Correlation of blood TCR repertoire with an organ manifestation in a mouse model of autoimmune disease

Tatsuya Ishikawa[1,2], Kenta Horie[1], Yuki Takakura[1], Houko Ohki[1,2], Yuya Maruyama[1,2], Mio Hayama[1,2], Maki Miyauchi[1,2], Takahisa Miyao[1,2], Nobuko Akiyama[1]*, Taishin Akiyama[1,2] *

[1]Laboratory of Immune Homeostasis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan
[2]Graduate School of Medical Life Science, Yokohama City University, Yokohama, Japan
*For correspondence: taishin.akiyama@riken.jp or nobuko.akiyama@riken.jp

Abstract

One hallmark of some autoimmune diseases is the individual variability of symptoms. Organs affected by the disease differ between patients, making diagnosis of affected organs a challenge. Although numerous studies have investigated the correlation between T cell antigen receptor (TCR) repertoires and development of infectious and immune diseases, the correlation between TCR repertoires and variations in disease symptoms among individuals remains unclear. This study investigated correlation of TCR repertoires in blood T cells with severity of an autoimmune phenotype that varies among individuals. We sequenced TCR repertoires in the blood and stomachs of mice deficient in autoimmune regulator (*Aire*) (AIRE KO), a mouse model of human autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED). Data analysis revealed that degree of similarity in TCR repertoires between the blood and stomach differs among individual AIRE KO mice and reflects the severity of stomach inflammation. This suggests the existence of blood TCR repertoires that can be used to predict affected organs. Indeed, we identified a set of TCR sequences whose frequencies in blood correlated with severity of stomach manifestations. Our results highlight the potential of using TCR repertoires not just for diagnosis of disease development, but also for diagnosis of affected organs in autoimmune diseases.

1. Introduction

Autoimmune diseases result from disrupted immune systems that trigger inflammation against various organs. Some autoimmune diseases present highly heterogeneous symptoms, as the severity and affected organs vary among patients. For example, systemic lupus erythematosus can affect various organs, including joints, skin, kidneys, heart, the hematopoietic system, and the nervous system, but with varying degrees of involvement among patients [1,2]. This individual variation is also seen in patients with APECED. APECED is caused by loss-of-function mutations in *AIRE*, a gene encoding the nuclear factor responsible for ectopic expression of peripheral-tissue antigens by thymic epithelial cells to mediate negative selection of self-reactive thymocytes in the thymus [3,4]. Patients are diagnosed as APECED by developing at least two symptoms among endocrine deficiency, ectodermal dystrophy, and chronic mucocutaneous candidiasis [5,6]. While some organs, such as adrenal and parathyroid glands, are constantly affected, others like thyroid glands and the pancreas are affected in only 2-12% of cases, and gastritis is seen in 13-15% of cases [5,6].

Mice deficient in *Aire* (AIRE KO) were previously established and showed autoimmune phenotypes.[3,7-14] Analysis of these mutant mice revealed that AIRE regulates thousands of tissue specific antigens (TSAs) in medullary thymic epithelial cells (mTECs). mTECs express and present these TSAs, thereby promoting negative selection of TSA-reactive T cells and generation of regulatory T cells in the thymus. Interestingly, as in APECED, some autoimmune phenotypes in AIRE KO mice vary among individuals [3,7,9,10], strains [9,11], and according to types of mutation [8,12-14]. In addition to deterministic factors such as genetic backgrounds, this variation may be ascribed partly to the stochastic nature of TSA expression regulated by AIRE, and to the inherent stochasticity that thymic T cells need to encounter mTECs, which are a rare population among total thymic cells, for thymic selection.

Because of the heterogeneous nature of autoimmune diseases, diagnosis of affected organs is necessary. In addition to the traditional approach by detecting organ-specific protein markers, it would be desirable to be able to use TCR repertoire analysis for diagnosis. TCR repertoire is formed by a combination of the hypervariable complementarity-determining region 3 (CDR3) loop of TCRs on their cell membranes and diverse V(D)J recombination in the thymus, allowing T cells to recognize myriad antigens in the body [15]. Since T cells play critical roles in development of autoimmune diseases by recognizing self-antigens, their TCR repertoires change depending on affected organs [16]. Therefore, diagnosis by detecting their TCR repertoires prior to disease exacerbation may provide an earlier window for disease diagnosis than a traditional approach detecting organ-specific protein markers. For example, by using peripheral blood samples from patients, multiple studies have shown altered utilization of V and J genes in patients with systemic lupus erythematosus, rheumatoid arthritis, and Sjogren's syndrome [17-19]. Moreover, Moore et al. have performed TCR repertoire analyses of spleen, salivary gland, and brain choroid plexus from an MRL/lpr mouse model of lupus and identified a set of TCR sequences significantly

detected in the brain [20]. However, despite growing evidence suggesting a correlation between TCR repertories and development of autoimmune diseases, the correlation between TCR repertoires and variations in disease severity of autoimmune diseases has rarely been studied.

In the present study, we performed TCRα and β sequencing in the blood and stomachs of Aire-deficient mice (AIRE KO), a mouse model of APECED, and tested the diagnostic potential of TCR repertoires to monitor individual variation in stomach symptoms in those mice. The data suggest that selected TCR repertoires in the blood correlate with autoimmune phenotypes in the stomach of AIRE KO mice, highlighting the potential of TCR sequence analysis in diagnosis of autoimmune diseases.

2. Materials and Methods

2.1 Mice

Female littermates were used in these experiments. B6.Cg-Aire<tm2Mmat>/Rbrc (RBC03515) mice were provided by the RIKEN RBC through the National Bio-Resource Project of MEXT in Japan. All mice were maintained under pathogen-free conditions and handled in accordance with Guidelines of the Institutional Animal Care and Use Committee of RIKEN, Yokohama Branch (2018-075).

2.2 Cell preparation and flow cytometry

Stomach tissue was excised and immediately washed with RPMI medium. It was then minced before digestion with 1.0mg/mL collagenase in RPMI medium (Wako) at 37°C. Supernatant containing cells was then passed through a 40-μm filter. Cells were isolated from blood by adding 2 mL of RBC lysis buffer (Biolegend) per 100 uL of blood collected. After incubation with RBC lysis buffer for 10-15 min at room temperature, cells were then washed with FACS buffer before antibody staining. For flow cytometry, cells were first incubated with blocking antibodies against Fc receptors (Biolegend) before staining with antibodies against TCRβ, CD4, CD8, CD44, and CD62L. Dead cells were excluded by staining with 7-aminoactinomycin D. Cells were sorted using a FACS Aria instrument (BD) for TCR sequencing.

