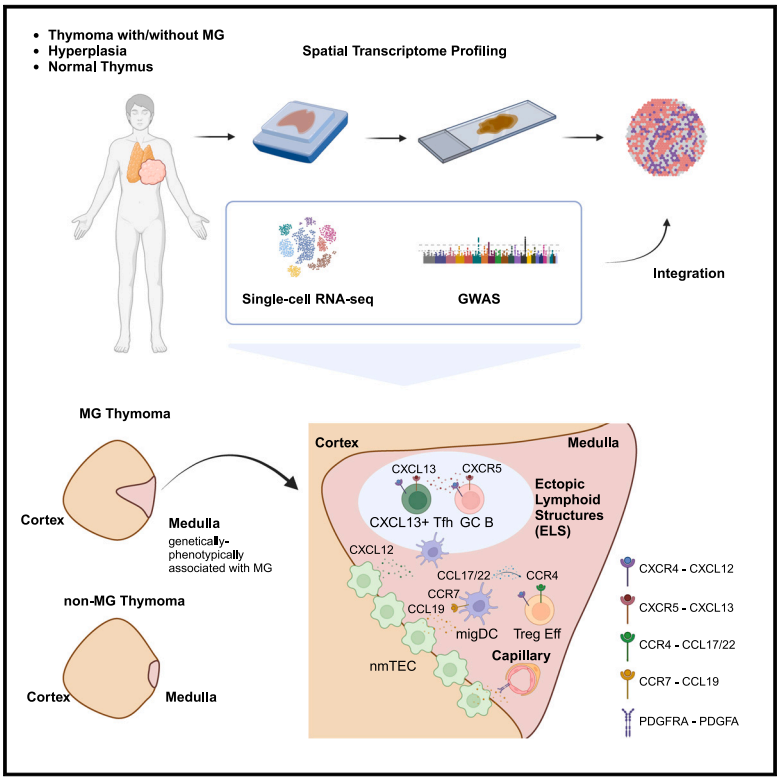
Cell Reports

Resource

Spatial transcriptomics elucidates medulla niche supporting germinal center response in myasthenia gravis-associated thymoma

Graphical abstract



Highlights

- Spatial transcriptomics reveals MG-specific features in the thymus
- Multimodal data integration highlights disease association in the medulla
- We identify specific immune niches in the medulla and nmTEC enrichment in the junction

Authors

Yoshiaki Yasumizu, Makoto Kinoshita, Martin Jinye Zhang, ..., Eiichi Morii, Tatsusada Okuno, Hideki Mochizuki

Correspondence

yoshiaki.yasumizu@yale.edu (Y.Y.), okuno@neurol.med.osaka-u.ac.jp (T.O.)

In brief

Yasumizu et al. identify MG-specific features in thymoma and thymic hyperplasia using spatial transcriptomics. Combining single-cell RNA-seq and GWAS data, they find MG susceptibility enriched in the medulla region of thymoma, characterized by specific immune cells. These findings provide a foundation for future MG research.

Yasumizu et al., 2024, Cell Reports *43*, 114677
September 24, 2024
https://doi.org/10.1016/j.celrep.2024.114677

CellPress

Cell Reports

CellPress
OPEN ACCESS

Resource

Spatial transcriptomics elucidates medulla niche supporting germinal center response in myasthenia gravis-associated thymoma

Yoshiaki Yasumizu,[1,2,3,4,13,*] Makoto Kinoshita,[1] Martin Jinye Zhang,[5,6] Daisuke Motooka,[3,7] Koichiro Suzuki,[8,9] Satoshi Nojima,[10] Naoshi Koizumi,[1] Daisuke Okuzaki,[3,7] Soichiro Funaki,[11] Yasushi Shintani,[11] Naganari Ohkura,[2,12] Eiichi Morii,[10] Tatsusada Okuno,[1,*] and Hideki Mochizuki[1,3]

[1]Department of Neurology, Graduate School of Medicine, Osaka University, Suita, Osaka, Japan
[2]Department of Experimental Immunology, Immunology Frontier Research Center, Osaka University, Suita, Osaka, Japan
[3]Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives (OTRI), Osaka University, Suita, Osaka, Japan
[4]Department of Neurology, Yale School of Medicine, New Haven, CT, USA
[5]Ray and Stephanie Lane Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA
[6]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[7]Genome Information Research Center, Research Institute for Microbial Diseases, Osaka University, Suita, Osaka, Japan
[8]BIKEN-RIMD NGS Laboratory, Research Institute for Microbial Diseases, Osaka University, Suita, Japan
[9]Biomedical Science Center, The Research Foundation for Microbial Diseases of Osaka University (BIKEN), Suita, Japan
[10]Department of Pathology, Graduate School of Medicine, Osaka University, Suita, Osaka, Japan
[11]Department of General Thoracic Surgery, Graduate School of Medicine, Osaka University, Suita, Osaka, Japan
[12]Department of Frontier Research in Tumor Immunology, Graduate School of Medicine, Osaka University, Suita, Osaka, Japan
[13]Lead contact
*Correspondence: yoshiaki.yasumizu@yale.edu (Y.Y.), okuno@neurol.med.osaka-u.ac.jp (T.O.)
https://doi.org/10.1016/j.celrep.2024.114677

