BCL2 (figure below).



We thank the reviewer for his/her constructive comments that helped us improve our manuscript.

Reviewer #2 (Remarks to the Author):

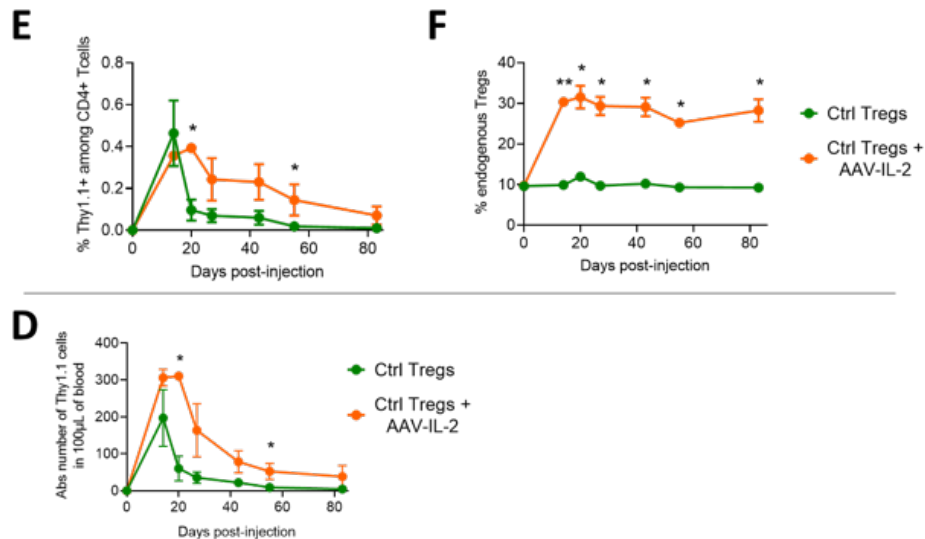
This manuscript by Robert et al., entitled “Enhancing regulatory T cell efficacy through self-sufficiency for an interleukin 2 partial agonist” aims to improve the long-term engraftment of adoptively transferred T regulatory cell by engineering Tregs that produce their own source of IL-2. Its well established that infused Tregs have a limited lifespan in vivo and the IL2/IL2R interaction is essential for the home stasis of thymically derived Tregs. Currently protocols for the clinical setting depend on ex vivo expansion in the presence of high concentration of IL-2 and TCR signaling, possibly making these expanded highly depending on high concentration of IL-2 to survive in vivo. The study was undertaken to test the hypothesis that transducing Tregs with an IL-2 gene might overcome some the limitations associated with adoptive Treg therapy. The authors investigate 2 different IL-2 approaches. One approach utilized the wild-type IL-2 gene that shows limited efficacy and survivability. The second approach utilized a partial agonist IL-2 which showed much more efficacy. Although this is an important area of research and the field can greatly benefit from improved strategies for Treg-based therapies, there are several weaknesses that dampen enthusiasm. Many figures lack statistical analysis, there is a low number of mice examined, and there is limited analysis of the tissues making interpretation of the results very difficult to determine the strength of the data. Please see specific comments below.

- ➔ We thank the reviewer for the constructive comments on the statistical analysis. We apologize for not explicitly stating that most of the reported experiments had been reproduced multiple times, which they were, and for not always providing statistical analyses. We have now corrected both points throughout the manuscript.

Major Comments:

1. Figure 1 lacks quantitative analysis of the results, including statistical analysis and is not adding much value to the overall conclusions. Much of this data is gleaned from the results presented in Figure 2.

- ➔ The purpose of Figure 1A and B is to explain the strategy (1A) and to show how the injected Tregs are distinguished from endogenous Tregs (1B), which we believe should help the reader to better understand the results that follow.
- ➔ Figures 1C and 1D were intended to highlight and illustrate with one representative experiment that Ctrl Tregs have poor survival after administration, which we believe is crucial for the reader to understand the challenge we are addressing. We now report the results of four independent experiments (new figure 1E, green line).
- ➔ New Figure 1E also shows that sustained exogenous IL-2 production provided by mean of an AAV-IL2 injection does not rescue the long-term persistence of Ctrl Tregs. Additionally, we report the absolute number of injected cells under both conditions in new Supp. Figure 1D. These results demonstrate that sustained IL-2 support can only improve the early survival of injected Ctrl Tregs. In contrast, continuous IL-2 production expanded endogenous Tregs, as shown in the new figure 1F.



This figure should be moved to the supplemental data and graphs and statistical analysis included.

- ➔ We believe that Figure 1 helps the reader to understand the problem and, with its current improvements, should remain in the main text. However, if the reviewer insists on moving it to the Supplementary Data, we would follow their recommendation.

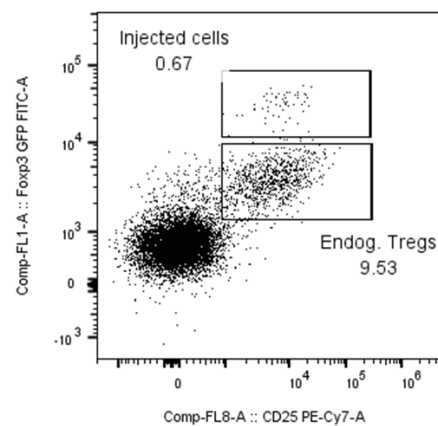
Moreover, the reader is left pondering why these in vitro expanded Tregs isolated from Thy1.1 B6 mice are being transduced on day 5 of culture with Thy1.2 reporter gene given that Thy1.1 serves as the congenic marker to track the engraftment and persistence of the infused Tregs into Thy1.2 B6 recipients. Rationale should be provided that this strategy serves as a control for the later experiments presented in Figure 2 where IL-2wt and IL-2pa are transduced along with the Thy1.2 reporter for readout of transduction efficiency.

- ➔ First, the appropriate control for IL-2 transduced cells are cells transduced with a neutral transgene. Indeed, the exposure to a viral preparation and the procedure itself may have some degree of deleterious effects on cell survival. Thus, the transduction of only a reporter gene allows to address this question.
- ➔ Second, we reasoned that the injection of the mixed population of transduced and untransduced cells coming from the same culture would provide the best possible control to show a potential selective advantage of the IL-2pa expressing cells, which is what we are seeing. Indeed, the cells that we inject come from the same culture and exposed to the same procedures. Since they are a mixture of cell that are transduced and express IL-2 and of others that are not, the latter are the perfect control for the former. By using a Thy1.2 reporter gene, these two populations can be monitored separately, and a potential survival advantage or disadvantage can be detected. This is illustrated in Figure 2C, which shows for cells expressing wild-type IL-2 (i) an initial increase in the proportion of transduced cells, indicating a growth advantage, followed by a sharp decrease in this proportion, indicating toxicity.

This figure could be strengthened by including data on the FoxP3 and CD25 expression on the donor and endogenous Tregs, including the MFI.

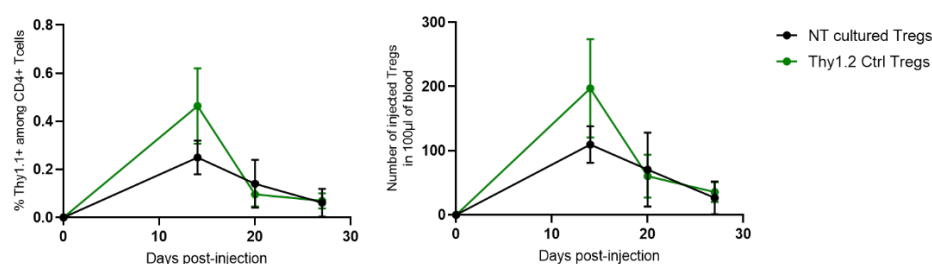
- ➔ We now added in Supp. Figure 1B, the percentage of CD25⁺Foxp3⁺ cells within the injected population, from the day of injection to 14 days post-injection. After day 14, the reduced cell number (20 cells in average in the 8 mice examined on day 36) makes the analysis impossible.

- ➔ The MFI of CD25 and Foxp3 from the four independent experiments performed could not be directly compared due to the fact they have not been analyzed together, leading to variability in MFIs between experiments (notably because of the setting of the flow cytometer). Additionally, for the control Tregs, the MFI analysis is less robust due to the small cell numbers; after day 14 post-injection, fewer than 100 cells are detected. It is also important to note that Foxp3 expression cannot be compared between endogenous and injected Tregs because we use GFP as a surrogate for Foxp3 and the knock-in of GFP in the two mouse strains leads to different levels of Foxp3/GFP expression. This is illustrated in the Figure below in which the injected and endogenous Tregs have similar CD25 expression while having distinct GFP/Foxp3 expression levels.



Furthermore, how does engraftment and persistence of these 9×10^5 control-Tregs compare to in vitro expanded cells that were not transduced on day 5? Is this strategy significantly impairing the in vitro expanded Tregs? This is an important comparison and would strengthen the data along with more rigorous quantitative analysis of the results

- ➔ As stated above, we believe that cells transduced with a neutral gene, such as Thy1.2, serve as the best control. As shown in Supp. Figure 1A and Figure 2C, the percentage of transduced cells in the Ctrl Tregs remains stable, indicating that the overall limited survival affects both transduced and non-transduced cells equally.
- ➔ To directly address the reviewer's comment, we evaluated the survival of cultured non-transduced Tregs ($n=2$) and compared it to cultured Thy1.2-transduced Ctrl Tregs ($n=8$). Our findings, depicted in the figures below, demonstrate that both the non-transduced cultured Tregs (shown in black) and the Thy1.2 transduced Ctrl Tregs (shown in green) exhibit comparably poor survival, as indicated by their proportions and absolute numbers in the blood of plain B6 mice.