2.3 TCR library preparation and sequencing

Total RNA was prepared using an RNeasy Micro Kit (QIAGEN) according to the manufacturer's protocol. A cDNA library for TCR was then prepared from RNA using a SMARTer Mouse TCR a/b Profiling Kit (Takara Bio) according to the manufacturer's protocol. To account for differences in numbers of sorted cells, RNA quantities were normalized prior to cDNA preparation. cDNA libraries of both TCRα and β were then subjected to paired-end sequencing using a Miseq (Illumina) according to the protocol for Miseq Reagent Kit v3 (Illumina).

2.4 TCR data processing and analysis

TCRα and β sequencing data were mapped using MiTCR software [21]. Output files from MiTCR were then converted to a format compatible with VDJtools, an analysis framework for repertoire sequencing data [22]. Diversity, clonality, CDR3 length, and significant V/J usage were calculated using VDJtools command lines. The Jaccard index of TCR repertoires between two samples was calculated by dividing the number of shared unique TCR sequences by the total number of unique TCR sequences. The distance matrix of selected TCR sequences was generated using the tcrdist3 python package [23], and hierarchical clustering was performed using the hclust function from the stats package in R.

2.5 Immunofluorescence

Rag1$^{-/-}$ tissues were snap-frozen in OCT compound. Tissues were sectioned with a cryostat (6 μm) and fixed in cold acetone for 10 min. After washing with PBS, sections were fixed in acetone and blocked with

10% normal goat serum. Sera from AIRE KO mice (100x) were then applied to the tissue sections and incubated for 1 h. Secondary anti-mouse antibody conjugated to Alexa-488 dye was then added along with 1 μg/mL propidium iodide (PI) and 5 μg/mL RNase A and incubated for 40 min. All images were captured using an SP8 confocal microscope (Leica).

2.6 Statistics

Statistical analysis employed Graphpad Prism 9 software. Student's t-test was used to compare means between pairs of groups, and two-way ANOVA was used to compare two groups with two independent variables. A p value less than 0.05 was considered significant. All outliers were included in the analysis.

3. Results

3.1 Individual variation in phenotypes of AIRE KO mice

Our aim in this study was to test the possibility of TCR repertoire analysis for organ-specific diagnosis of autoimmune manifestations in a murine model. As a murine model of autoimmune diseases, we chose gastric autoimmunity of AIRE KO mice because: 1) the stomach is large enough to obtain sufficient numbers of T cells for TCR sequencing and 2) autoimmune inflammation in the stomach is heterogeneously induced in AIRE KO mice [3,9].

We first confirmed the heterogeneous features of autoimmune phenotypes in the stomachs of AIRE KO mice. Inflammatory status of the stomachs was evaluated for AIRE KO mice approximately 30 weeks old and their wild-type littermates (AIRE WT) as controls. Flow cytometric analysis indicated that the proportion of $CD44^+$ $CD62L^-$ activated effector memory T cells (Teffs) among $CD4^+$ and $CD8^+$ T cells was significantly increased, and cell numbers were significantly increased for $CD8^+$ Teffs and showed an increasing trend for $CD4^+$ Teffs in stomachs of AIRE KO mice compared to controls, manifesting stomach inflammation with activated T cell infiltration in mutant mice (Figure 1a-b). We further evaluated generation of autoantibodies against stomach tissues by immunohistochemical staining of the stomach section from RAG1-deficient (*Rag1*$^{-/-}$) mice with serum of AIRE KO mice (Figure 1c, left). Quantification of these immunofluorescent images showed that the level of autoantibodies to the stomach was significantly higher in the serum of AIRE KO mice than that of AIRE WT mice (Figure 1c, right), which is consistent with previous studies. Notably, standard deviations (SD) in the number of Teffs and serum levels of anti-stomach antibodies are higher in AIRE KO mice than in controls (Figure1a-c, Standard deviation (SD) = 27780 and 7076 each for the $CD4^+$ Teff number and SD = 17061 and 614 for the $CD8^+$ Teff number in AIRE KO and control mice, respectively; SD = 19.3 and 4.17 each for the serum antibody level in AIRE KO and control mice, respectively). Thus, there was large variation among individual AIRE KO mice in the number of infiltrated Teffs in the stomach and serum levels of anti-stomach antibodies. Consequently, significant individual variation in stomach manifestations was confirmed in this line of AIRE KO mice.

If this line of AIRE KO is utilized to validate the possibility of TCR repertoire analysis for a stomach-specific diagnosis, it is necessary to confirm that individual differences in the degree of autoimmune symptoms in the stomach are not the same as in other organs. We therefore investigated other target organs of autoimmunity of the AIRE dysfunction. Titers of autoantibodies to various organs in the serum were quantified by immunohistochemical staining. These results showed that the profile of targeted organs varied markedly among individual mice (Figure 1d-e). Mouse #1 was the most severely affected, as it produced the broadest profile with the greatest number of organs having high autoantibody signal intensity. On the other hand, Mouse #6 showed a sporadic profile with the highest signal intensity for the stomach, lung, and pancreas. Mice #3, #4, and #5 also produced sporadic profiles with relatively high

intensity for organs such as skin, eyes, and salivary glands. Finally, Mouse #2 produced the smallest signal in the greatest number of organs examined. We also evaluated the correlation of autoantibody signal intensity of stomach with that of other organs. While organs such as large intestine, pancreas, and lung showed high correlation, skin, eyes, and ovaries showed low correlation with stomachs, suggesting that titers of serum autoantibodies against each tissue vary among individuals in these AIRE KO mice (Figure 1f). Taken together, these results indicate that autoimmune phenotypes of AIRE KO mice used in this study are indeed heterogeneous. Accordingly, we concluded that AIRE KO mice comprise a useful model to assess the disease-monitoring potential of TCR repertoires.

3.2 Limited correlation of blood T cell count, TCR diversity, clonality, CDR3 length, and V/J usage with disease severity in AIRE KO mice

Before performing TCR sequencing, we asked whether blood T cell counts might reflect the severity of disease in AIRE KO mice. To this end, we used flow cytometry to detect $CD4^+$ and $CD8^+$ Teffs in the blood of AIRE KO and AIRE WT mice. We found that both the proportion and the number of $CD4^+$ and $CD8^+$ Teffs were comparable in AIRE KO and AIRE WT mice (Figure 2a-b). In addition, we found a weak correlation between the number of blood $CD4^+$ (Rs = 0.37) or $CD8^+$ Teffs (Rs = 0.37) with the mean signal intensity of stomach immunofluorescence (Figure 2c), a measure of the severity of stomach manifestations. Overall, the Teff count in blood does not reflect disease severity in stomachs of AIRE KO mice.