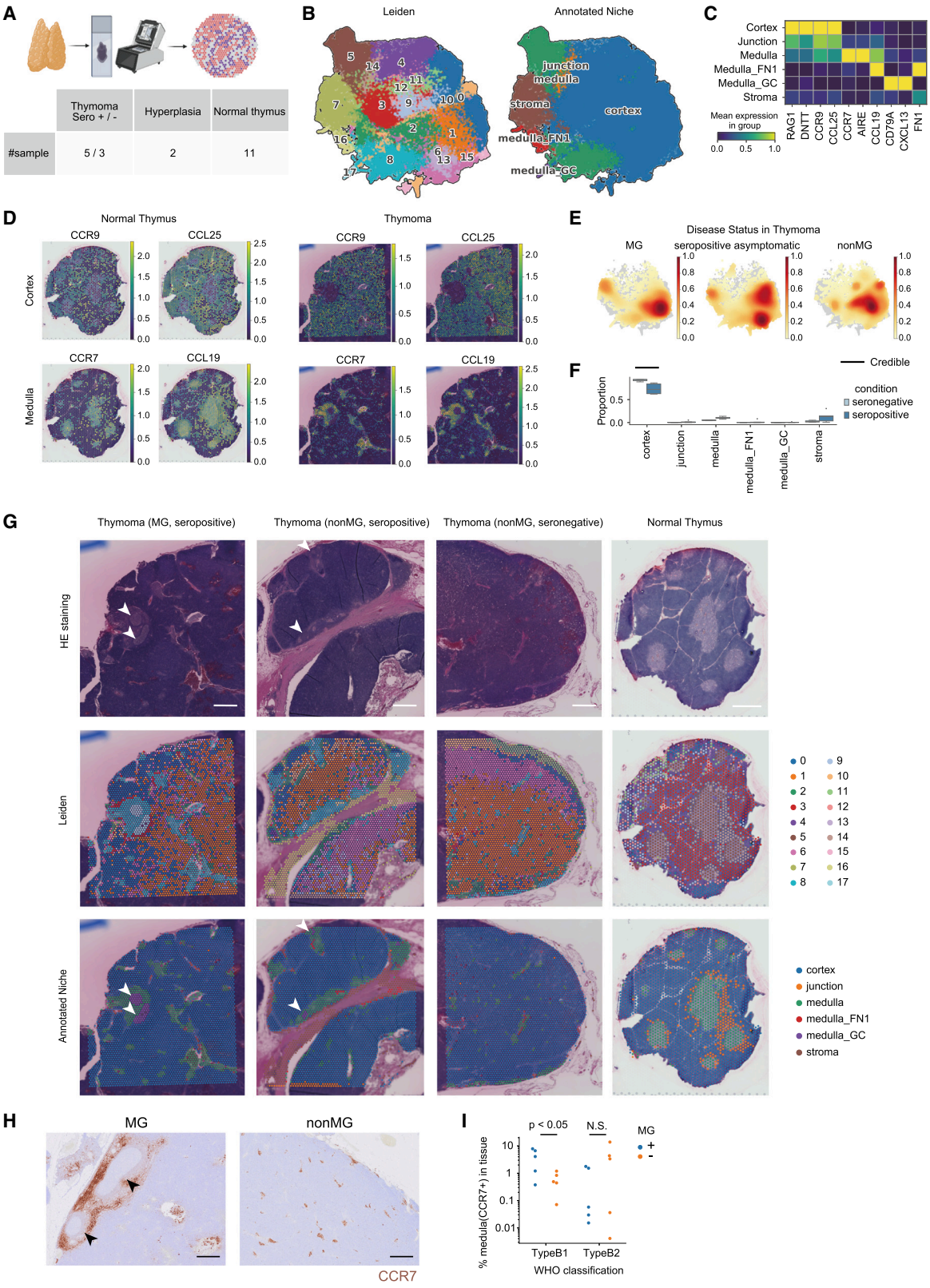
SUMMARY

Myasthenia gravis (MG) is etiologically associated with thymus abnormalities, but its pathology in the thymus remains unclear. In this study, we attempt to narrow down the features associated with MG using spatial transcriptome analysis of thymoma and thymic hyperplasia samples. We find that the majority of thymomas are constituted by the cortical region. However, the small medullary region is enlarged in seropositive thymomas and contains polygenic enrichment and MG-specific germinal center structures. Neuromuscular medullary thymic epithelial cells, previously identified as MG-specific autoantigen-producing cells, are enriched in the cortico-medullary junction. The medulla is characterized by a specific chemokine pattern and immune cell composition, including migratory dendritic cells and effector regulatory T cells. Similar germinal center structures and immune microenvironments are also observed in the thymic hyperplasia medulla. This study shows that the medulla and junction areas are linked to MG pathology and provides insights into future MG research.

INTRODUCTION

Myasthenia gravis (MG) is an autoimmune disease that causes systemic muscle weakness due to the production of autoantibodies that target the neuromuscular junction. Similar to other autoimmune diseases, genome-wide association studies (GWASs) have identified MG as a polygenic disease, with variants associated with T cell and B cell functions.[1,2] MG is also associated with thymoma and thymic hyperplasia. Currently, a thymectomy is the first choice of treatment for MG with thymoma, and the effectiveness of this treatment underscores the role of the thymus in disease pathogenesis.[3,4] However, a thymectomy is an invasive surgical procedure that can adversely affect the immune system.[5] In addition to thymectomy, only symptomatic treatments targeting the immune system or neuromuscular junctions are available, underscoring the need for the development of novel, less-invasive treatments that act upstream of the disease pathway. Therefore, the identification of thymic abnormalities related to MG is urgently needed.

The thymus is the primary lymphoid organ responsible for T cell education; it eliminates autoreactive T cells and induces regulatory T cells (Tregs), which serve as the site of central tolerance. However, due to the complexity of thymic function and structure, its physiological role and involvement in MG remain unclear. We previously identified the abnormal expression of neuromuscular-related molecules in MG-specific medullary thymic epithelial cells (mTECs) and germinal center (GC) formation in MG-associated thymomas using single-cell RNA sequencing (scRNA-seq) analysis.[6] However, spatial interpretation using scRNA-seq remains challenging. Therefore, there has



(legend on next page)

been no spatial prioritization to determine the areas within the complex thymic tissue that are truly related to the disease thus far. Although our scRNA-seq results suggested interactions between mTECs and immune cells, their spatial proximity has not been elucidated. Furthermore, a comprehensive understanding of the immune cells that form niches within the thymus is lacking.

Additionally, MG is associated not only with thymomas but also with thymic hyperplasia in younger patients.[7] Thymic hyperplasia is a benign condition characterized by the enlargement of the normal thymus and, similar to thymomas, is reported to involve the formation of lymphoid follicles with GCs.[8] Although both thymomas and thymic hyperplasia are thymic abnormalities associated with MG, whether there is a pathogenic link between the two remains controversial.

In recent years, spatial transcriptomics technology has evolved, greatly advancing our spatial understanding of disease processes.[9–11] Spatial transcriptomics has enabled significant improvements in the interpretation of cellular niches compared to observational methods with fewer parameters, such as H&E staining or immunohistochemistry (IHC). However, despite the significant amount of information it provides, assigning pathological significance and considering causality using spatial transcriptomics alone have been challenging. By integrating scRNA-seq data from the corresponding tissue, we can extract more information and estimate the cellular composition of each spot for a more detailed interpretation.[12] Nonetheless, there is currently no consensus on how to appropriately prioritize susceptible regions.

In this study, we conducted a spatial evaluation of MG thymomas using spatial transcriptome analysis to identify disease-related niches and characterize distinctive gene expression. We developed a method, single-cell disease-relevance score (scDRS)-spatial. This method leverages polygenic enrichment to identify disease-relevant spatial localizations by integrating single-cell spatial transcriptomics with disease GWAS, extending an existing method, scDRS,[13] that analyzes scRNA-seq data. In particular, scDRS-spatial considers physical contact between multiple cells, in addition to cell-type-specific polygenic enrichment, by assessing spatial niches rather than single cells. Furthermore, we reconstructed the largest single-cell atlas of thymomas by integrating data from previous reports.[6,14] By integrating this atlas with spatial transcriptomic data, we were able to estimate the detailed spatial interactions between cell populations. Through these integrated analyses, we attempted to identify hotspots of MG pathology in MG thymomas and the immune responses at these sites. Finally, we conducted a spatial transcriptome analysis of MG-associated thymic hyperplasia and discussed the similarities between the immune microenvironments of MG thymomas and hyperplasia.

RESULTS

Spatial-transcriptome profiling of thymoma, hyperplasia, and normal thymus

To investigate the spatial characteristics of thymuses associated with MG, we conducted a spatial transcriptome analysis. We previously reported a stronger association between thymomas and the presence of anti-acetylcholine receptor antibodies (AChR-Abs) than with the presence of MG-related symptoms. In this study, we primarily profiled thymomas (the World Health Organization [WHO] classification type B1 or B2) in patients positive for AChR-Ab (seropositive) as MG-type thymomas. We profiled the thymomas of four seropositive patients, two of whom exhibited MG symptoms, yielding five samples. Additionally, three samples were obtained from three seronegative patients (AChR-Ab$^-$, WHO type B1 or B2), and two thymic hyperplasia samples were obtained from two seropositive patients with MG symptoms. Formalin-fixed and paraffin-embedded sections were profiled using the 10x Genomics Visium platform (Figure 1A; Table S1). For comparison with normal thymuses, we integrated the Visium data of 11 samples from 11 individuals with normal fetal and pediatric thymuses. After quality control, 59,796 spots were retained for downstream analyses. Because each Visium spot is estimated to contain approximately 1–10 cells, each spot can be considered to represent a niche. Initially, the Leiden algorithm was used to define 18 clusters (Figures 1B, S1A, and S1B). Based on these 18 Leiden clusters, we defined 6 annotated clusters: the cortex, medulla, junction, stroma, and 2 medulla-specific clusters characterized by *FN1* expression (medulla_FN1) and a high concentration of GCs (medulla_GC) (Figures 1B and 1C). For instance, the distinct expression of chemokine-receptor pairs, such as *CCR9-CCL25* and *CCR7-CCL19*, significantly differentiated the cortex from the medulla (Figures 1D and S1C). The transcriptome profiles of the medulla and cortex were maintained, even in tumors (Figure 1D). We also investigated the expression of *AIRE*, a master regulator involved in self-antigen expression in mTECs, and found that in thymomas, its expression was lower compared to normal or thymic hyperplasia and was limited to the medulla (Figures S1D and S1E). Spatially, in normal thymuses, the cortex typically formed an outer layer with the medulla inside, whereas in thymomas, small medullary structures (on average, 9.77% in thymoma and 19.8%

Figure 1. Spatial transcriptomic analysis revealed histological structures in MG thymoma

(A) Schematic representation of the spatial transcriptomic analysis and enrolled sample numbers.

(B) Unsupervised clusters (Leiden) and annotation (annotated niche) of spots on uniform manifold approximation and projection (UMAP) plots.

(C) Heatmap showing mean expression of marker genes in annotated niche groups. Also see Figure S1A for automatically extracted marker genes.

(D) Representative spatial gene expression of normal thymus and thymoma samples.

(E) Distribution of disease status on UMAP plots.

(F) Comparison of the proportion of annotated niches in thymoma samples. Statistical analysis was performed using scCODA.[16].