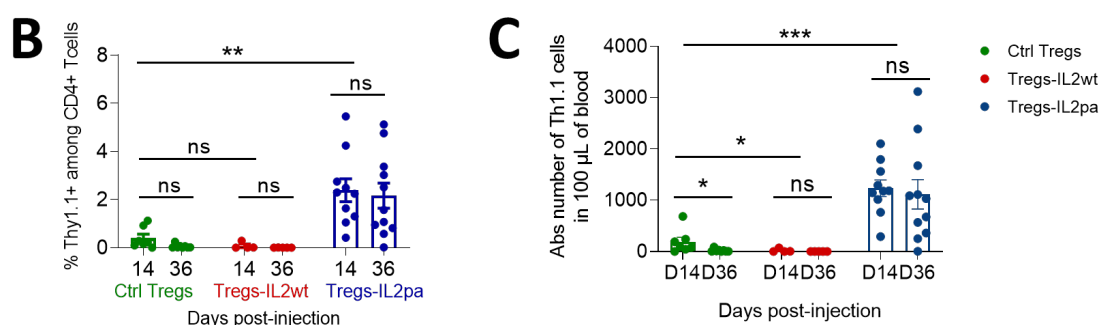


and increase in the number of mice examined as Figure 1D has only $n=2$.

➔ We agree with the reviewer that having more mice would give more confidence in the results. However, in planning our experiments, we do have to abide to the 3R attitude (Replace, Reduce, and Refine, providing a framework for performing more humane animal research). As this figure is only illustrative of the problem in its nature, and does not address the solution to the problem, which is the topic of the article, we think that 2 mice are enough. However, we are now reporting the results of a new experiment assessing the effect of sustained IL-2 supplementation on Ctrl Treg survival (n=3) (Figure 1E and F).

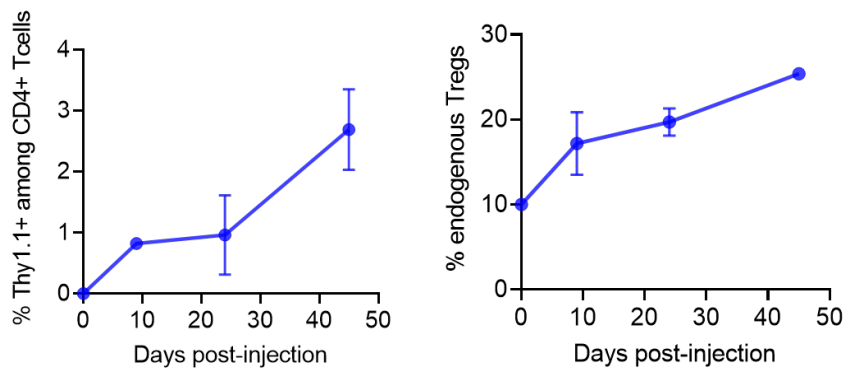
2. Figure 2 is missing statistical analysis. It seems to appear based on the figure legend that “n=2 per group” was performed. However, there are several statements in the figure legend that are confusing. For example, data in (B) is represented of 2 animals per group as stated in the figure legend. However, the figure legend also states “Data were extracted from one of six independent experiments conducted.” This is confusing, was it 2 animals per group in 6 independent experiments, so 12 mice in total per group? Why only show results from 1 experiment and not all 6?

➔ We presented in the article one representative experiment out of a total of six. All these experiments showed a similar pharmacodynamic of the injected cells. Of note, the primary focus of the each repeated experiments was different, such as to perform TCR characterization or single-cell analysis, which resulted in some variation in the time points of blood sampling. These variations limited our ability to pool long-term data across all experiments. However, two critical time points, D14 and D36, were consistently analyzed across independent experiments. Therefore, these time points have been used for pooled data representation in the figures below. Data were extracted from 4 independent experiments, Ctrl Tregs (n=8), Tregs-IL2wt (n=5), and Tregs-IL2pa (n=11). Figure B below, which summarizes the proportion of injected cells on days 14 and 36 post-injection across these multiple experiments, has been included as new Figure 2B. Additionally, Figure C below, showing the absolute numbers of injected cells at the two time points, has been added as Supplementary Figure 2C. The results demonstrate that, starting from day 14, Tregs-IL2pa exhibit significantly improved survival compared to Ctrl Tregs and Tregs-IL2wt. Furthermore, Ctrl Tregs show a decline in survival from day 14 to day 36, whereas Tregs-IL2pa persist at consistent levels.



Data in panel L also states “data were extracted from one of three independent experiments”. Why only show results from 1 out 3?

➔ The data were pooled and now shown on the new Figure 2L.

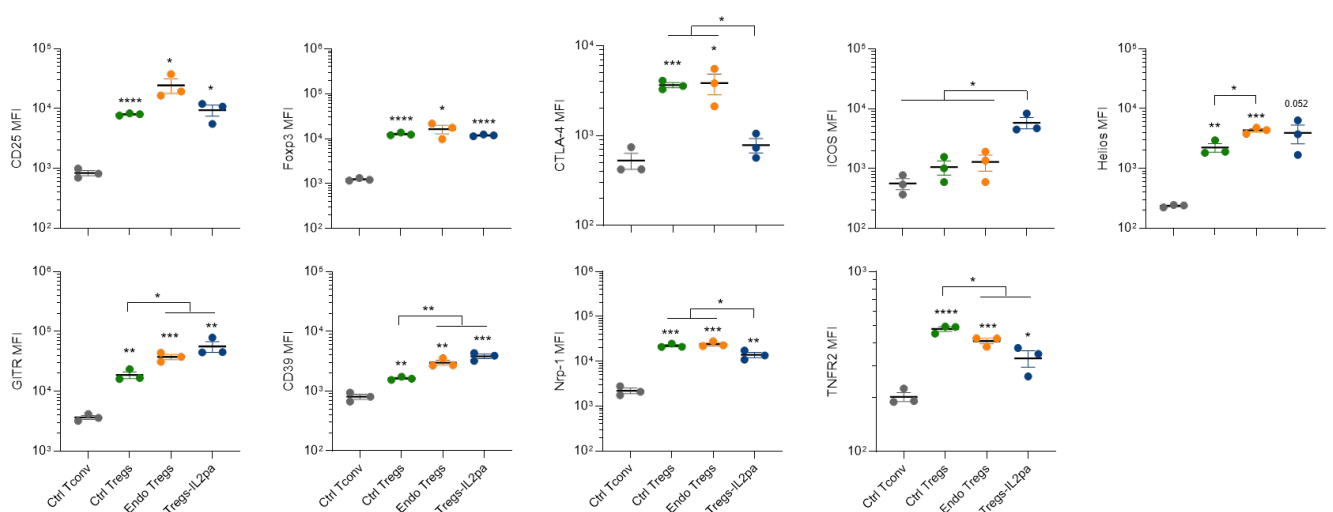


Moreover, several timepoints are given for the analysis of the blood, but there I only limited analysis of the tissues with only a single timepoint examined, day 14 post-Treg infusion. Without any statistical analysis, the low number of mice examined, and the limited analysis of the tissues. it is very difficult to determine the strength of the data.

➔ The tissue analysis was conducted primarily to confirm cell survival profiles observed in the blood and to demonstrate that the cells are also present in various tissues. Our objective was not to track the kinetics of cell distribution across tissues, as this would require analyzing multiple mice at each time point, significantly increasing the number of animals to be sacrificed. However, as indicated, the TCR and single-cell analyses shown in Figure 3 were performed using spleens collected from mice 4 to 6 months post-injection, confirming Tregs-IL2pa survival in the spleen at late time points.

3. In Figure 3, how many mice were examined? Figure 3D should include the MFI on the histograms and a graph showing all the MFI from all the mice examined and statistical analysis performed on the 4 T-cell populations examined.

➔ Three mice per group were analyzed. We now added the MFI values (shown below) for each mouse for all the analyzed markers in a new Supp. Figure 3B.



4. Again, in Figure 4 many graphs lack statistical analysis, only panels E and H have statistics.

➔ We have now added more statistical analyses to this figure.

5. In figure 5, examination of expansion and transduction was performed from sorted human PBMCs CD4+CD25+ CD127-/low from n=2 per group in panel B-D. What was the transduction efficiency in these human cells? Statistics performed in panel B-D?

➔ As indicated in the figure legend, 35% of transduced cells were obtained with both vectors. As the experiment has been performed twice (two different blood donors) no statistical analysis could be performed.