To look for blood TCR repertoires that might correlate with stomach inflammation, we determined TCR repertories of stomachs and blood of AIRE KO mice and searched for common TCR repertoires in the blood and stomach because stomach Teffs are expected to circulate more in the blood of mice with severe stomach inflammation. We then subjected 6 AIRE KO (labeled from #1-6) and 7 AIRE WT (labeled from #7-13) mice, which were used in flow cytometry and immunofluorescence studies, to TCRα and β sequencing. In particular, we chose $CD4^+$ Teffs in the blood (Blood KO) and stomach (Stomach KO) of AIRE KO mice and in the blood (Blood WT) of AIRE WT for TCR sequencing (Table 1). Approximately 5000 to 15000 $CD4^+$ Teffs were FACS-sorted from the blood. For the stomach, all $CD4^+$ Teffs in the organ were sorted for library preparation. The total number of $CD4^+$ Teffs obtained from the stomach (hereafter referred to as stomach Teff count) ranged from 2000 to 40000 per stomach and showed a strong correlation with the mean signal intensity of stomach immunofluorescence (Rs = 0.77). Thus, both the stomach Teff count and the mean signal intensity of stomach immunofluorescence serve as a measure of the severity of stomach manifestations in AIRE KO mice (Figure 2d).

We performed TCR repertoire sequencing of these sorted $CD4^+$ Teffs. To consider the difference in cell number, the amount of RNA extracted from each sample was normalized for cDNA library preparation. By analyzing sequencing data, some parameters of TCRα and β repertoires such as diversity, clonality,

CDR3 length, and V/J usage, were determined in each sample. We then sought to find parameters that predict disease onset in the stomachs of these AIRE KO mice according to two criteria: 1) the difference between Blood KO and Blood WT, for prediction by measurement of blood TCR repertoires, and 2) the similarity between Blood KO and Stomach KO, for supporting the reflection of stomach manifestations.

First, we calculated the number of unique clonotypes and normalized Shannon indices to evaluate diversity and clonality, respectively, of Blood WT, Blood KO, and Stomach KO. We found that diversity among Blood WT, Blood KO, and Stomach KO was comparable for both TCRα and β (Figure 3a). On the other hand, we found that TCRα clonality was significantly increased in Stomach KO while TCRβ clonality was comparable among all groups (Figure 3a). Similar to clonality, the amino acid length of CDR3 was significantly decreased in Stomach KO for TCRα and was comparable among all groups for TCRβ (Figure 3b). In support of the trend of clonality and CDR3 length, most significantly detected V and J genes were strongly present in Stomach KO for TCRα, but were comparably present in all groups for TCRβ (Figure 3c-e). Since diversity, clonality, CDR3 length, and V/J usage for TCRα and β repertoires did not meet criteria for correlation, we concluded that these measures do not reflect disease severity in AIRE KO.

3.3 Individual similarities of TCR repertoires in the blood and stomachs of AIRE KO

We then shifted our attention to individual TCR sequences in TCR repertoires and focused on shared TCR sequences in the blood and stomach from the same AIRE KO mice. To this end, we calculated Jaccard indices to evaluate the similarity of TCR repertoires based on the proportion of shared TCR sequences between each paired blood (Blood#1-13) and stomach sample (Stomach #1-6). Scatterplots and heat maps of Jaccard indices suggested that Blood KO correlated with the corresponding Stomach KO (Figure 4a and b). Indeed, Jaccard indices of the blood and stomach from the same AIRE KO mice (on average about 4%) were four times higher than those of blood and stomachs from different mice (on average about 1%) for both TCRα and β (Figure 4c). Thus, the blood TCR repertoire correlates with that of the stomach repertoire of AIRE KO mice, and more importantly, the correlation of the TCR repertoire between stomach and blood is individualized. Notably, scaled Jaccard indices correlated with the stomach Teff count, a measure reflecting the severity of the stomach manifestation (Figure 4d). Thus, mice with more severe gastric inflammation have more common TCR repertoires in blood and stomach, implying the high circulation of stomach Teff in the blood. Overall, these data support the potential for monitoring the severity of stomach manifestation by detecting TCR sequences common to the blood and stomach in AIRE KO.

3.4 Identification of candidate TCR repertoires to monitor disease severity in AIRE KO

In order to further confirm that the blood TCR repertoire reflects stomach manifestations, we sought to determine actual sequences of these TCR repertoires. Such TCR repertoires could monitor the severity of stomach manifestations in AIRE KO mice as a biomarker. To this end, we first screened the CDR3 of

TCR sequences according to three criteria: 1) CDR3 sequences that were detected in both the Stomach KO and Blood KO of the same mice, 2) were found in at least two AIRE KO mice and 3) were not present in Blood WT. As a result, we successfully identified 29 CDR3 sequences for both TCRα and β. We then categorized these CDR3 sequences based on their amino acid sequences using the tcrdist3 package to generate a distance matrix, which was then used for hierarchical clustering [23]. Based on the clustering result, we categorized identified CDR3 amino acid sequences into 4 clusters (k-means = 4) for TCRα (Figure 5a) and 3 clusters (k-means = 3) for TCRβ (Figure 5b).

Since a biomarker should discriminate between diseased and non-diseased subjects, we searched for clusters containing CDR3 sequences whose frequencies from the blood and stomach correlated with the severity of stomach manifestations, represented by the stomach Teff count. We found that the mean frequencies of CDR3 sequences classified as Cluster3 in TCRα and classified as Cluster1 in TCRβ (hereafter referred to as Cluster3/TCRα and Cluster1/TCRβ, respectively) showed some positive correlation with both the blood and stomach (Rs > 0.5) (Figure 5c-d). Stomach TCR frequencies of Cluster 3/TCRα and Cluster1/TCRβ were able to discriminate the two mice with the highest stomach Teff counts (Mouse #1 and Mouse #6) from the other mice. However, blood TCR frequencies of these clusters were not clearly correlated with stomach Teff counts (Figure 5c-f). Especially for Cluster1/TCRβ, high TCR frequencies in the stomach of Mouse #6 were not reflected in the blood (Figure 5d, f). Therefore, we focused on Cluster3/TCRα and found that when we divided Cluster3 into two sub-clusters, Clusters 3.1 and 3.2 (Figure 5g), TCR frequencies of Cluster 3.1 were relatively high in the blood of Mouse #2 (Figure 5e, h, and Table 2), a mouse with a low stomach Teff count (Figure 5h). By contrast, TCR frequencies of Cluster 3.2 showed a relatively strong correlation with stomach Teff counts for both the blood (Rs = 0.64) and stomach (Rs = 0.64) (Figure 5h). Therefore, we used the TCR frequencies of Cluster 3.2 from the blood or stomach and finally confirmed their correlations with two distinct dimensions reflecting the severity of stomach manifestations, stomach Teff count and mean signal intensity in immunofluorescent staining of the stomach (Figure 5i). The result of 3D plots showed that by using TCR frequencies of Cluster 3.2 from both the blood and stomach, Mouse #1 and Mouse #6 are clearly distinguished from the others. Consequently, data analysis suggests that these TCR repertoires in the blood can monitor autoimmune manifestations in the stomachs of AIRE KO mice.