(G) H&E staining, Leiden clusters, and annotated niche groups of representative samples. The arrowheads indicate a lymphoid follicle. The scale bars indicate 100 μm.

(H) IHC staining of CCR7 in MG and non-MG thymoma. GC was found in the medulla region (arrowhead). The scale bar represents 500 μm.

(I) The proportion of CCR7$^+$ area in thymoma was quantified using microscopic images (MG: WHO type B1, $n = 5$; non-MG: $n = 5$, and WHO type B2, MG: $n = 5$; non-MG: $n = 5$). Statistical analysis was performed using a two-sided Welch's t test.

in normal thymus) were interspersed predominantly within cortical structures (on average, 80.6% in thymoma and 73.9% in normal thymus) (Figures S1B and S2), as previously suggested.[15] The junction area was positioned both transcriptomically and spatially between the medulla and cortex (Figures 1C and S1F–S1H). Examination of the regional proportions of thymomas between seropositive and seronegative cases revealed no significant differences at the Leiden cluster level. However, the cortical region exhibited a significant decrease in seropositive cases (Figures 1E–1G and S1I). We further performed IHC staining for medulla-specific CCR7 to validate the medulla and cortex bias in MG thymoma. While no difference was observed in WHO type B2, we found that in type B1, the proportion of the $CCR7^+$ area was significantly larger in seropositive MG compared to seronegative non-MG (Figures 1H and 1I). The expression of an MG-specific gene set ("yellow module")[6] was highest in the junction (Figures S1J and S1K). We conducted a transcriptome comparison between seropositive and seronegative samples. The analysis revealed a relatively shared upregulation of pathways associated with acquired immune responses, particularly in the B cell lineage, interferon-γ signaling, major histocompatibility complex class II antigen presentation, and respiratory electron transport in both the medulla and cortex. Additionally, we observed an increase of the expressions of previously identified MG thymoma-specific genes, such as *NEFM*, *KRT6A*, and *KRT15* in seropositive thymomas (Figure S3).[6] Thus, by clustering the spatial transcriptome data of the thymus, we identified the predominant cortical and interspersed medullary structures in thymomas and revealed a reduction in the proportion of the cortex in MG-associated thymomas.

Prioritization of pathogenic niche in MG thymoma

Similar to other autoimmune diseases, MG is polygenic. We hypothesized that identifying the niches with genetic susceptibility to MG accumulation would allow us to prioritize these niches (Figure S4A). To this end, we extended scDRS[13] to spatial data, namely, the scDRS-spatial framework. scDRS integrates scRNA-seq data with GWAS to identify cell types with polygenic enrichment. scDRS-spatial goes beyond the cellular level by further assessing the polygenic enrichment of spatial niches that are hotspots of physical intercellular contact. We conducted null simulations using random gene sets and confirmed that scDRS-spatial was well calibrated for spatial transcriptome data (Figure S4B). Specifically, an imputation using Markov affinity-based graph imputation of cells (MAGIC)[17] produced conservative estimates (Figure S4B). Based on these findings, we imputed spatial data using MAGIC to reduce technical noise and estimated polygenic enrichment at each spot using scDRS-spatial. Furthermore, we added Visium data from various tissues across the human body as controls (Table S3). At the tissue level, the spots in the thymus were significantly associated with MG (Figure S4C). At the level of Leiden clusters, niche 8 (corresponding to the medulla) was significantly associated with a false discovery rate (FDR) of <0.2 (Figure S4D). Across all regions, the medulla was significantly associated with MG (FDR < 0.2; Figure S4E). Moreover, when stratified by condition, the proportion of associated niches and the heterogeneity in the medulla, especially niche 8, were higher in seropositive thymomas than in seronegative thymomas and the normal thymus (Figures 2A–2C). These results suggest that genetic susceptibility accumulates in the medullary regions of thymomas.

Cellular composition in MG-thymoma niche

To elucidate the cellular composition of MG-thymoma niches, we performed cell deconvolution by integrating scRNA-seq data. For deconvolution, we created a single-cell reference for thymomas by adding our data to the single-cell data reported by Xin et al.[14] After quality control, 113,948 cells were retained, defining 50 clusters, including immune, epithelial, and stromal cells (Figures 3A, 3B, and S5A–S5F). Notably, we achieved a higher-precision annotation of the TEC population, which was less represented in our previous study.[6] The mTECs were characterized by the expression of *CLDN4* (Figure S5A). Within the mTECs, several sub-clusters were defined, including $AIRE^{high}$ mTECs (mTEC *AIRE*), $KRT14^{high}$ mTECs (mTEC *KRT14*), and neuromuscular mTECs (nmTECs), which were characterized by a high yellow module and *GABRA5* expression (Figure S5B). In $CD4^+$ T cells, we identified *CXCL13*-producing T follicular helper cells (CD4 Tfh *CXCL13*). This population expressed Tfh-related functional genes *CXCR5*, *PDCD1*, *IL21*, and *BCL6* (Figure S5G), suggesting the active help of GC reaction by $CD4^+$ T cells in MG thymoma.[18,19] The dendritic cell (DC) fraction also included plasmacytoid DCs (pDC), conventional DCs type 1 (cDC1), type 2 (cDC2), and migratory DCs (migDCs), which were characterized by *CCR7* and *LAMP3* expression. migDCs expressed both *CD274* (PD-L1) and *CD80*, suggesting the involvement of T cell activation[20] (Figure S5H). We then assessed MG-specific features in the references to confirm their consistency. Deconvolution using bulk RNA-seq data of thymomas generated by The Cancer Genome Atlas (TCGA)[21] consortium revealed that the frequency of nmTECs was the most significantly associated with MG (Figure S6A, adjusted p value = 6×10^{-6}), similar to a previous result.[6] In addition, the expression of yellow module genes was highest in nmTECs (Figure S6B). This observation indicates that nmTECs were the most associated cell type at the single-cell level, even in the updated single-cell reference, which elaborated on the TEC populations.

Next, we leveraged the single-cell reference to analyze the spatial transcriptome data at cell-type resolution. By integrating the single-cell reference with Visium data using cell2location,[12] we estimated the cellular composition of each spot (Figures 3C and 3D). As in the normal thymus, immature T cells such as $CD4^-$ $CD8^-$ double-negative cells and $CD4^+$ $CD8^+$ double-positive T cells were concentrated in the cortical region, whereas mature T cells and mTECs were abundant in the medullary region (Figures 3C and 3D). GC B cells (B GC) and CD4 Tfh *CXCL13* were also enriched in the medullary GC region (Figures 3C and 3D). The stroma and medulla_FN1 regions were characterized by high numbers of endothelial cells, fibroblasts, and vascular smooth muscle cells (Figures 3C and S7A). In the seropositive cases, an increase in nmTECs and immune cells, such antibody-secreting cells (ASCs), switched memory B cells (B SM), B GC, migDCs, and effector T regs (Treg Eff), and reduced cortical TECs (cTECs) were confirmed (Figures 3E and S7B). Next, we explored the co-localization patterns of constituent cells using non-negative matrix factorization (NMF), defining