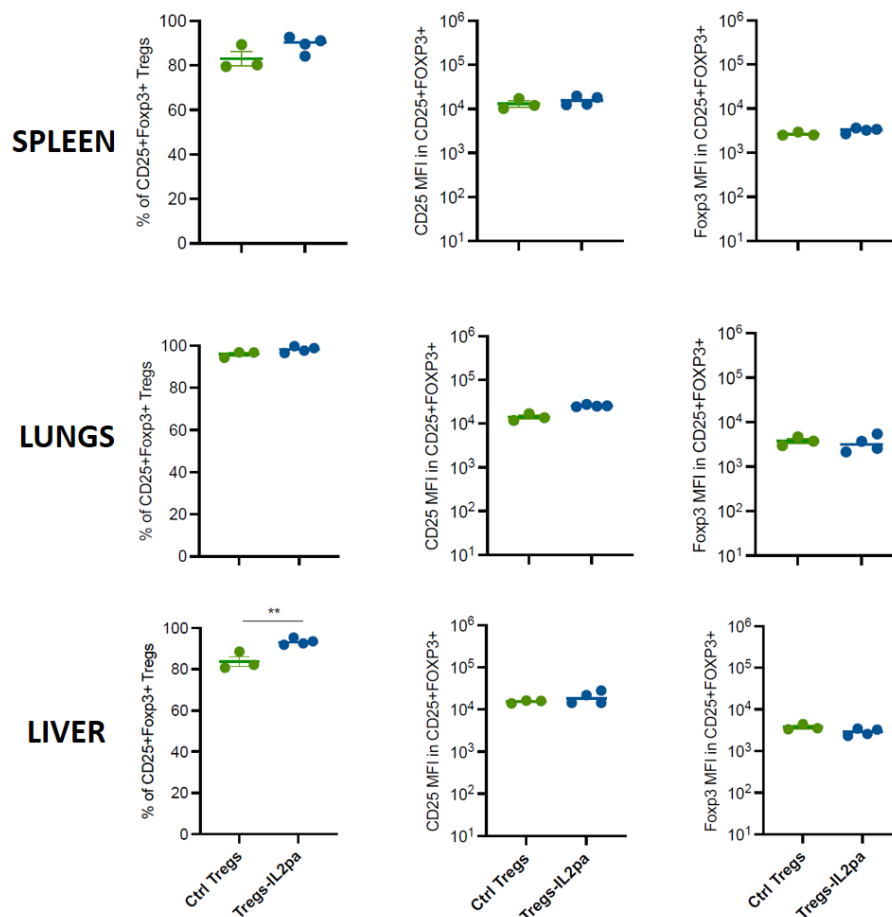
In panel G, data is small due to the inclusion of IL-2 in the upper right corner. An insert that zooms in on the lower portion (not include the IL-2 in the upper right) to improve presentation of the data would be helpful.

➔ We thank the reviewer for the recommendation, we have revised the figure to enhance clarity and readability.

6. Figure 6, Again statistics missing for some of the data. Panels C and D need to include MFI, quantitative analysis through presentation of graphs and statistics of all the data points.

➔ Due to the different group sizes and the variability between the mice (n=3 control Tregs, n=4 Tregs-IL2pa), statistical differences were not obtained in the new Figure 6B that shows the absolute number of hCD45 cells in the organs. However, we have now added as Figure 6A the follow up of human Treg survival in the blood of NSG mice and we show a statistically significant difference between Ctrl Tregs and Tregs-IL2pa survival (n=3 per group).

➔ To complete the panel C, we now provide the percentage of CD25+Foxp3+ cells and the MFI of CD25 and Foxp3 of all mice. Only significant statistical differences are indicated. This figure has been added to the article as Supp. Figure 6A.



➔ For panel D, the analysis was performed on one mouse, the one with the highest injected cell count on the day of sacrifice. Since the percentage of GFP+ transduced cells was low (5%), we prioritized the analyses shown now in panel B, using most of the available cells, leaving no remaining cells for further analyses. However, for one Tregs-IL2pa mouse with a high cell count, we analyzed three additional markers to supplement the data. Due to the limited number of mice, we decided to transfer this figure to Supp. Figure 6B.

7. Although the efficacy data in vivo in figure 4 and 6 shows great promise, there was no discussion on the impact that the IL-2pa-Tregs could have on tumor development or chronic infections. The results shown show high level engraftment and survivability of the donor Tregs cells but also a substantial increase in endogenous Tregs as well. This is a real concern as too many Tregs pose challenges as well and the authors do not even discuss this or a strategy to pursue in the future to explore the safety of their engineering approach.

- ➔ We thank the reviewer for this comment, and finding that “figure 4 and 6 shows great promise”
- ➔ We had the same concern when we developed IL-2 as a Treg stimulator for the treatment of autoimmune diseases (Saadoun et al NEJM, 2011). People were asking about cancer and infection in face of stimulated and expanded Tregs. We addressed these concerns specifically in our paper Churlaud et al, Clinical Immunology, 2014. Results from this article show that with sustained very high levels of Tregs in CD4 cells (>40%), there were no deleterious effects on anti-infectious and anti-cancer immune response. We now discuss this more thoroughly in the article.

Minor Comments

1. Figure 2 is missing labels A, B and C. Please add these labels.

➔ The missing labels have been added.

Again, we thank the reviewer for his/her constructive comments and apologize for giving the impression of having not reproduced our results and for omitting to show some of the statistical analyses.

We thank the reviewer for his/her constructive comments that helped us improve our manuscript.

Reviewer #3 (Remarks to the Author):

This is a timely and consequential manuscript, with the potential to greatly affect the clinical applicability of tTreg adoptive cell therapy, both for autoimmune disease and hematopoietic stem cell transplantation.

➔ We thank the reviewer for these encouraging comments.

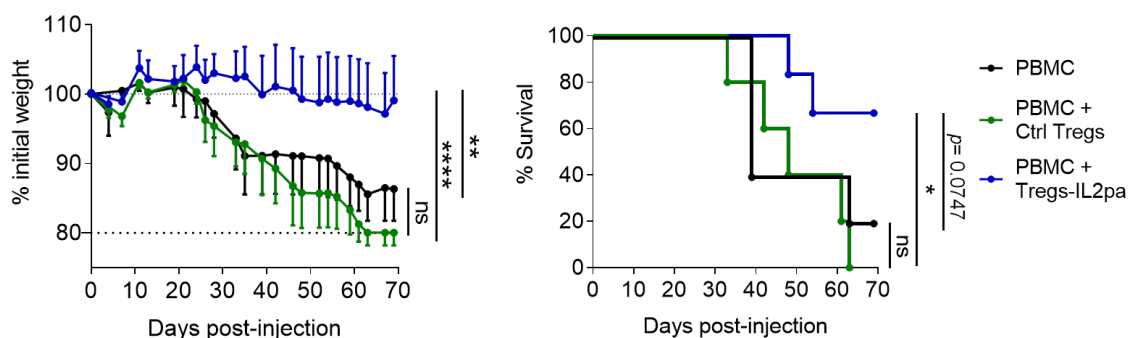
There is one major point and a few minor points that need to be addressed.

Major point:

- Figure 6E: The clinical applicability of Tregs-IL2pa is paramount for this manuscript. While it is understood that using a sub-optimal ratio of Treg is an approach to help demonstrate increased XGVHD efficacy for Tregs-IL2pa, only a single assay is shown, and control Treg do not even show a hint of efficacy. The authors also do not actually list the p-value (and statistical method) for Tregs-IL2pa. This reviewer would like to see the xenoGVHD assay repeated. Important factors would be a significant amelioration of disease with Ctrl Tregs, and enhanced efficacy with Tregs-IL2pa. The weight figure was very effective and should be repeated.

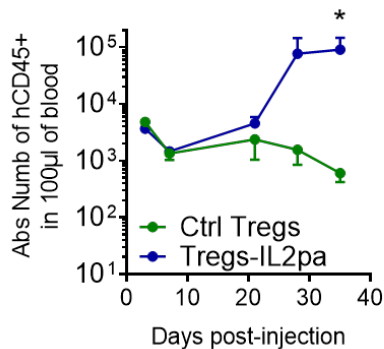
➔ The reason for which plain Tregs “*did not show a hint of efficacy*” is because we on purpose used a ratio of 2 Teff for 1 Treg, while people usually use a 1/1 ratio. This is because we reasoned that to show superiority, one needs a setting where the comparator has poor efficacy.

The data were reanalyzed with the advice of a statistician. Using a Kruskal Wallis test, we now show statistically significant differences in weight loss as shown below, and now reported in the article. Using a Gehan-Breslow-Wilcoxon test, we also show statistically significant differences in survival.



Many investigators have also assessed Treg survival by monitoring Treg persistence in blood (similar to murine assays these authors used). With their previous data, it seems very likely that Tregs-IL2pa will have increased persistence in the blood.

➔ The reviewer is correct, human Tregs-IL2pa also exhibits increased persistence in the blood of NSG mice, as shown in the kinetics below. These kinetics, with n=3 per group, were obtained from a new experiment, distinct from the one presented in Figure 6B. We have now added this figure to the main figures in the new Figure 6A.



Minor points:

- I do not think the designation the authors have chosen for their cell product (e.Tregs) is appropriate since there are already >89 papers that use eTreg as an abbreviation for effector regulatory T cells.