4. Discussion

In the present study, we used an AIRE KO mouse model of APECED to test whether TCR repertoires can serve as biomarkers to monitor organ manifestations in autoimmune disease. We identified a set of TCR sequences in Cluster 3.2 of TCRα that discriminate mice with relatively severe phenotypes from other mice. These data support the potential of TCR repertoires to monitor disease severity, at least in this AIRE KO mouse model of autoimmune disease.

Results of TCR sequencing showed that the clonality, CDR3 length, and V/J usage of Stomach KO of TCRα repertoires were significantly altered compared to those of Blood WT and Blood KO. This suggests that TCRs responsive to stomach-derived antigens may be more abundant in the TCRα repertoire than in the TCRβ repertoire. In addition, we found no significant difference in diversity, clonality, CDR3 length, and V/J usage between Blood WT and Blood KO. This is consistent with a previous study showing that diversity of the TCR repertoire is comparable between the blood of AIRE WT and AIRE KO mice [24].

To date, most analyses of TCR repertoire in autoimmune diseases have been conducted to monitor disease development rather than variation in disease manifestations [16]. Most previous studies simply compared TCR repertoires between patients and normal individuals and searched for common TCR sequences that were abundant in patients [17-19]. In this study, we went a step further by investigating differences in TCR repertoires among diseased samples to address the potential of TCR repertoires to dissect disease heterogeneity.

Despite identification of candidate TCR repertoires to monitor stomach manifestation in AIRE KO mice, there is no guarantee that these TCR repertoires specifically recognize a self-antigen derived from the stomach. To address this issue, it is important to determine TCR repertoire specificities against stomach-derived antigens. In fact, we could not find potential target epitopes of the identified TCR sequences in Cluster 3.2 by searching VDJdb, a curated database of TCRs with known antigen specificities [25]. Previously, MUCIN6 has been reported as a self-antigen involved in development of autoimmunity in AIRE KO mice [26], and it is of interest to know whether identified TCR repertories can respond to MUCIN6. However, determining the antigen specificity of TCRs based on current experimental approaches is time-consuming, and there is a significant lack of data on known TCR-peptide pairs to develop a reliable algorithm for predicting antigen specificity of any TCR [27]. Therefore, invention of a high-throughput experimental approach that allows simultaneous identification of myriad TCR-peptide pairs is awaited for accurate prediction of tissue specificity of TCR repertoires.

The advantage of using a mouse model of autoimmune disease is that unlike human samples, it is easy to obtain samples from blood and peripheral organs. Especially for the purpose of this study, it would be very difficult to obtain sufficient T cells from stomachs of patients with autoimmune diseases for TCR

repertoire analysis. However, TCR sequences of mice and humans are different. Clearly, there is a great need for bioinformatics tools that can predict human TCR sequences from those of mice, so that we can seamlessly translate findings from animal models to the clinic.

Author contributions

Tatsuya Ishikawa: Conceptualization, Writing - Original Draft, Investigation, Formal analysis, Funding acquisition, Visualization **Kenta Horie**: Investigation, Formal analysis **Yuki Takakura**: Investigation **Houko Ohki**: Resources **Yuya Maruyama:** Resources, Writing - Review & Editing **Mio Hayama**: Resources, Writing - Review & Editing **Maki Miyauchi**: Resources, Methodology **Takahisa Miyao**: Methodology, Writing - Review & Editing **Nobuko Akiyama**: Project administration, Writing - Review & Editing, Funding acquisition **Taishin Akiyama**: Conceptualization, Supervision, Writing - Original Draft, Project administration, Funding acquisition

Disclosure of conflicts of interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by Grants-in-Aid for Scientific Research from JSPS (20K07332, 20H03441) (TA, NA), and CREST from Japan Science and Technology Agency (JPMJCR2011) (TA, TI). TI is supported by the establishment of university fellowships towards the creation of science technology innovation from Japan Science and Technology Agency (JPMJFS2140) and RIKEN Junior Associate Program. Computations were performed on the NIG supercomputer at ROIS, National Institute of Genetics and HOKUSAI supercomputer at ISD, RIKEN.