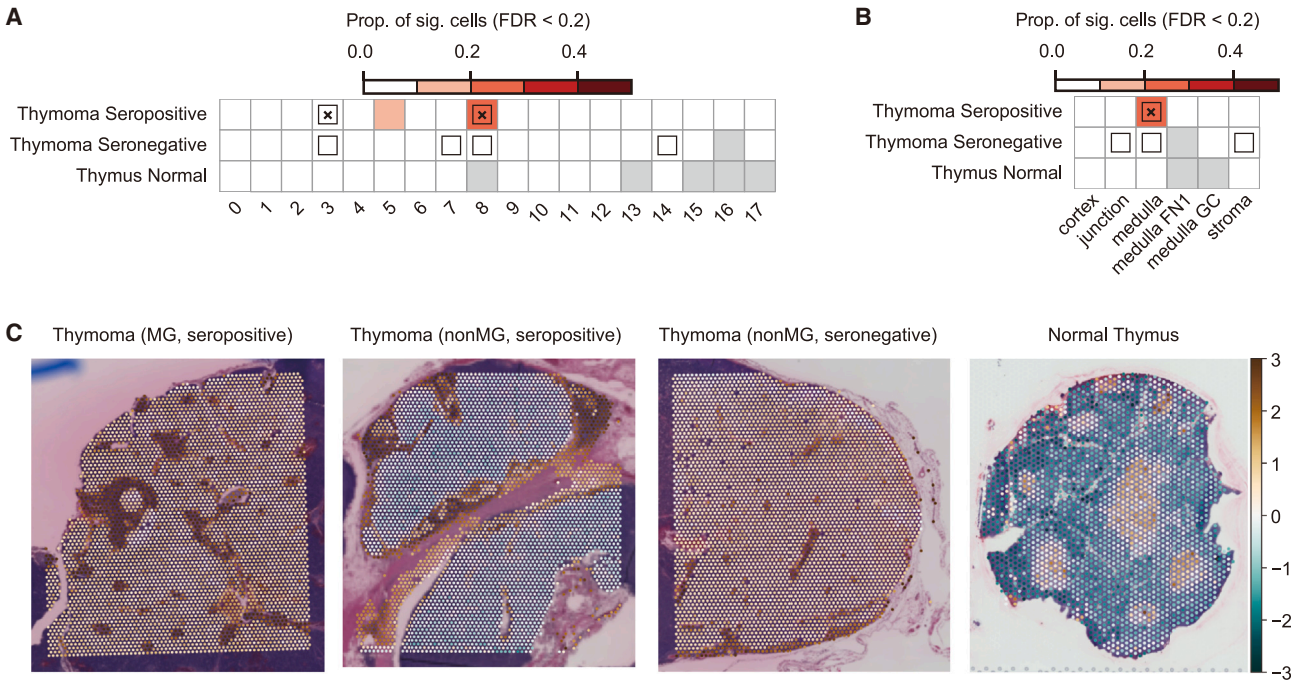


Figure 2. scDRS-spatial unveiled polygenic enrichment in the medulla in MG thymoma

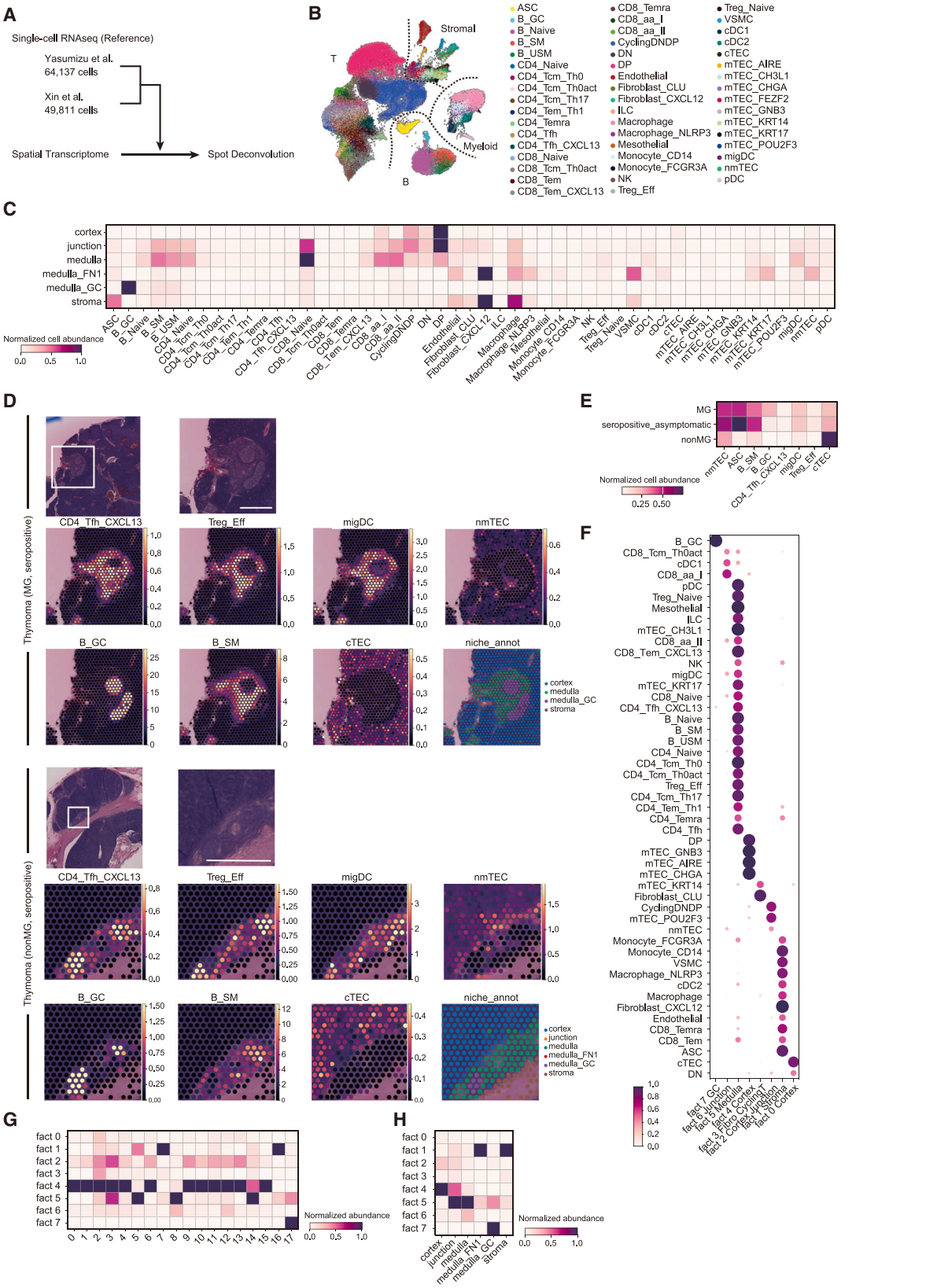
(A and B) Heatmaps show disease association in Leiden clusters (A) and annotated niche (B). Samples were stratified with disease conditions. Heatmap colors depict the proportion of significant cells (FDR < 0.2) evaluated using scDRS.[13] Squares denote significant disease associations (FDR < 0.05), and cross symbols denote significant heterogeneity in association (FDR < 0.05).

(C) scDRS scores on representative Visium slides.

eight co-localization factors (factors 0–7) (Figures 3F–3H and S7C). By analyzing the cellular contributions and enriched regions of each factor, we found that certain factors were predominantly associated with specific regions: factors 0, 2, and 4 with the cortex; factor 5 with the medulla; factors 1 and 3 with the stroma; and factor 7 with both the junction and GCs. Factor 7, composed of B GC, was localized within GCs, while CD4 Tfh *CXCL13* was present both inside and around GCs in the medullary region, forming the GC niche (Figures 3F–3H and S7C). Factor 5, composed of mature immune cells, such as Treg Eff, migDCs, and B SM, constituted an immune microenvironment in the medulla (Figures 3F–3H and S7C). Factor 6, comprising nmTECs, cDC1, and migDCs, was particularly abundant at the junction area (Figures 3F–3H and S7C). The ASC niche was not identified within the cortex or medulla, but it was present in the stromal region (Figures 3C and S7A). Endothelial cells were concentrated in the medulla and stroma, highlighting a lower vascular density in the cortex (Figures 3C and S7A). In summary, cell deconvolution identified eight co-localizing communities and their constituent cells.