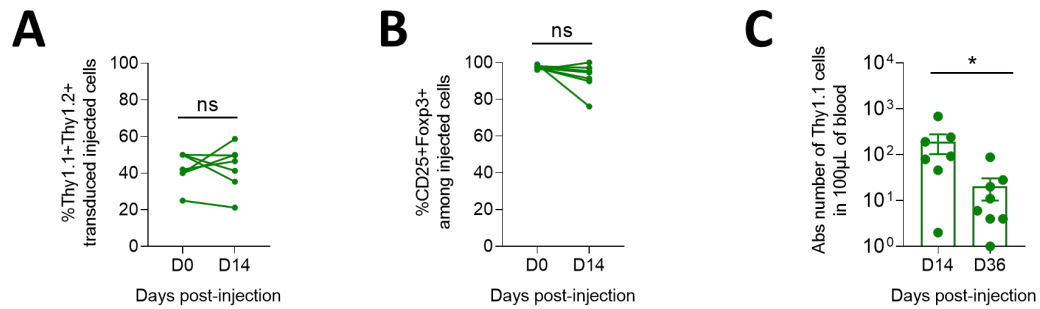
➔ We thank the reviewer for this relevant comment on the designation of our cell product. We have now replaced eTregs with is.Tregs for interleukine-2sufficient Tregs.

- Figure 1: All the data presented appear to be from a single purification/expansion or from single animals receiving adoptive Treg transfer. Summary data should be presented for each panel (A-D) to show the relative variability of the initial set-up.

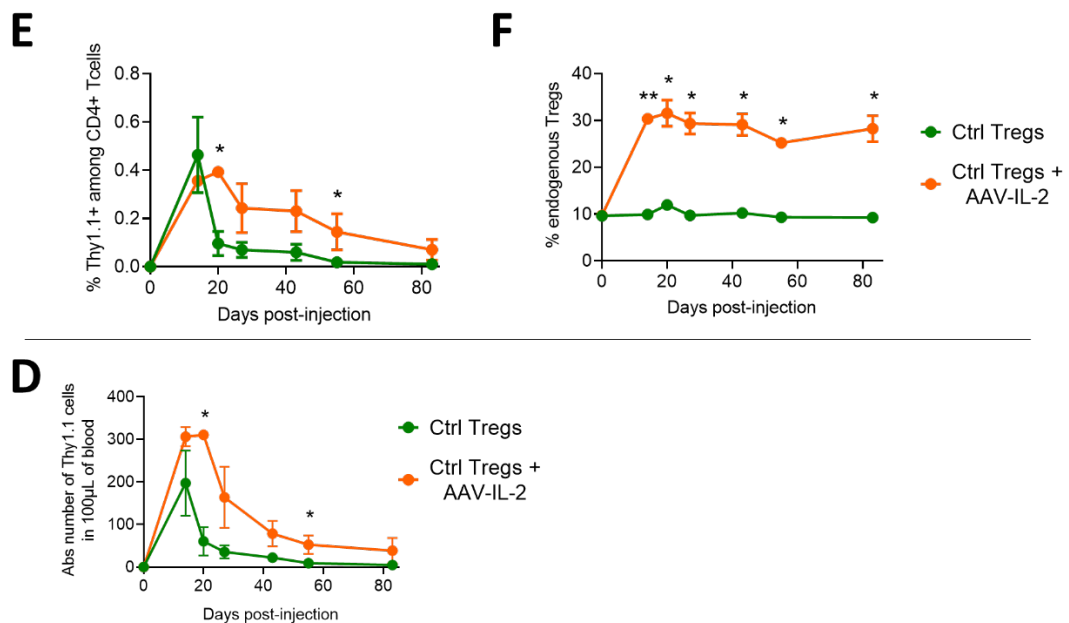
➔ Figure 1A and B purpose is to explain the strategy and shows how the injected Tregs are distinguished from endogenous Tregs.

➔ Figures 1C and 1D illustrate that Ctrl Tregs exhibit poor survival following administration. These figures demonstrate that normal Tregs do not persist in the blood or tissues. To provide quantitative analysis, we pooled data from blood samples across four independent experiments (n=8) and incorporated this into the updated figure.

In Supp. Figure 1A, we added the percentage of transduced cells (n=8), and in Supp. Figure 1B, the percentage of CD25⁺Foxp3⁺ cells within the injected population (n=8), both from the day of injection to 14 days post-injection (see Figures A and B below). Additionally, we included the absolute number of injected cells in the blood in Supp. Figure 1C (see Figure C below), showing data from days 14 and 36 post-injection. These results demonstrate a significant reduction of Ctrl Treg number in the blood from 14 to 36 days post-injection.



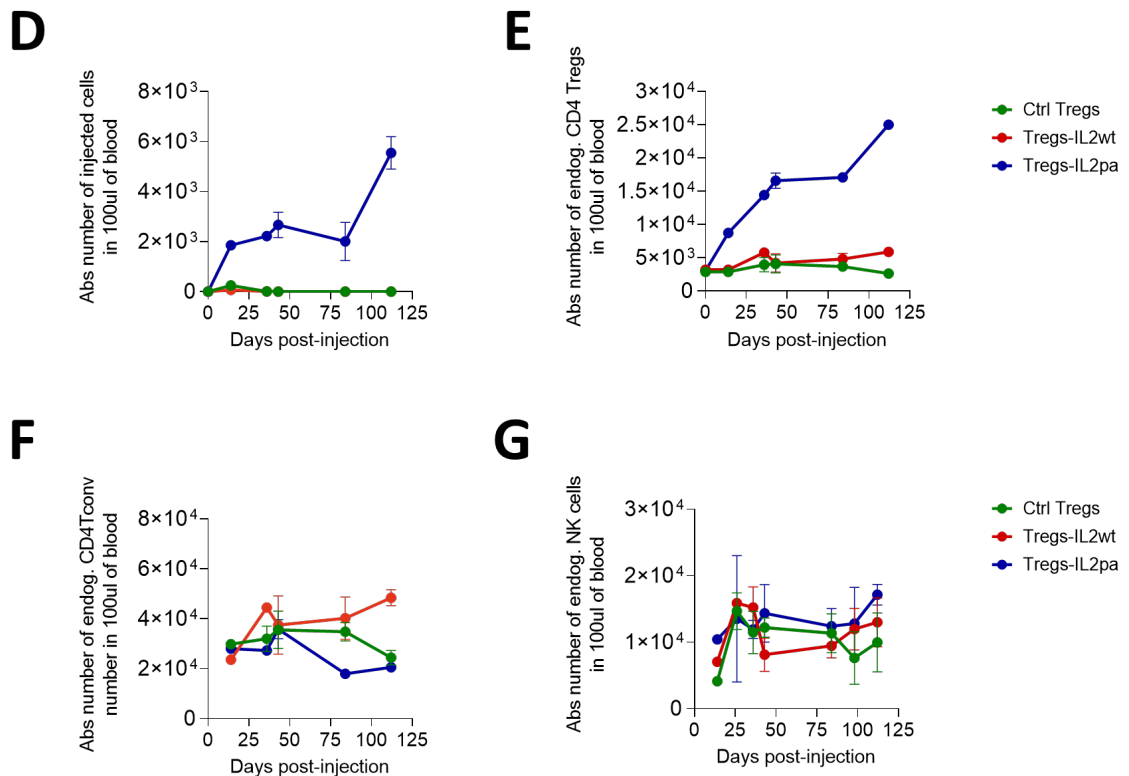
➔ Figure 1D demonstrates that exogenous IL-2 does not rescue the survival of Ctrl Tregs. Initially, this figure included a sample size of $n=2$. Rather than adding more mice to increase the sample size, we opted for a more robust approach analyzing the survival of Tregs in the presence a persisting IL-2 treatment. We injected an AAV-IL2 which provides a sustained IL-2 production for as long as the follow up. We compared the survival of Ctrl Tregs co-injected with AAV-IL2 ($n=3$) to Ctrl Tregs without AAV ($n=8$). These results demonstrate that AAV-IL2 injection improves only the early survival of the injected cells. We now show the results of this experiment.



• Figure 2 (summary data is well represented here): Because there is a time-dependent change in both endogenous and Tregs-IL2pa, it is necessary to assess effects via cell# instead of % (of CD4+ or endogenous Tregs). This will also help to show whether there is an overall expansion in endogenous CD4+ conv. T cells as well.

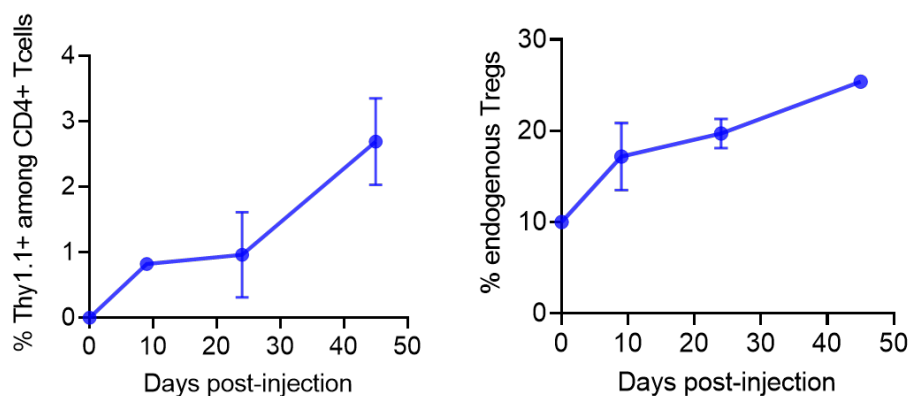
➔ Cell numbers may vary significantly between analysis days, partly due to potential variability in red blood cell lysis, which can differentially affect white blood cell counts. Despite these variations, as shown in the figures below, the observed profiles in percentages are consistent with the absolute cell numbers. Tregs-IL2pa able to persist and expand (Figure D), leading to

an increase in endogenous Tregs (Figure E), while the numbers of Tconv and NK cells remain comparable to those in control mice (Figures F and G). We have now included these figures in the manuscript in Supp. Figure 2D to G).