References

1. Wu, H. et al. Organ-specific biomarkers in lupus. *Autoimmun Rev* **16**, 391-397 (2017).
2. Allen, M.E., Rus, V. & Szeto, G.L. Leveraging Heterogeneity in Systemic Lupus Erythematosus for New Therapies. *Trends Mol Med* **27**, 152-171 (2021).
3. Anderson, M.S. et al. Projection of an immunological self shadow within the thymus by the aire protein. *Science* **298**, 1395-401 (2002).
4. Perniola, R. Twenty Years of AIRE. *Front Immunol* **9**, 98 (2018).
5. Perheentupa, J. Autoimmune polyendocrinopathy--candidiasis--ectodermal dystrophy (APECED). *Horm Metab Res* **28**, 353-6 (1996).
6. Betterle, C., Greggio, N.A. & Volpato, M. Clinical review 93: Autoimmune polyglandular syndrome type 1. *J Clin Endocrinol Metab* **83**, 1049-55 (1998).
7. Ramsey, C. et al. Aire deficient mice develop multiple features of APECED phenotype and show altered immune response. *Hum Mol Genet* **11**, 397-409 (2002).
8. Kuroda, N. et al. Development of autoimmunity against transcriptionally unrepressed target antigen in the thymus of Aire-deficient mice. *J Immunol* **174**, 1862-70 (2005).
9. Jiang, W., Anderson, M.S., Bronson, R., Mathis, D. & Benoist, C. Modifier loci condition autoimmunity provoked by Aire deficiency. *J Exp Med* **202**, 805-15 (2005).
10. Venanzi, E.S., Melamed, R., Mathis, D. & Benoist, C. The variable immunological self: genetic variation and nongenetic noise in Aire-regulated transcription. *Proc Natl Acad Sci U S A* **105**, 15860-5 (2008).
11. Niki, S. et al. Alteration of intra-pancreatic target-organ specificity by abrogation of Aire in NOD mice. *J Clin Invest* **116**, 1292-301 (2006).
12. Su, M.A. et al. Mechanisms of an autoimmunity syndrome in mice caused by a dominant mutation in Aire. *J Clin Invest* **118**, 1712-26 (2008).
13. Hubert, F.X. et al. Aire-deficient C57BL/6 mice mimicking the common human 13-base pair deletion mutation present with only a mild autoimmune phenotype. *J Immunol* **182**, 3902-18 (2009).
14. Goldfarb, Y. et al. Mechanistic dissection of dominant AIRE mutations in mouse models reveals AIRE autoregulation. *J Exp Med* **218**(2021).
15. Schatz, D.G. & Ji, Y. Recombination centres and the orchestration of V(D)J recombination. *Nat Rev Immunol* **11**, 251-63 (2011).
16. Mitchell, A.M. & Michels, A.W. T cell receptor sequencing in autoimmunity. *J Life Sci (Westlake Village)* **2**, 38-58 (2020).
17. Liu, X. et al. T cell receptor β repertoires as novel diagnostic markers for systemic lupus erythematosus and rheumatoid arthritis. *Ann Rheum Dis* **78**, 1070-1078 (2019).
18. Ye, X. et al. High-Throughput Sequencing-Based Analysis of T Cell Repertoire in Lupus Nephritis. *Front Immunol* **11**, 1618 (2020).
19. Lu, C. et al. Clinical significance of T cell receptor repertoire in primary Sjogren's syndrome. *EBioMedicine* **84**, 104252 (2022).
20. Moore, E. et al. The T Cell Receptor Repertoire in Neuropsychiatric Systemic Lupus Erythematosus. *Front Immunol* **11**, 1476 (2020).
21. Bolotin, D.A. et al. MiTCR: software for T-cell receptor sequencing data analysis. *Nat Methods* **10**, 813-4 (2013).
22. Shugay, M. et al. VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. *PLoS Comput Biol* **11**, e1004503 (2015).
23. Dash, P. et al. Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature* **547**, 89-93 (2017).
24. Oftedal, B.E. et al. T cell receptor assessment in autoimmune disease requires access to the most adjacent immunologically active organ. *J Autoimmun* **81**, 24-33 (2017).
25. Shugay, M. et al. VDJdb: a curated database of T-cell receptor sequences with known antigen specificity. *Nucleic Acids Res* **46**, D419-d427 (2018).
26. Gavanescu, I., Kessler, B., Ploegh, H., Benoist, C. & Mathis, D. Loss of Aire-dependent thymic expression of a peripheral tissue antigen renders it a target of autoimmunity. *Proc Natl Acad Sci U S A* **104**, 4583-7 (2007).

27. Hudson, D., Fernandes, R.A., Basham, M., Ogg, G. & Koohy, H. Can we predict T cell specificity with digital biology and machine learning? *Nat Rev Immunol*, 1-11 (2023).