Cell-cell interaction analysis reveals niche-specific chemokine profiles

Next, we analyzed cell-cell interactions (CCIs) within the cell groups constituting the niches. Using CellphoneDB,[22] we explored CCIs by considering the co-localizing communities identified by cell2location analysis. Numerous CCIs were identified, among which chemokines were particularly cell specific and appeared to be involved in niche-specific cell migration (Figures S8A, S8B, and S9). In both tumor and normal tissues, *CCL25-CCR9* and *CCL19-CCR7* interactions were specific to the cortex and medulla, respectively (Figures 4A and 4B). Previously, we reported that nmTECs have an intermediate profile between that of mTECs and cTECs,[6] and indeed, they expressed both *CCL25* and *CCL19* (Figure 4A). Interestingly, in thymomas, both single-positive T cells and migDCs expressed *CCR7*, suggesting that the medullary characteristics of thymomas facilitate the mobilization of migDCs. Ligands for *CCR4* specific to Treg Eff, such as *CCL17* and *CCL22*, were expressed by migDCs in thymomas, suggesting their role in maintaining Treg Eff in the medulla[23] (Figures 4A and 4B). Similarly, *CXCL16*, the ligand for *CXCR6* specific to Treg Eff, was expressed in cDC1, cDC2, and migDCs (Figures 4A and 4B). migDCs also expressed *CXCL10*, which potentially interacts with *CXCR3*$^{+}$ effector T cells (Figures 4A and 4B). We previously demonstrated that mature infiltrating T/B cells in the thymus specifically express *CXCR4*.[6] The *CXCL12* ligand was expressed by nmTECs,[6] suggesting its role in maintaining the medullary niche (Figures 4A and 4B). Finally, *CXCL13,* a key chemokine for the maintenance of the GC, was expressed by CD4 Tfh *CXCL13* (Figure 4A). The expression of *CCR4*, *CXCL16*, and *CXCR5-CXCL13* was lower in the normal thymus than in the thymoma, suggesting their thymoma-specific roles in maintaining niches (Figure 4A). In contrast, chemokines such as *CCL25*, *CCL19*, and *CXCL12* and their receptors were expressed in both the normal thymus and thymoma, suggesting that some factors might be



(legend on next page)

synchronized with normal conditions and MG thymoma (Figure 4A). Taken together, we identified spatially characteristic chemokine ligand-receptor pairs in thymomas, supporting the involvement of these niches in the pathogenesis of thymoma-associated MG.

Extrapolation of thymoma niche to thymic hyperplasia

Finally, we verified whether our findings were consistent with those in thymic hyperplasia. Histologically similar to the normal thymus, the structure with the cortex on the outside and the medulla on the inside was maintained (Figure 5A). GCs are present in the medulla, similar to thymomas, suggesting that the microenvironment supporting GC formation is common in both thymomas and thymic hyperplasia (Figure 5A). Polygenic signals identified by scDRS-spatial analysis were generally more enriched in thymoma samples and were particularly observed in the medulla, similar to our findings in thymomas (Figures 5B, 5C, S10A, and S10B). Although there is no scRNA-seq reference for thymic hyperplasia, application of the thymoma reference revealed that the eight-cell communities identified in thymomas were consistently formed in accordance with histological features (Figure 5D). Furthermore, the expression of chemokines and their receptors was consistent with thymomas, and *CCR4*, *CXCL16*, and *CXCR5-CXCL13*, which had lower expressions in the normal thymus, were abundantly expressed in hyperplasia (Figures 5E and 5F). These findings indicate that an immune microenvironment supporting GCs is present in the medulla in thymic hyperplasia as well as thymoma.

DISCUSSION

In this study, spatial transcriptomics was used to identify the niche involved in the pathogenesis of MG thymoma and to explore its molecular characteristics. We successfully identified the MG-associated niche and its constituents in both thymomas and thymic hyperplasia. Our analysis revealed that cortical-like areas, medullary-like areas, and immune hotspots coexisted within a single patient, highlighting the heterogeneity of the tumor environment within an individual. A relative enlargement of medulla was observed in seropositive thymoma. In line with these results, genetic and phenotypic associations of the medulla were also suggested. Furthermore, we identified the formation of ectopic lymphoid structures in the MG thymus and the chemokines that support these structures.

The significance of the medulla has been frequently discussed, including in our previous single-cell analyses.[6,24–26] mTECs play a crucial role in negative selection by eliminating autoreactive T cells through self-antigen production.[27–29] The abnormalities in this process in MG highlight that negative selection has a potential risk of inducing autoimmunity. In particular, the expression of neuromuscular-related antigens by nmTECs in MG thymomas has been suggested to feed on autoreactive T cells.[6] These nmTECs were localized at the junction of the medulla and cortex, suggesting that the origin of these nmTECs in tumor development was at this junction. Additionally, our analysis demonstrated the accumulation of migDCs in the medulla. migDCs expressing *CCR7* migrate to tertiary lymphoid structures or lymph nodes with high concentrations of *CCL19* and play an important role in T cell priming.[30,31] The medulla, due to mTEC-induced *CCL19* expression, may physiologically trap migDCs and mediate T cell help. Furthermore, CCI between migDCs and mTECs has been noted, even in the normal thymus,[26] and this collaboration may be attributed to normal thymic function.

Furthermore, our analysis provides insights into the role of specific immune cells in the pathogenesis of MG. A concurrent abundance of $CXCL13^{+}$ $IL21^{+}$ Tfh cells within the lymphoid follicles and accumulation of migDCs in the medulla were observed. These findings suggest that follicle formation in the thymus induces potent affinity maturation and B cell proliferation, possibly contributing to the pathogenesis of MG.[32,33] $CXCR5^{-}$ $PDCD1^{+}$ T peripheral helper (Tph) cells, observed at inflammatory sites in rheumatoid arthritis, systemic lupus erythematosus, and Sjögren syndrome,[34–36] were not clearly identified as distinct cell populations in our single-cell analysis. In contrast, other studies and ours have reported an increase in circulating Tph cells in MG, highlighting the need for further investigation on Tfh and Tph cells in relation to the progression of MG.[2,37] Additionally, effector Tregs were observed to be abundant in the medulla of MG thymoma. Although the accumulation of GWAS signals in Tregs[2,6] and their dysfunction in MG patients[38] have been suggested, whether normalizing their function could lead to therapeutic effects remains an important subject for future research.

Notably, as MG is an antibody-mediated disease, it has been reported that ASCs are increased in MG thymoma.[39] However, the niche for ASCs was not found within the thymic cortex or medulla but was rather abundant in the stromal region. Even after thymectomy, the circulation of autoreactive B cell clones in the periphery has been reported,[40] suggesting that extrathymic niches, such as the bone marrow,[41] may harbor ASCs. Nevertheless, it was suggested that the immune microenvironment within the medulla primarily contributes to B cell maturation.

This study also profiled MG-associated thymic hyperplasia. Because there are no single-cell datasets specific to thymic hyperplasia, a detailed comparison of thymic epithelial cell profiles was not possible. Nonetheless, our analysis revealed notable similarities in immune cells, chemokine profiles, and polygenic signals in the medulla of MG thymomas. Consequently, this

Figure 3. Cell deconvolution analysis revealed cellular composition in MG thymoma

(A) Schematic of scRNA-seq reference construction and cell deconvolution analysis.

(B) Cell clusters of the reference scRNAseq data on UMAP plot.

(C) Cellular decomposition in each annotated niche group. Deconvolution was performed using cell2location.[12]

(D) Normalized cellular decomposition and H&E staining of representative Visium slides. The scale bars indicate 100 μm.

(E) Normalized cellular decomposition in each disease condition.

(F) Cell compartments identified using NMF. The normalized NMF weights of cell types across NMF components are shown.

(G and H) Distributions of cell compartments across Leiden clusters (G) and annotated niche groups (H). The abundance was normalized for each column.