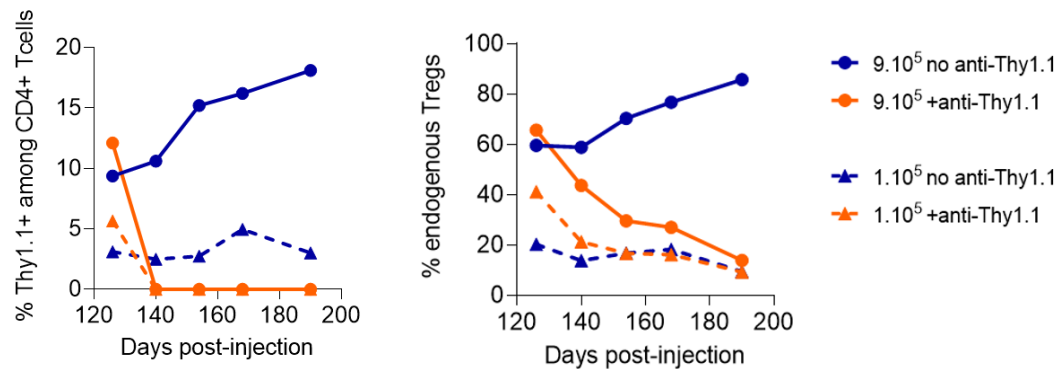
Finally, Figures 2L-M seem to be only n=1. Is this correct?

➔ Yes. In Figure 2L, we are just showing that a transfer from one mice to another is feasible, which only requires one mouse. Nevertheless, we are now reporting 3 repeats of this same experiment, with similar results. The figures below show the kinetics obtained in 3 secondary recipient mice.



➔ In Figure 2M, we demonstrate that cells can be effectively depleted for at least two months using a depleting antibody. The depletion was performed in two mice, as shown in the figures below, with one having received 9×10^5 Tregs-IL2pa and the other 1×10^5 Tregs-IL2pa. In both

cases, the target cells were successfully depleted, and endogenous Tregs returned to their baseline percentages following Treg-IL2pa depletion. As again, these results only illustrate feasibility, we have chosen to show only the 9×10^5 cell dose treated with anti-Thy1.1 in Figure 2M to improve the clarity of the representation.



While perhaps a bit of semantics, I question the authors statement that IL2pa increases the ‘survival potential’ of Tregs. Specifically, it appears that Tregs-IL2pa expand in number even over the first 14 days. Fig. 2H shows that when only 9×10^5 Tregs-IL2pa are injected on day 0, $>1 \times 10^6$ Tregs-IL2pa were present in just the spleen.

➔ We agree with the reviewer’s observation that the data from day 14 suggest an expansion of cells within the first 14 days. As a result, we have revised the term to “survival and expansion potential” to better reflect this finding.

- Figure 3E: The authors anticipated questions on TCR clonality in Tregs-IL2pa, and I applaud the stringent conditions they chose to address the issue with (following 6 months of expansion in vivo). These data showed that Tregs-IL2pa had much lower diversity than any of the controls (control naïve and activated memory Tregs, or endogenous Treg present in Tregs-IL2pa injected animals). The authors statement in both the abstract and discussion is that Tregs-IL2pa have a diverse TCR repertoire. While technically true, I think this is somewhat misleading. The authors suggest that Tregs-IL2pa had the lowest diversity because ‘they originate from a necessarily less diverse repertoire of cultured and expanded cells’, but I don’t think CD3/28 polyclonal stimulation has been shown to effect clonality. Also, Tregs-IL2pa had lower diversity than the endogenous Tregs with which they co-expanded in vivo.

➔ As mentioned by the reviewer, the repertoire is truly polyclonal with 3000 unique clonotypes per sample of 500,000 sorted transduced Tregs-IL2pa. We make this clearer in the text. Regarding the statement ‘they originate from a necessarily less diverse repertoire of cultured and expanded cells,’ we think that the culture period does reduce TCR diversity, as not all clones are transduced and are not likely to expand equally. In any case, the polyclonality of the repertoire is highlighted by the control of a multiorgan autoimmunity as tested in Figure 4. More thorough repertoire studies, including at the single cell levels, will help better understand the Treg-IL2pa diversity, but fall outside the main scope of this paper.

Figure 5G/H (and I): no indication is given as to whether the control Treg were given exogenous IL-2. This is somewhat confusing, especially given that the preceding figures (5B-F) looked at both possibilities. If exogenous IL-2 was included, please expressly denote. If not, I would question the

comparison between an actively expanding population (Tregs-IL-2pa) and a dying population (Ctrl Tregs) (from 5C).

- ➔ In contrast to Figure 5C, the cells from Figures 5G/H were cultured without exogenous IL-2 for only 3 days instead of 5 days. From their microscopic evaluation, and the RNA amounts recovered, they were not dying.

Points that need to be addressed in discussion: In the NSG setting, did Tregs-IL-2pa result in increased expansion/persistence of co-transferred non-transduced cells (GFP- cells shown in 6D)?

- ➔ Yes, the GFP- cells survived as well. Of note, there are no cells that could consume IL-2 in NSG mice, and all the excess IL-2 produced by Treg-IL2pa can be used by the non-transduced cells.

The authors document that both endogenous Tregs and Tregs-IL2pa significantly increase over time, to supra-physiological levels. While they discuss their own publication assessing immune suppression with low-dose IL-2 in NOD mice, they should also discuss publications showing that high Treg numbers can compromise immune control in mice (The Journal of Immunology, 2011, 186: 4862–4871). Generalized immunosuppression is a very important consideration in human adoptive tTreg therapy and needs to be discussed.

- ➔ We thank the reviewer for this suggestion. We had the same concern when we developed IL-2 as a Treg stimulator for the treatment of autoimmune diseases (Saadoun et al NEJM, 2011). People were asking about cancer and infection in face of stimulated and expanded Tregs. We addressed these concerns specifically in our paper Churlaud et al, Clinical Immunology, 2014. Results from this article show that with sustained very high levels of Tregs in CD4 cells (>40%), there were no deleterious effects on anti-infectious and anti-cancer immune responses. We now discuss this more thoroughly in the article.

We thank the reviewer for his/her constructive comments that helped us improve our manuscript.

Reviewer #4 (Remarks to the Author):

In this manuscript, Robert et al provide intriguing results on the genetic modification of regulatory T cells (Treg) to confer what they call “self-sufficiency” for IL-2 signaling and improve on current limitations (namely survival, expansion, and stability) incurred by Treg-based therapies for different immunological disease. The key novelty is the identification of the need to use a gene encoding a partial agonist form of IL-2 (IL-2pa) to enhance the properties of adoptively transferred cells. It is noteworthy that the study includes both murine and human Treg investigations, providing important proof that the proposed strategy has translational potential and is clinically relevant. Overall, the results presented are promising and could have a meaningful impact on the design of more effective Treg therapies. Unfortunately, there are also multiple major points that are not addressed by the authors, reducing significantly the enthusiasm for the manuscript in the current form.

Major Points:

1. The most important weakness of the manuscript is the absence of information on the partial agonist of IL-2 used in this work. The authors do not provide a rationale of how they arrived at the

version implemented and this sequence is not provided (and others tested). Moreover, it is not clear if this IL-2pa sequence is the same or not between mouse and human Treg studies.

- ➔ We did indeed discuss in the article the rationale for how we arrived at the version that was implemented. We stated: "*We noticed that the percentage of Thy1.2+ transduced cells within the Tregs-IL2wt population (which also contains untransduced cells) increased sharply from about 40% at injection to 80% at day 14, and then decreased sharply to disappear by day 25 (Figure 2C). This indicates an initial growth advantage of the IL-2wt transduced cells, followed by a marked growth disadvantage, indicating toxicity. We hypothesised that this could be due to excessive tonic signalling by the continuous production of IL-2wt. We therefore aimed to evaluate the fate of Tregs expressing IL-2 with partial agonist activity (IL-2pa)*".
- ➔ We then gave information on how we selected the sequence, based on the evaluation of the partial agonist activity of different human IL-2 mutants (based on their reduced Stat-5 phosphorylation upon activation of the IL-2 receptor), as shown in Supplementary Figure 2B. We then performed preliminary survival experiments in mice to select our preferred human IL-2pa.
- ➔ The specific sequence we chose is proprietary. In order to give this strategy a chance to be developed in the clinic, we cannot disclose it at this time. However, **we would like to emphasize that (i) we are providing enough mechanistic information to understand our results and (ii) we are disclosing enough information for others to identify their own suitable IL-2pa for the purpose of generating efficient IL-2 self-sufficient Tregs. In fact, we are disclosing the only important parameter, which is the degree of loss of biological activity of IL-2pa (approximately 1000 fold based on pSTAT5).**

2. It is remarkable that the transferred Treg(IL-2pa) keep expanding in the recipient and, even more impressive, that they support the expansion of endogenous Treg to the point that they constitute more than 60% of total CD4 T cells. As pointed out by the Authors themselves, this progressive expansion could become an issue of global immunosuppression. The authors should test this concern (they have published a previous paper using established models). The simple mention in the Discussion that their previous study (using a viral vector for sustained IL-2 production in vivo) did not indicate there should be a concern is not a valid argument for the conditions of this very different model system.