Figure1

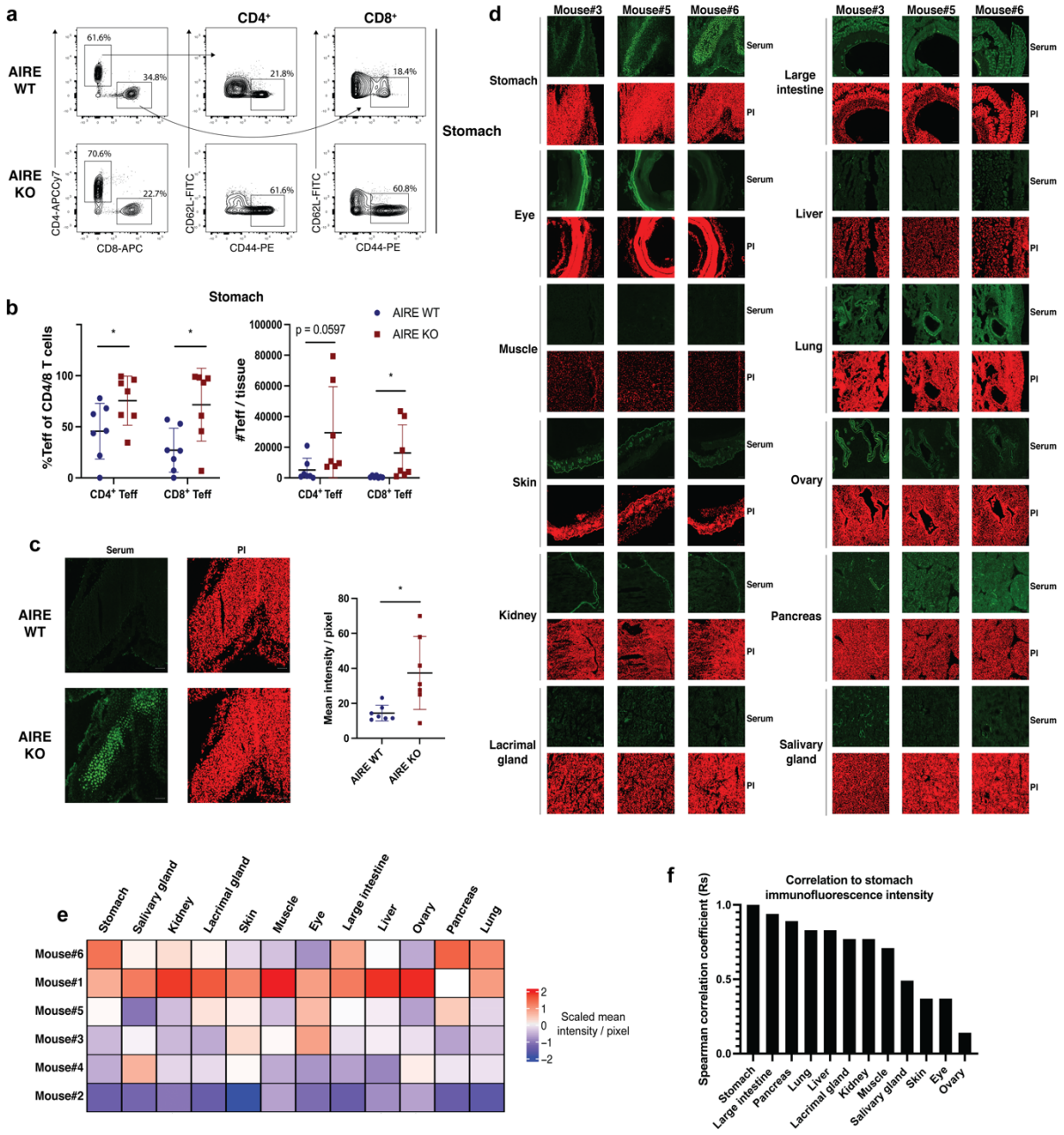

Figure 1. **Individual variation in multi-organ phenotypes of AIRE KO mice**. (a) Flow cytometric analysis of CD4$^+$ and CD8$^+$ Teffs from stomachs of AIRE WT and AIRE KO mice. (b) Numbers and percentages of CD4$^+$ and CD8$^+$ Teffs from stomachs of AIRE WT and AIRE KO mice. (c) Immunofluorescent staining of AIRE WT and AIRE KO sera against stomach sections of Rag1$^{-/-}$ mice. (d) Immunofluorescent staining of sera from AIRE KO mice against different tissue sections from Rag1$^{-/-}$ mice. (e) Heatmap of scaled mean intensity of immunofluorescent signals. (f) Spearman correlation

coefficients (Rs) between stomach immunofluorescence signals and those of other organs. Samples were analyzed using unpaired, two-tailed Student's t-test. ns, not significant ($p > 0.05$). *$p < 0.05$.

Figure 2

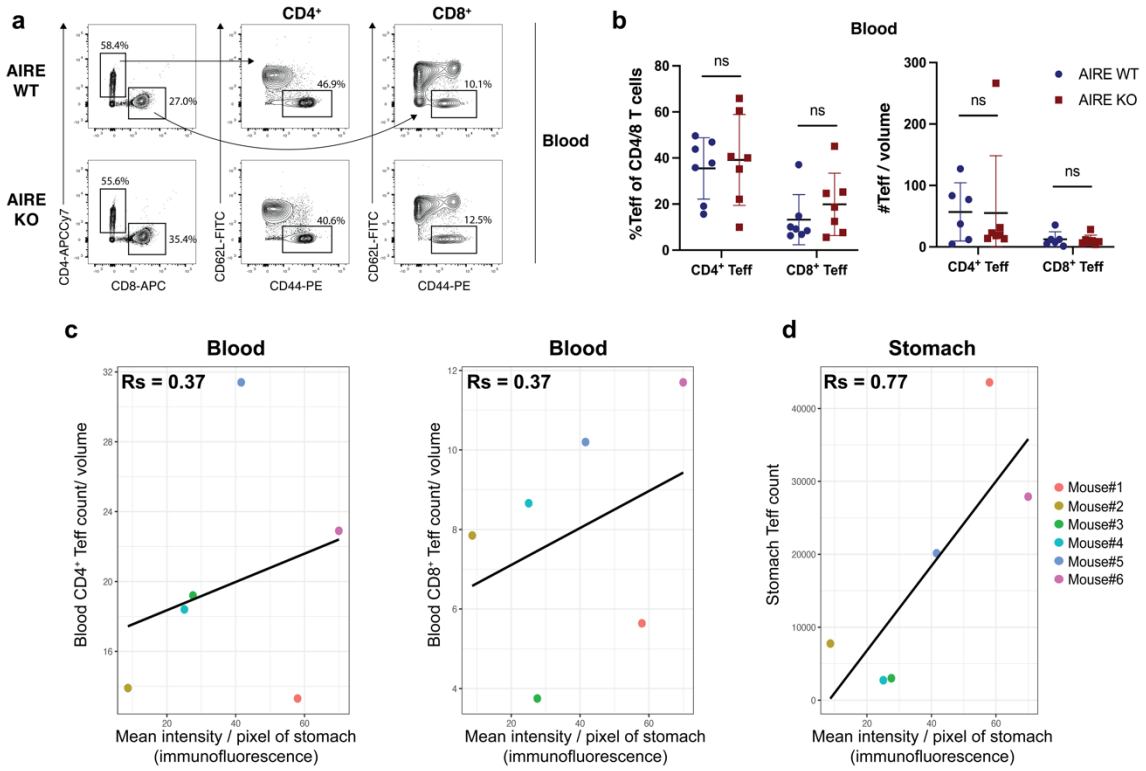

Figure 2. **Correlation of activate T cell number in the stomach and blood with titer of anti-stomach antibody in AIRE KO mice**. (a) Flow cytometric analysis of $CD4^+$ and $CD8^+$ Teffs from the blood of AIRE WT and AIRE KO mice. (b) Numbers and percentages of $CD4^+$ and $CD8^+$ Teffs from the blood of AIRE WT and AIRE KO mice. (c) Spearman's correlation of blood $CD4^+$ Teffs (left) or $CD8^+$ Teffs (right) with the mean signal intensity of stomach immunofluorescence. (d) Spearman's correlation of stomach $CD4^+$ Teff count with mean signal intensity of stomach immunofluorescence, reflecting the titer of anti-stomach antibody. Samples were analyzed by unpaired two-tailed Student's t-test. ns, not significant ($p > 0.05$). Rs, Spearman's correlation coefficient.