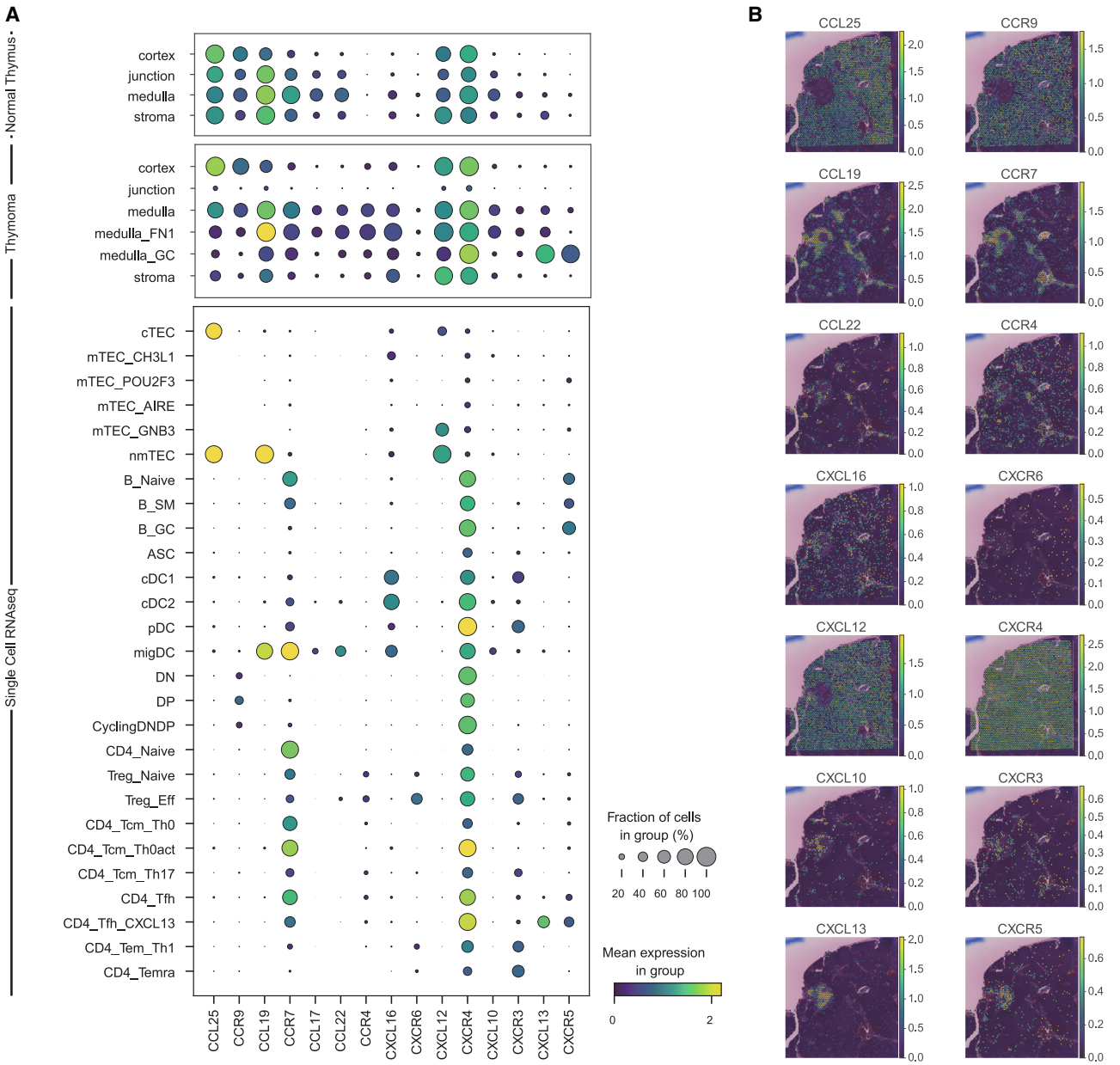


Figure 4. Niche-specific cytokine organization identified by CCI analysis

(A) Dot plot showing cytokine expressions across annotated niche groups in the normal thymus (top) and thymoma (center), and across major cell types (bottom). Gene expressions for annotated niche groups were from the Visium dataset, and those for the cell types were from scRNAseq.

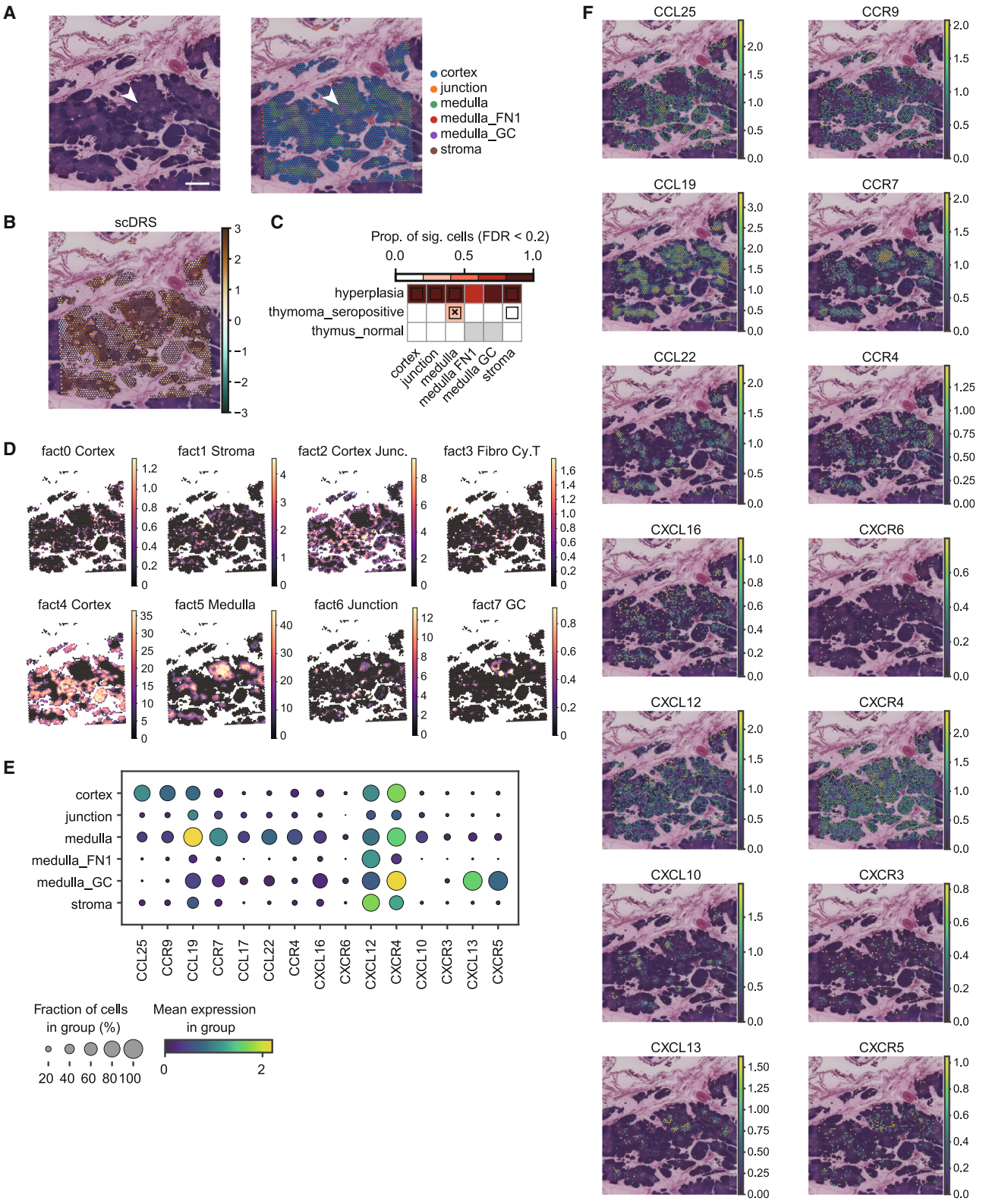
(B) Representative cytokine expression of a thymoma sample. Ligands (left) and receptors (right) are shown correspondingly per line.

study offers an invaluable resource for understanding the pathogenesis of MG by presenting a comprehensive overview of thymomas and thymic hyperplasia.

In summary, using spatial transcriptomic analysis, we successfully identified the immune microenvironment in the medulla, revealing that many of its characteristics resonate with the physiological features of the thymus. Current treatments for MG, aside from thymectomy, are mainly supportive and target the immune system and neuromuscular junctions. We hope that this study will contribute to a complete understanding of MG pathogenesis and the development of novel treatments targeting upstream pathological processes.

Limitations of the study

There are several limitations worth noting. First, the limited sample size should be acknowledged. In our previous study,[6] the scarcity of samples in scRNA-seq analysis of MG thymoma prompted the integration of large-scale bulk RNA-seq data,



(legend on next page)

revealing the significance of nmTECs and immune cell interactions. This study is consistent with these findings and provides further insights. Second, the constraints of spatial transcriptomics technology, such as the 50-μm scale of Visium, necessitated cell deconvolution due to the inclusion of multiple cells per spot. Recent advancements in subcellular resolution techniques[42] and *in situ* single-nucleus barcoding[43] offer potential improvements in this regard. Lastly, our study focuses on observations from post-onset samples, limiting our ability to show the causality of pathogenesis. Furthermore, we were unable to elucidate the temporal sequence of disease onset. Ultimately, establishing natural onset models and conducting intervention experiments will be essential.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Yoshiaki Yasumizu (yoshiaki.yasumizu@yale.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Raw sequence data have been deposited at Japanese Genotype-phenotype Archive (JGA) under accession number JGAS000672. Microscope images and Space Ranger output files are available at Genomic Expression Archive (GEA) under accession number E-GEAD-747.
- The code is available on GitHub (https://github.com/yyoshiaki/MG_Visium_2023). The processed spatial transcriptome and scRNA-seq data are available in figshare (https://doi.org/10.6084/m9.figshare.25052546).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported by the Grants-in-Aid for Scientific Research Grant 23H02826 from the Ministry of Education, Culture, Sports, Science and Technology of Japan. Y.Y. was supported by the Takeda Science Foundation. We acknowledge the next-generation sequencing core facility of the Genome Information Research Center at the Research Institute for Microbial Diseases, Osaka University for its support in sequencing, as well as the BIKEN Foundation for preparing tissue sections for visual analysis. This work was partly achieved through the use of SQUID at the Cybermedia Center, Osaka University. The illustrations were generated using BioRender.com. We acknowledge Dr. Ryoji Yasumizu and Nardos Cheru for their critical comments on the manuscript.