We agree with the reviewer that this would be an important question if human Tregs-IL2pa behaved like murine Tregs. However, the kinetics of Tregs-IL2pa expansion in humans is likely to be very different from that observed in mice and cannot be predicted at this stage. Indeed:

- ➔ Treg proliferation requires IL-2 + TCR stimulation. It is likely that a Treg has to travel longer to finally meet its antigen in a 70kg human than in a 30g mouse.
- ➔ The time for division after activation is faster in mice than in humans, as in most biological processes.
- ➔ In addition, and most importantly, since TCR stimulation would be eliminated in CAR-transduced Tregs, this will dramatically change the kinetics of expansion.
- ➔ Finally, this will not be an issue if allogeneic Tregs were to be used.
- ➔ In all cases, the expression of a kill switch will be used in first-in-man studies.

We improved the discussion on these points.

- ➔ We respectfully disagree on the fact that our previous work "is not a valid argument". We had the same concern of excessive immunoregulation when we developed IL-2 as a Treg stimulator

for the treatment of autoimmune diseases (Saadoun et al NEJM, 2011). People were asking about cancer and infection in face of stimulated and expanded Tregs. We addressed these concerns specifically in our paper Churlaud et al, Clinical Immunology, 2014. Results from this article show that with sustained very high levels of Tregs in CD4 cells (>40%), there were no deleterious effects on anti-infectious and anti-cancer immune response. This teaches that very high activated Treg levels do not generate a true immunosuppression. Whether IL-2 is endogeneous or exogeneous, it activates and expands Tregs that are functional. And as discussed above, we will not reach the same levels of expansion in humans. We now discuss this more thoroughly in the article.

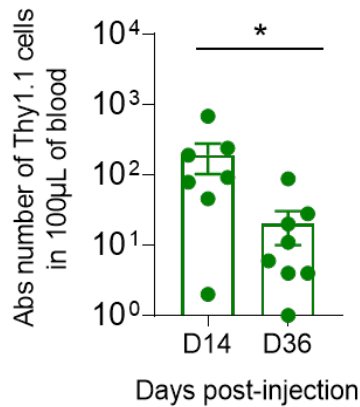
3. The in vivo model of Treg(IL-2pa) suppression, based on the Foxp3-DTR mouse model, is not really a “therapeutic model”. The results clearly show that Treg(IL-2pa), injected 20 days before DT treatment, can replace endogenous Treg and prevent the inflammation induced by the depletion induced by DT. However, it does not prove that Treg(IL-2pa) are therapeutic. They should be tested in a more clinically relevant model (as done in the past by the group).

➔ We agree that we show that Treg-IL2pa can replace endogenous Tregs but we also think that it is the best possible model to prove the functionality of these Tregs. It is well documented that Treg depletion triggers the stimulation of multiple Teffs that will cause a multiorgan autoimmunity. By using this model, we thus not only show that our Tregs-IL2pa can control the activation of Teff, but also for numerous autoantigens.

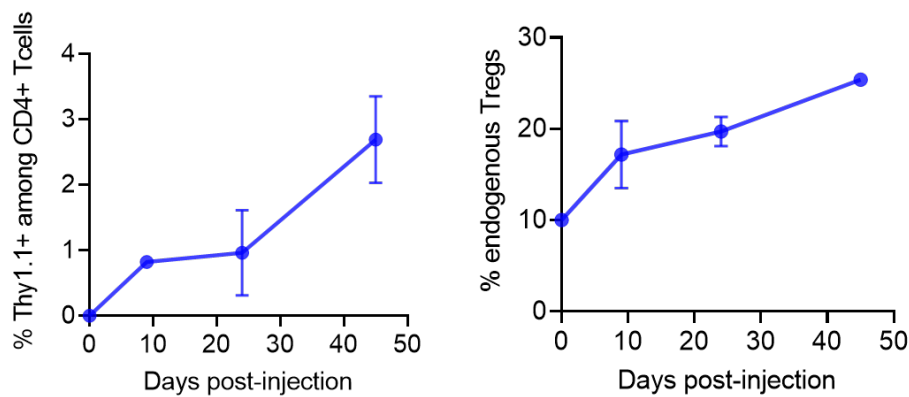
4. In many parts of this manuscript figures, there is no cumulative data presented (only representative) limiting our ability to draw robust conclusions (e.g. Figure 1B-D, Figure 2L-M, Figure 3D). As importantly, in many instances there is a lack of error bars and statistical analysis. The most prominent example is Figure 6E. This is probably the piece of data that provides the most important indication of the therapeutic potential of human Treg (IL-2pa). Unfortunately, there is no statistical analysis presented and we have to believe there is no statistically significant difference between groups.

We apologize that we gave the impression that we had not reproduced experiments, which is of course not the case.

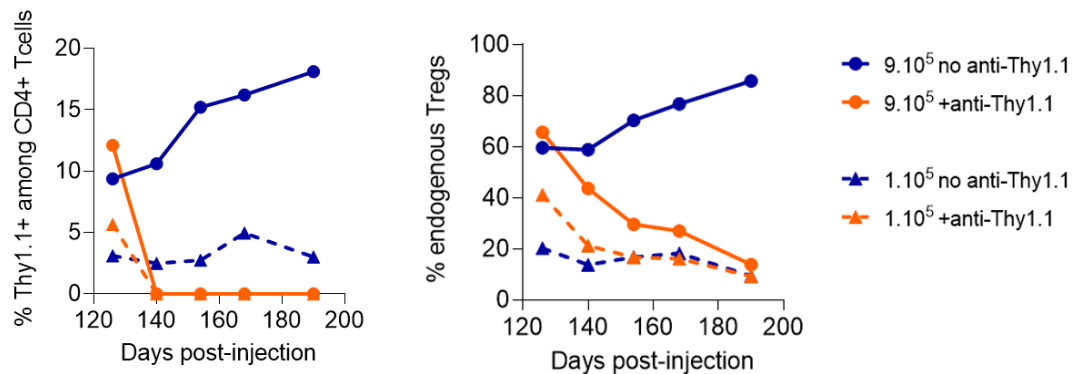
➔ Figure 1 aims to enhance the reader's understanding of the rationale for our work. It demonstrates that normal Tregs have poor survival in both blood and organs. We designed it as an “illustration”. This experiment was reproduced many times, but not always with the exact same time points. Nevertheless, to reassure the reviewer, we compiled the results of experiments for which we have data at day 14 and 36 in all the mice (n=8 from 4 independent experiments) and showed a significant decrease in the number of Tregs at day 36. The figure below has been added to the paper in Supp. Figure 1C.



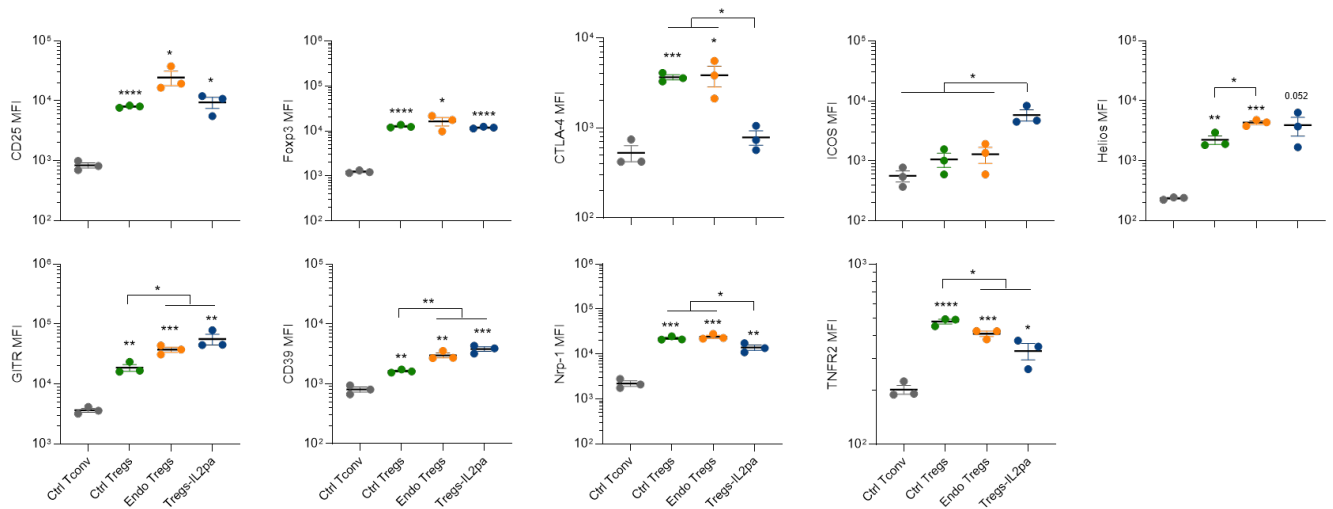
- ➔ In Figure 2L, we are just showing that a transfer from one mice to another is feasible, which only requires one mouse. Nevertheless, we had performed this same experiment 3 times with similar results. The figures below show the kinetics obtained in 3 secondary recipient mice and replaced the previous ones in the paper.