Figure 3

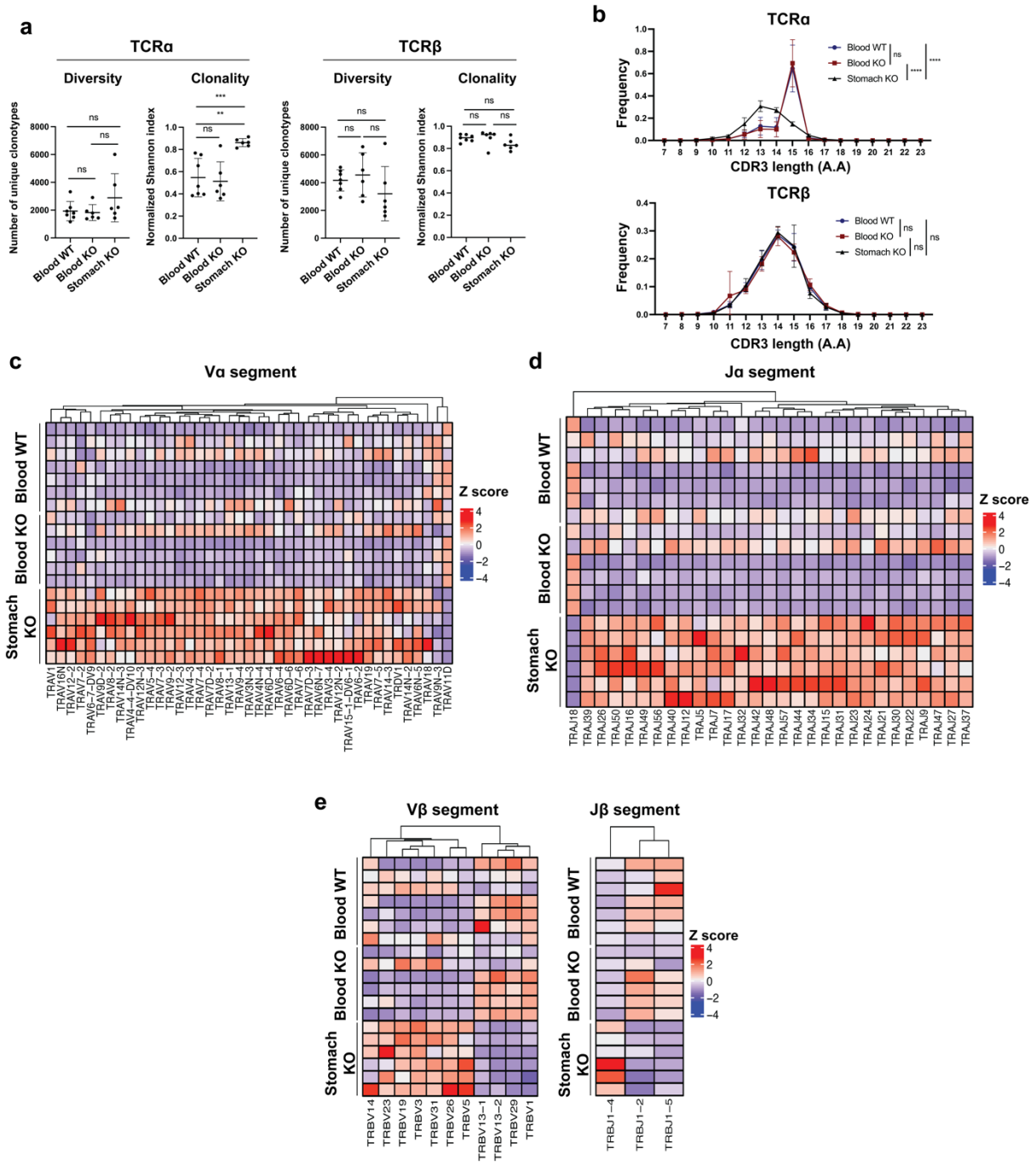

Figure 3. **TCR repertoire analysis of blood and stomachs of AIRE KO mice**. (a) Diversity and clonality of TCRα and β of Blood WT, Blood KO, and Stomach KO. (b) CDR3 amino acid length distributions of TCRα and β of Blood WT, Blood KO, and Stomach KO. Two-way ANOVA with random variables. ****$p < 0.0001$. ns, not significant ($p > 0.05$). Heatmap showing scaled frequencies of significantly detected Vα

(c), Jα (d), Vβ, and Jβ (e) genes between Blood WT and Blood KO and between Stomach KO and Blood WT/KO.