AUTHOR CONTRIBUTIONS

Y.Y., M.K., N.K., S.N., and T.O. designed all the experiments. Y.Y. performed the bioinformatics analysis, prepared the figures, and drafted the manuscript. T.O., M.K., and M.J.Z. reviewed and edited the manuscript. N.K., D.M., K.S., D.O., and S.N. performed the experiments. S.N., S.F., Y.S., and E.M. provided samples for analysis. M.J.Z., S.N., and E.M. provided expert advice. All authors critically reviewed and edited the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this study, the authors used ChatGPT4 to improve language and readability. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Human samples
- METHOD DETAILS
  - Spatial transcriptomics (CytAssist Visium)
  - Visium data analysis
  - scDRS-spatial
  - Single-cell RNA-seq analysis
  - Cell deconvolution of Visium samples using Cell2location
  - Cell deconvolution of TCGA bulk RNA-seq samples using scaden
  - Cell-cell interaction analysis by CellphoneDB
  - Immunohistochemistry staining and quantification
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114677.

Received: February 5, 2024
Revised: July 30, 2024
Accepted: August 8, 2024
Published: August 23, 2024

REFERENCES

1. Chia, R., Saez-Atienzar, S., Murphy, N., Chiò, A., Blauwendraat, C., International Myasthenia Gravis Genomics Consortium; Roda, R.H., Tienari, P.J., Kaminski, H.J., Ricciardi, R., et al. (2022). Identification of genetic risk loci and prioritization of genes and pathways for myasthenia gravis: a genome-wide association study. Proc. Natl. Acad. Sci. USA *119*, e2108672119. https://doi.org/10.1073/pnas.2108672119.
2. Yasumizu, Y., Takeuchi, D., Morimoto, R., Takeshima, Y., Okuno, T., Kinoshita, M., Morita, T., Kato, Y., Wang, M., Motooka, D., et al. (2024).

Figure 5. Spatial characteristics of thymic hyperplasia

(A) H&E staining and annotated niche groups of a thymic hyperplasia sample. The arrowhead indicates a lymphoid follicle. The scale bar indicates 100 μm.
(B) scDRS scores of representative Visium slides.
(C) Heatmap shows disease association in annotated niches stratified by disease conditions. Heatmap colors depict the proportion of significant cells (FDR < 0.2) evaluated using scDRS.[13] Squares denote significant disease associations (FDR < 0.05), and cross symbols denote significant heterogeneity in association (FDR < 0.05).
(D) Distributions of cell compartments defined by NMF.
(E) Dot plot showing cytokine expressions across annotated niche groups in thymic hyperplasia.
(F) Representative cytokine expression of a thymoma sample. Ligands (left) and receptors (right) are shown correspondingly per line.

Single-cell transcriptome landscape of circulating CD4+ T cell populations in autoimmune diseases. Cell Genom. *4*, 100473.