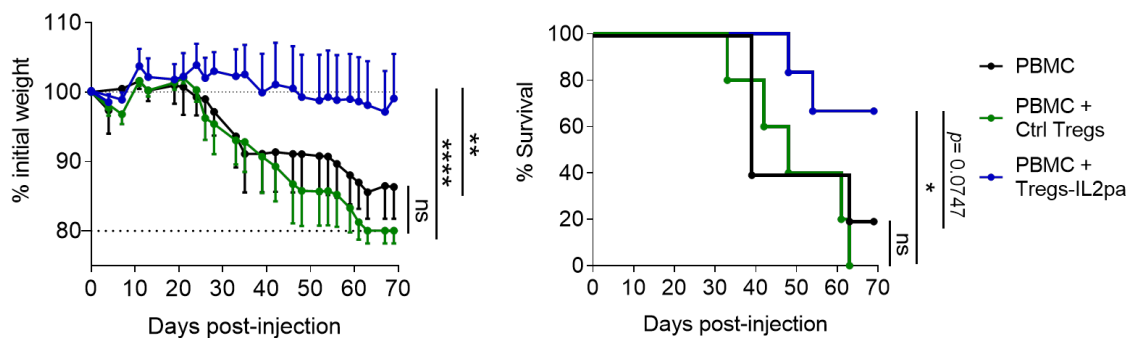
- ➔ In Figure 2M, we demonstrate that cells can be effectively depleted for at least two months using a depleting antibody. The depletion was performed in two mice, as shown in the figures below, with one having received 9×10^5 Tregs-IL2pa and the other 1×10^5 Tregs-IL2pa. In both cases, the target cells were successfully depleted, and endogenous Tregs returned to their baseline percentages following Treg-IL2pa depletion. We have chosen to show only the 9×10^5 cell dose treated with anti-Thy1.1 in Figure 2M for the clarity of the illustration.



➔ In Figure 3D, three mice per group were analyzed, we now added in Supp. Figure 3B, the MFI values (shown below) for each mouse for all the analyzed markers.

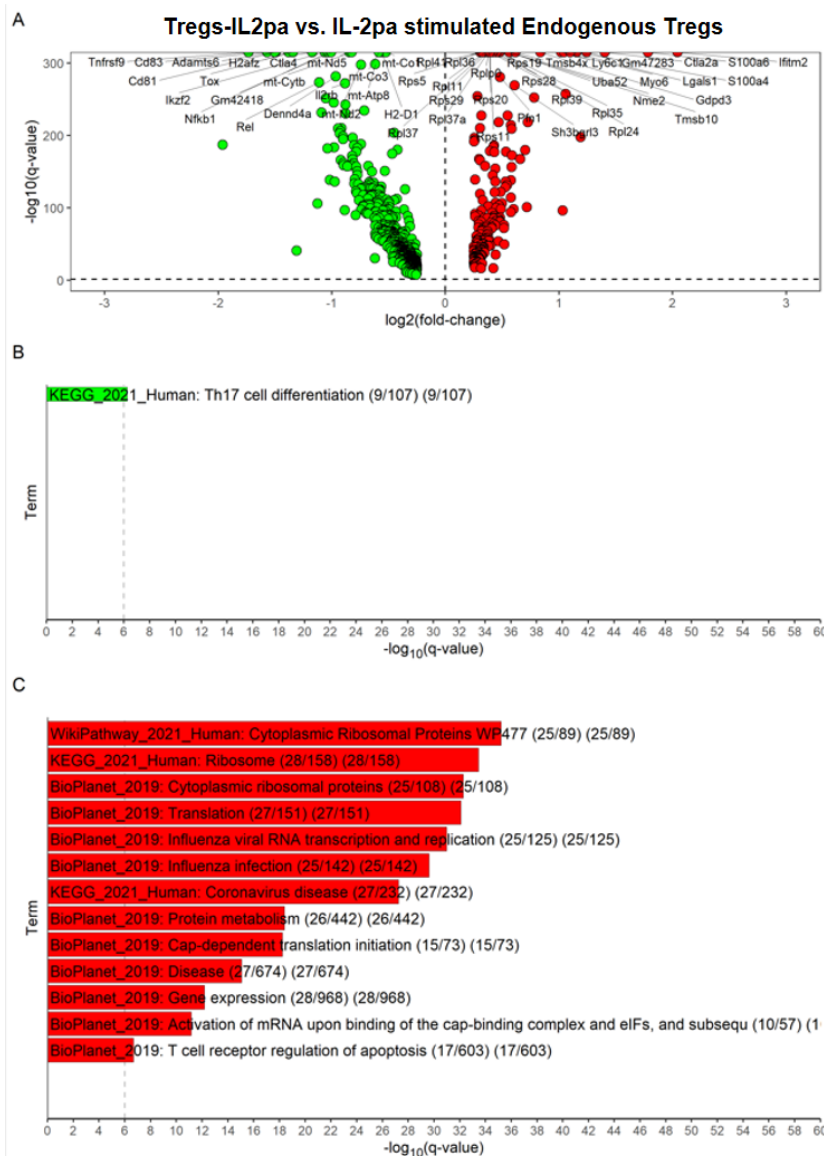


➔ The data were reanalyzed with the advice of a statistician. Using a Kruskal Wallis test, we now show statistically significant differences in weight loss as shown below, and now reported in the article. Using a Gehan-Breslow-Wilcoxon test, we also show statistically significant differences in survival.



5. The scRNAseq analysis of murine Treg groups is potentially very interesting, but the results presented only focus on showing a similarity between Treg(IL-2pa) and endog Treg. The UMAP shows that these two populations also have significant elements of distinction, and it would be really useful to know where these differences reside.

- ➔ The UMAP can indeed separate the populations, but also shows that they overlap, and does not provide an estimate of the effect size for the differences.
- ➔ Figure A below shows a volcano plot comparing Tregs-IL2pa with IL-2pa-stimulated endogenous Tregs. Figure B and C shows the functional enrichment analysis, highlighting the biological pathways differentially expressed by Tregs-IL2pa and endogenous Tregs, respectively. These data indicate that the main difference between the two populations are related to translation and metabolism.



In similar fashion, the scRNAseq of human Treg(IL-2pa) is presented to indicate an increase in the IL-2 signaling pathway and in glycolysis. What is the meaning of it?

- ➔ Figure 5E show that there are minimal differences between Ctrl Tregs and Tregs-IL2pa. When the analysis includes fresh Tregs, all the cultured cells that had access to IL-2 are indistinguishable. It is only when comparing them without the fresh Tregs that some (minor) differences appear. It shows that Tregs-IL2pa have a slightly more activated IL-2 signaling pathway and an increased metabolism, in line with the demands of their high activation state.

And could the authors draw any parallelism between the mouse and human analysis?

- ➔ It is challenging to draw parallels between the mouse and human analyses. However, both experiments indicate that the Tregs-IL-2pa have an increased activation state and metabolism.

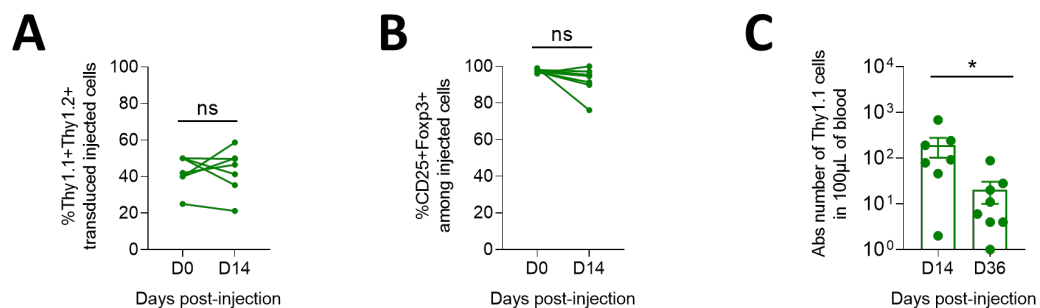
Finally, could these dataset shed any light on why Treg with wt IL-2 disappear suddenly? The authors imply toxicity but no data is showed to support this interpretation.

- ➔ We respectfully disagree. As shown in Figure 2C, Tregs-IL2wt exhibit first a rapid expansion with an increase in the percentage of the transduced vs non-transduced cells among the adoptively transferred cells. This clearly indicates an initial survival advantage. Thereafter, this is followed by a sharp disappearance of these cells, to levels even lower than for the control Tregs. This clearly indicate toxicity. We hypothesized that this toxicity is likely due to excessive IL-2 stimulation, especially considering that Tregs with IL-2pa exhibit a 1,000-fold reduced IL-2 biological activity compared to IL-2wt (as shown in the pSTAT5 analysis in Supp. Figure 2B). IL-2pa solve this issue, supporting our hypothesis.

Minor points:

1) Figure 1, panels B and C would be even more informative if, in addition to providing cumulative data (see above), would show the ratio between transduced and non-transduced injected cells. It would be nice to see Foxp3 expression as well.