Figure 4

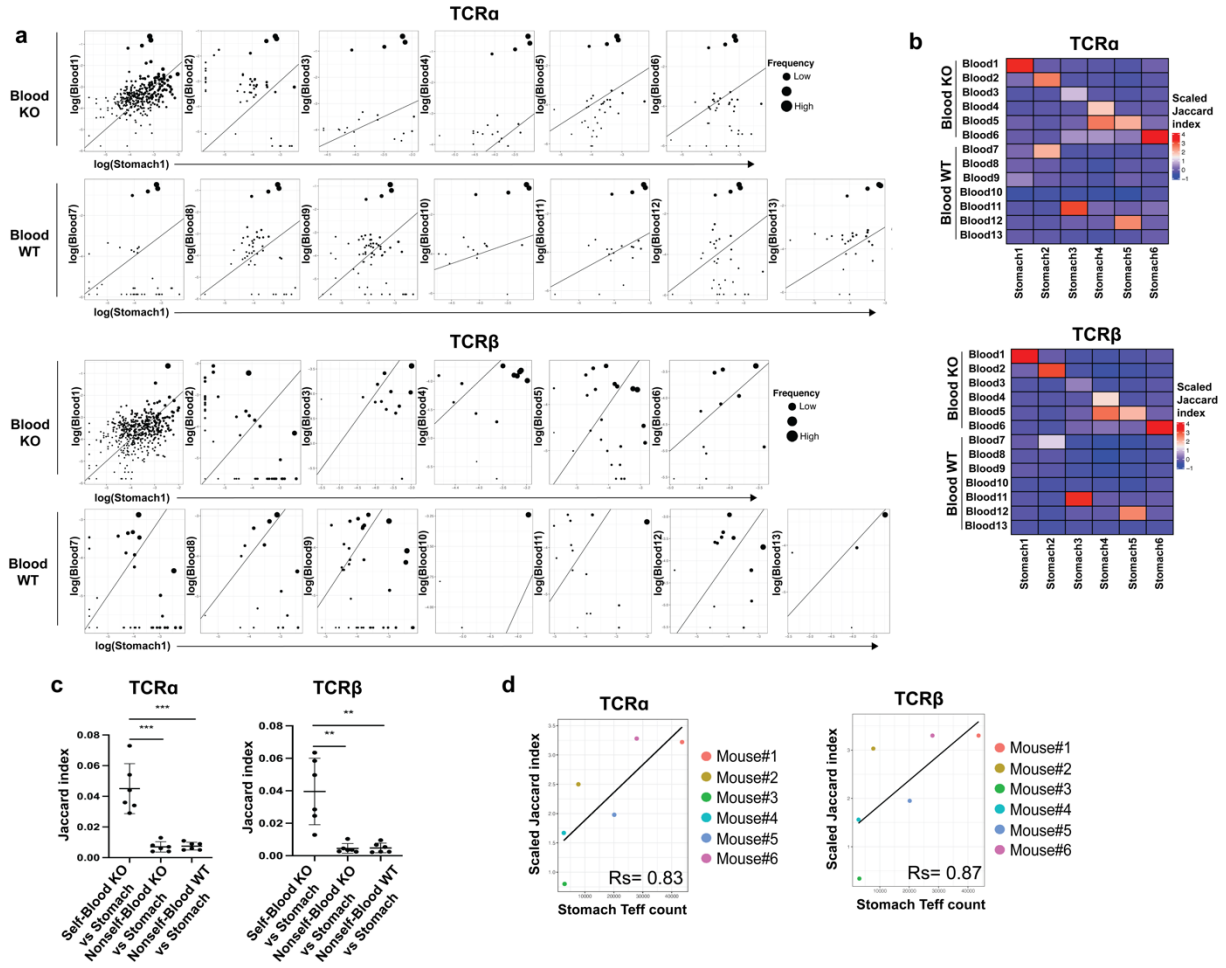

Figure 4. **Individual similarity of TCR repertoires between blood and stomachs of AIRE KO**. (a) Scatterplot showing frequencies of TCRs shared between Stomach1 and each of the blood samples. (b) Heatmap of scaled Jaccard indices of each of blood and stomach pair. (c) Jaccard indices of the stomach and blood from the same AIRE KO mice, average Jaccard indices of the stomach and blood from different AIRE KO mice, and average Jaccard indices of the stomach and blood from AIRE WT mice. Unpaired two-tailed Student's t-test. **$p < 0.01$, ***$p < 0.001$. (d) Spearman's correlation between stomach Teff counts and scaled Jaccard indices of the blood and stomach from each AIRE KO mouse. Rs, Spearman's correlation coefficient.

Figure 5

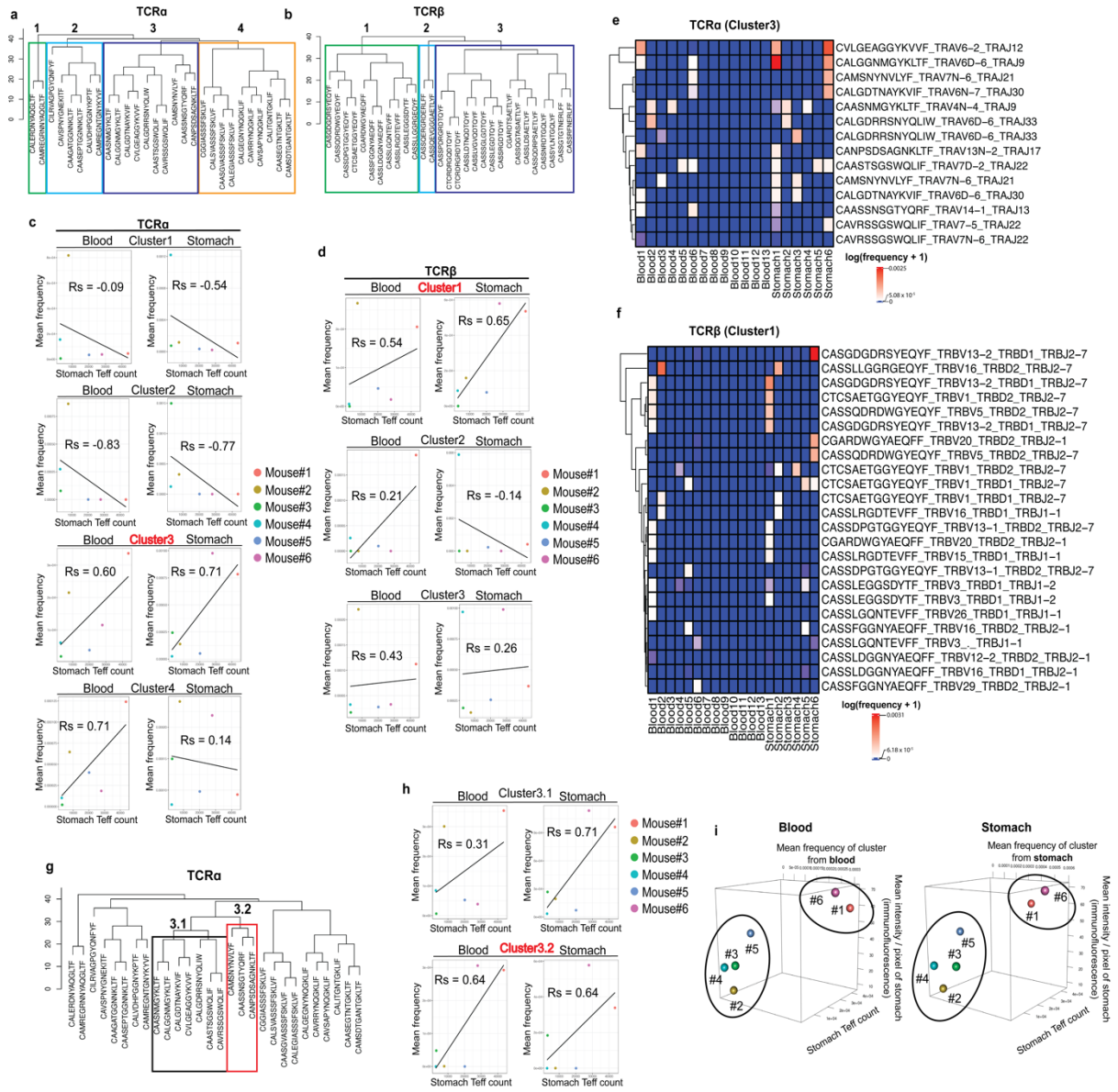

Figure 5. **Identification of candidate TCR repertoires correlating with severity of stomach manifestation in AIRE KO.** Hierarchical clustering of candidate TCRα (a) and β (b) repertories. Cluster numbers are shown at the tops of trees. Spearman's correlation of stomach Teff counts and mean frequencies of TCRα (c) and β (d) sequences contained in each cluster detected in the blood (left) and stomach (right) of each AIRE KO mouse. Heatmap showing log-transformed frequencies of TCRα sequences contained in Cluster3 (e) and TCRβ sequences contained in Cluster1 (f). Different CDR3 nucleotide sequences encoding the same CDR3 amino acid sequences are also shown in the heatmap. (g) Hierarchical clustering tree showing subclusters of Cluster3 of TCRα (Cluster3.1 and 3.2). (h) Spearman's correlation of stomach Teff counts and mean frequencies of TCRα sequences in Clusters 3.1 (top) and 3.2 (bottom) detected in the blood (left) and stomach (right) of each AIRE KO mouse. (i) 3-

dimensional plots of mean frequencies of TCRα sequences in Cluster 3.2 detected in the blood (left) or stomach (right), stomach Teff counts, and mean signal intensities of stomach immunofluorescence. Rs, Spearman's correlation coefficient.