- ➔ Figure 1A and B purpose is to explain the strategy and shows how the injected Tregs are distinguished from endogenous Tregs.
- ➔ Figures 1C and 1D illustrate that Ctrl Tregs exhibit poor survival following administration. These figures demonstrate that normal Tregs do not persist in the blood or tissues. To provide quantitative analysis, we pooled data from blood samples across four independent experiments (n=8) and incorporated this into the updated figure.
- In Supp. Figure 1A, we added the percentage of transduced cells (n=8), and in Supp. Figure 1B, the percentage of CD25⁺Foxp3⁺ cells within the injected population (n=8), both from the day of injection to 14 days post-injection (see Figures A and B below). Additionally, we included the absolute number of injected cells in the blood in Supp. Figure 1C (see Figure C below), showing data from days 14 and 36 post-injection. These results demonstrate a significant reduction of Ctrl Treg number in the blood from 14 to 36 days post-injection.



2) Figure 2 is missing A, B, C labels

- ➔ The missing labels have been added.

3) Figure 3, organize the markers in panel B and D in the same order to improve readability.

- ➔ The order of the panels has been reorganized.

Additionally, could the authors better explain in the text what the rationale is to wait 6 month post injection to analyze Treg?

- ➔ The rationale for waiting several months before analyzing the cells is to demonstrate that the results are sustained over the long term. For instance, even one year post-injection, Tregs-

IL2pa maintain full demethylation of the TSDR, and six months post-injection, Tregs-IL2pa continue to express key functional and activation genes.

4) Figure 4, after panel A, there is no inclusion of ctrl Treg in all other panels.

➔ Ctrl Tregs were not included in this experiment because the disease was induced 20 days post-injection. At this time point, we know that Tregs-IL2pa still persist, whereas Ctrl Tregs are not.

5) Figure 5G, it is hard to read the volcano plot, can IL-2 be excluded to better show the changes in gene expression?

➔ The figure has been improved for more readability.

We thank the reviewer for his/her constructive comments that helped us improve our manuscript.

Reviewer #1 (Remarks to the Author):

In my view, the revised version of the manuscript addresses most of the reviewer's concerns and would be suitable for publication. However, one significant issue remains unresolved, as the revised manuscript still lacks sufficient information about the compound used. After careful consideration, I am willing to support the publication of the revised manuscript on the condition that the authors provide the relevant information about the reagents or make the reagents available to the scientific community under specific terms (such as an MTA, NDA, or formal collaboration), and that Nature Communications proceed with retraction of the article should the authors fail to comply.

→ We sincerely thank the reviewer for their positive feedback and for supporting the publication of our manuscript. In response to the reviewer's request for additional information regarding the IL-2 compound used in our study, we are committed to providing the relevant details. Specifically, as suggested by the reviewer, we are willing to make this information available under appropriate terms, such as through a Materials Transfer Agreement (MTA), Non-Disclosure Agreement (NDA), or formal collaboration.

Reviewer #2 (Remarks to the Author):

This revised manuscript by Robert et al., entitled "Enhancing regulatory T cell efficacy through self-sufficiency for an interleukin 2 partial agonist" aims to improve the long-term engraftment of adoptively transferred T regulatory cell by engineering Tregs that produce their own source of IL-2. Overall, the authors have addressed many of the reviewers' comments and added to the discussion. A major comment and some minor comments are below.

Major comments:

1) Still have concerns about the safety of their Treg product due to the high-level engraftment and survivability of the donor Tregs cells but also a substantial increase in endogenous Tregs as well. This is a real concern as too many Tregs pose challenges as well and the authors do not even discuss this or a strategy to pursue in the future to explore the safety of their engineering approach. Although this group has previous data from other models that a substantial increase does not impair immune responses or increase susceptibility to tumors, future studies should be performed to specially address this potential concern and nor so easily dismissed because of their previous work with a different approach that resulted in an increase and sustained Treg population.

→ We agree with the reviewer that this is an important issue. However, we believe that the level of expansion observed in our experimental context will likely be much reduced in real-world applications, as the outcome will depend heavily on the

specific context of use. For example: (i) in the case of CAR-is.Tregs, their expansion will be influenced by the intensity of stimulation provided by the CAR; (ii) for allogeneic cells, expansion will be limited by their potential ablation due to the allogeneic response. Obviously, the persistence of is.Tregs will be one of the main biological parameters to be monitored in first-in-man use.

➔ We have added a sentence in the discussion to reflect these points.

Minor comments:

1) Page 7 line 154: It states figure 1B but believe it should be Figure 2B. Please clarify.

➔ We thank the reviewer for pointing out this error. Indeed, the reference should be to Figure 2B, and we have corrected this mistake in the manuscript.

2) Page 7, lines 155-157. Should add that this is data from Figure 2A, the observed 10% at later timepoints for Tregs-IL2pa Recipients.

➔ We thank the reviewer for the suggestion. As recommended, we have added the reference to Figure 2A in the manuscript for the relevant lines.

However, Figure 2B shows that day 14 and day 36 the %Thy1.1 among the CD4 T cells averages around 2% (with a couple of mice higher around 5%) which differs from the Figure 2A right panel significant which shows around 5% with no error bars indicating that the data is consistent. What is the difference between panel A and B. Authors state in their rebuttal that day 14 and day 36 were consistent time points over multiple experiments (Panel B, n=5-11 mice) and it seems that Panel A is only 2 mice per group. In panel B, which data points are the 2 mice per group from panel A? Thinking in the Tregs-IL2pa group that they are the 2 highest points in panel B, then where are the error bars in panel A. Please clarify,

➔ The long-term follow-up experiment shown in panel A initially included 3 mice per group. At D14, one mouse per group was sacrificed, and the remaining two mice were followed from D25 to D127. In Tregs-IL2pa group, since the injected cell number remained stable from D14 to D36 at least, as observed even in the pooled data from different experiments (panel B), we decided to include the data from the first mouse at D14 in the panel A. This mouse had one of the higher injected cell percentages (4.72%) at D14 in panel B, although it was not the highest (the highest value came from another experiment). As this was a single value, no error bar could be obtained for this time point. At D36, the remaining two mice were considered for the long-term follow-up in panel A. These two mice corresponded to the highest percentages of injected cells at D36 in panel B (5.12% and 4.76%). The consistency between the two mice at this time point explains the absence of an error bar at D36 in panel B.

3) Please clarify in the figure legend and text for Figure 4A that the in vitro Treg suppressor assay is using day 3 in vitro transduced Tregs and not Thy1.1+ transduced Tregs isolated from adoptively transferred mice. It is possible that the Tregs-IL2pa isolated from adoptively transferred mice could have higher suppressive capacity compared to the Tregs-CTRL if

isolated early timepoint on a per cell basis. Although the in vivo experiment shows that the Tregs-IL2pa have suppressor function, it would also be interesting to know whether the Tregs-IL2pa or even the stimulated endogenous Tregs are more robust suppressor compared to Tregs from unmanipulated recipients.

- ➔ The legend and text for Figure 4A has been modified to clarify that the Tregs used in this experiment are culture cells, obtained three days post-transduction.
- ➔ Regarding the in vitro suppression assay with the cells that expanded in vivo, we thank the reviewer for the suggestion. It is indeed an interesting experiment, and while we are unable to conduct it at this stage, we plan to include it in future experiments.

Reviewer #3 (Remarks to the Author):

Thank you for addressing the comments point by point.

Reviewer #4 (Remarks to the Author):

In this revised version of the manuscript, Robert et al provide some additional data in response to the concerns raised by the reviewers. Here are comments specific to the points raised by this reviewer.

Major Points:

1. The authors indicate that the modifications introduced in the sequence of human IL-2 to generate a partial agonist are proprietary and cannot be disclosed. This is unfortunate but understandable. However, they do not provide enough information for other investigators to “identify their own suitable IL-2pa”. Additional information that does not undermine the patentability of their invention could be provided to support the reproducibility of their work.

- ➔ We thank the reviewer for understanding that we have not yet disclosed the exact nature of the IL-2 used in this work. To comply with Nature Communications' policy on data sharing and to ensure reproducibility, we are willing to make this information available under appropriate terms, such as through a Materials Transfer Agreement (MTA), Non-Disclosure Agreement (NDA), or formal collaboration.

2. Thank you for adding to the Discussion a few points outlining the probable differences between the results presented in the murine model and the behavior that human Treg might have clinically. Such a point is important to define potential limitations of the work that have not been addressed yet as well as potential risks that will need to be accounted for.

3. In relation to the in vivo model of mouse Treg(IL-2pa) suppression, if the authors don't want to add any more convincing experimental model of therapeutic efficacy, they should change the definition from "therapeutic" to "prophylactic"

➔ We indeed assessed the efficacy of Tregs-IL2pa in the prevention of two models: autoimmunity induction for murine Tregs and xenogeneic GVHD for human cells. We had previously mentioned "therapeutic efficacy" at the end of the introduction to describe the results reported, but we have now replaced it with "prophylactic efficacy".

4. Thank you for addressing the issue of lack of cumulative data and statistics. The data can be better interpreted now.

5. The new figure (panels A, B, and C) presented in the rebuttal should be included as Supplementary Data in the manuscript.

➔ We thank the reviewer for the comment. We have added the functional enrichment analysis in the Supplementary Data.

Minor Points

All properly addressed.