3. Wolfe, G.I., Kaminski, H.J., Aban, I.B., Minisman, G., Kuo, H.-C., Marx, A., Ströbel, P., Mazia, C., Oger, J., Cea, J.G., et al. (2016). Randomized Trial of Thymectomy in Myasthenia Gravis. N. Engl. J. Med. *375*, 511–522.
4. Masaoka, A., Yamakawa, Y., Niwa, H., Fukai, I., Kondo, S., Kobayashi, M., Fujii, Y., and Monden, Y. (1996). Extended thymectomy for myasthenia gravis patients: a 20-year review. Ann. Thorac. Surg. *62*, 853–859.
5. Kooshesh, K.A., Foy, B.H., Sykes, D.B., Gustafsson, K., and Scadden, D.T. (2023). Health Consequences of Thymus Removal in Adults. N. Engl. J. Med. *389*, 406–417.
6. Yasumizu, Y., Ohkura, N., Murata, H., Kinoshita, M., Funaki, S., Nojima, S., Kido, K., Kohara, M., Motooka, D., Okuzaki, D., et al. (2022). Myasthenia gravis-specific aberrant neuromuscular gene expression by medullary thymic epithelial cells in thymoma. Nat. Commun. *13*, 4230.
7. Cron, M.A., Maillard, S., Villegas, J., Truffault, F., Sudres, M., Dragin, N., Berrih-Aknin, S., and Le Panse, R. (2018). Thymus involvement in early-onset myasthenia gravis. Ann. N. Y. Acad. Sci. *1412*, 137–145.
8. Levine, G.D., and Rosai, J. (1978). Thymic hyperplasia and neoplasia: a review of current concepts. Hum. Pathol. *9*, 495–515.
9. Kanemaru, K., Cranley, J., Muraro, D., Miranda, A.M.A., Ho, S.Y., Wilbrey-Clark, A., Patrick Pett, J., Polanski, K., Richardson, L., Litvinukova, M., et al. (2023). Spatially resolved multiomics of human cardiac niches. Nature *619*, 801–810.
10. Lerma-Martin, C., Badia-i-Mompel, P., Ramirez Flores, R.O., Sekol, P., Hofmann, A., Thäwel, T., Riedl, C.J., Wünnemann, F., Ibarra-Arellano, M.A., Trobisch, T., et al. (2022). Spatial cell type mapping of multiple sclerosis lesions. Preprint at bioRxiv. https://doi.org/10.1101/2022.11.03.514906.
11. Rao, A., Barkley, D., França, G.S., and Yanai, I. (2021). Exploring tissue architecture using spatial transcriptomics. Nature *596*, 211–220.
12. Kleshchevnikov, V., Shmatko, A., Dann, E., Aivazidis, A., King, H.W., Li, T., Elmentaite, R., Lomakin, A., Kedlian, V., Gayoso, A., et al. (2022). Cell2location maps fine-grained cell types in spatial transcriptomics. Nat. Biotechnol. *40*, 661–671.
13. Zhang, M.J., Hou, K., Dey, K.K., Sakaue, S., Jagadeesh, K.A., Weinand, K., Taychameekiatchai, A., Rao, P., Pisco, A.O., Zou, J., et al. (2022). Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. Nat. Genet. *54*, 1572–1580.
14. Xin, Z., Lin, M., Hao, Z., Chen, D., Chen, Y., Chen, X., Xu, X., Li, J., Wu, D., Chai, Y., and Wu, P. (2022). The immune landscape of human thymic epithelial tumors. Nat. Commun. *13*, 5463.
15. Marx, A., Ströbel, P., and Weis, C.-A. (2018). The pathology of the thymus in myasthenia gravis. Mediastinum *2*, 66.
16. Lopez, R., Regier, J., Cole, M., Jordan, M., and Yosef, N. (2018). Bayesian Inference for a Generative Model of Transcriptome Profiles from Single-cell RNA Sequencing. Preprint at bioRxiv. https://doi.org/10.1101/292037.
17. van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A.J., Burdziak, C., Moon, K.R., Chaffer, C.L., Pattabiraman, D., et al. (2018). Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. Cell *174*, 716–729.e27.
18. Yoshitomi, H., and Ueno, H. (2021). Shared and distinct roles of T peripheral helper and T follicular helper cells in human diseases. Cell. Mol. Immunol. *18*, 523–527.
19. Crotty, S. (2019). T Follicular Helper Cell Biology: A Decade of Discovery and Diseases. Immunity *50*, 1132–1148.
20. Sugiura, D., Maruhashi, T., Okazaki, I.-M., Shimizu, K., Maeda, T.K., Takemoto, T., and Okazaki, T. (2019). Restriction of PD-1 function by *cis*-PD-L1/CD80 interactions is required for optimal T cell responses. Science *364*, 558–566.
21. Hoadley, K.A., Yau, C., Hinoue, T., Wolf, D.M., Lazar, A.J., Drill, E., Shen, R., Taylor, A.M., Cherniack, A.D., Thorsson, V., et al. (2018). Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell *173*, 291–304.e6.
22. Garcia-Alonso, L., Handfield, L.-F., Roberts, K., Nikolakopoulou, K., Fernando, R.C., Gardner, L., Woodhams, B., Arutyunyan, A., Polanski, K., Hoo, R., et al. (2021). Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. Nat. Genet. *53*, 1698–1711.
23. Nakandakari-Higa, S., Canesso, M.C.C., Walker, S., Chudnovskiy, A., Jacobsen, J.T., Bilanovic, J., Parigi, S.M., Fiedorczuk, K., Fuchs, E., Bilate, A.M., et al. (2023). Universal recording of cell-cell contacts in vivo for interaction-based transcriptomics. Preprint at bioRxiv, 2023.03.16.533003. https://doi.org/10.1101/2023.03.16.533003.
24. Yamada, Y. (2023). Histogenetic and disease-relevant phenotypes in thymic epithelial tumors (TETs): The potential significance for future TET classification. Pathol. Int. *73*, 265–280.
25. Marx, A., Yamada, Y., Simon-Keller, K., Schalke, B., Willcox, N., Ströbel, P., and Weis, C.-A. (2021). Thymus and autoimmunity. Semin. Immunopathol. *43*, 45–64.
26. Park, J.-E., Botting, R.A., Domínguez Conde, C., Popescu, D.-M., Lavaert, M., Kunz, D.J., Goh, I., Stephenson, E., Ragazzini, R., Tuck, E., et al. (2020). A cell atlas of human thymic development defines T cell repertoire formation. Science *367*. 2020.01.28.911115.
27. Kadouri, N., Nevo, S., Goldfarb, Y., and Abramson, J. (2020). Thymic epithelial cell heterogeneity: TEC by TEC. Nat. Rev. Immunol. *20*, 239–253.
28. Anderson, M.S., Venanzi, E.S., Klein, L., Chen, Z., Berzins, S.P., Turley, S.J., von Boehmer, H., Bronson, R., Dierich, A., Benoist, C., and Mathis, D. (2002). Projection of an immunological self shadow within the thymus by the aire protein. Science *298*, 1395–1401.
29. Aschenbrenner, K., D'Cruz, L.M., Vollmann, E.H., Hinterberger, M., Emmerich, J., Swee, L.K., Rolink, A., and Klein, L. (2007). Selection of Foxp3+ regulatory T cells specific for self antigen expressed and presented by Aire+ medullary thymic epithelial cells. Nat. Immunol. *8*, 351–358.
30. Li, J., Zhou, J., Huang, H., Jiang, J., Zhang, T., and Ni, C. (2023). Mature dendritic cells enriched in immunoregulatory molecules (mregDCs): A novel population in the tumour microenvironment and immunotherapy target. Clin. Transl. Med. *13*, e1199.
31. Magen, A., Hamon, P., Fiaschi, N., Soong, B.Y., Park, M.D., Mattiuz, R., Humblin, E., Troncoso, L., D'souza, D., Dawson, T., et al. (2023). Intratumoral dendritic cell-CD4+ T helper cell niches enable CD8+ T cell differentiation following PD-1 blockade in hepatocellular carcinoma. Nat. Med. *29*, 1389–1399.
32. Zuckerman, N.S., Howard, W.A., Bismuth, J., Gibson, K., Edelman, H., Berrih-Aknin, S., Dunn-Walters, D., and Mehr, R. (2010). Ectopic GC in the thymus of myasthenia gravis patients show characteristics of normal GC. Eur. J. Immunol. *40*, 1150–1161.
33. Sims, G.P., Shiono, H., Willcox, N., and Stott, D.I. (2001). Somatic hypermutation and selection of B cells in thymic germinal centers responding to acetylcholine receptor in myasthenia gravis. J. Immunol. *167*, 1935–1944.
34. Bocharnikov, A.V., Keegan, J., Wacleche, V.S., Cao, Y., Fonseka, C.Y., Wang, G., Muise, E.S., Zhang, K.X., Arazi, A., Keras, G., et al. (2019). PD-1hiCXCR5- T peripheral helper cells promote B cell responses in lupus via MAF and IL-21. JCI Insight *4*, e130062. https://doi.org/10.1172/jci.insight.130062.
35. Pontarini, E., Murray-Brown, W.J., Croia, C., Lucchesi, D., Conway, J., Rivellese, F., Fossati-Jimack, L., Astorri, E., Prediletto, E., Corsiero, E., et al. (2020). Unique expansion of IL-21+ Tfh and Tph cells under control of ICOS identifies Sjögren's syndrome with ectopic germinal centres and MALT lymphoma. Ann. Rheum. Dis. *79*, 1588–1599.
36. Rao, D.A., Gurish, M.F., Marshall, J.L., Slowikowski, K., Fonseka, C.Y., Liu, Y., Donlin, L.T., Henderson, L.A., Wei, K., Mizoguchi, F., et al. (2017). Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. Nature *542*, 110–114.

37. Çebi, M., Durmus, H., Aysal, F., Özkan, B., Gül, G.E., Çakar, A., Hocaoglu, M., Mercan, M., Yentür, S.P., Tütüncü, M., et al. (2020). CD4+ T Cells of Myasthenia Gravis Patients Are Characterized by Increased IL-21, IL-4, and IL-17A Productions and Higher Presence of PD-1 and ICOS. Front. Immunol. *11*, 809.

38. Gradolatto, A., Nazzal, D., Truffault, F., Bismuth, J., Fadel, E., Foti, M., and Berrih-Aknin, S. (2014). Both Treg cells and Tconv cells are defective in the Myasthenia gravis thymus: roles of IL-17 and TNF-α. J. Autoimmun. *52*, 53–63.

39. Yamamoto, Y., Matsui, N., Uzawa, A., Ozawa, Y., Kanai, T., Oda, F., Kondo, H., Ohigashi, I., Takizawa, H., Kondo, K., et al. (2021). Intrathymic Plasmablasts Are Affected in Patients With Myasthenia Gravis With Active Disease. Neurol. Neuroimmunol. Neuroinflamm. *8*, e1087. https://doi.org/10.1212/NXI.0000000000001087.

40. Jiang, R., Hoehn, K.B., Lee, C.S., Pham, M.C., Homer, R.J., Detterbeck, F.C., Aban, I., Jacobson, L., Vincent, A., Nowak, R.J., et al. (2020). Thymus-derived B cell clones persist in the circulation after thymectomy in myasthenia gravis. Proc. Natl. Acad. Sci. USA *117*, 30649–30660.

41. Fujii, Y., Monden, Y., Hashimoto, J., Nakahara, K., and Kawashima, Y. (1985). Acetylcholine receptor antibody production by bone marrow cells in a patient with myasthenia gravis. Neurology *35*, 577–579.

42. Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., et al. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell *185*, 1777–1792.e21.

43. Russell, A.J.C., Weir, J.A., Nadaf, N.M., Shabet, M., Kumar, V., Kambhampati, S., Raichur, R., Marrero, G.J., Liu, S., Balderrama, K.S., et al. (2023). Slide-tags enables single-nucleus barcoding for multimodal spatial genomics. Nature *625*, 101–109.

44. Büttner, M., Ostner, J., Müller, C.L., Theis, F.J., and Schubert, B. (2021). scCODA is a Bayesian model for compositional single-cell data analysis. Nat. Commun. *12*, 6876.

45. Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A.C., Kuemmerle, L.B., Rybakov, S., Ibarra, I.L., Holmberg, O., Virshup, I., et al. (2022). Squidpy: a scalable framework for spatial omics analysis. Nat. Methods *19*, 171–178.

46. Heimli, M., Flåm, S.T., Hjorthaug, H.S., Trinh, D., Frisk, M., Dumont, K.-A., Ribarska, T., Tekpli, X., Saare, M., and Lie, B.A. (2022). Multimodal human thymic profiling reveals trajectories and cellular milieu for T agonist selection. Front. Immunol. *13*, 1092028.

47. Suo, C., Dann, E., Goh, I., Jardine, L., Kleshchevnikov, V., Park, J.-E., Botting, R.A., Stephenson, E., Engelbert, J., Tuong, Z.K., et al. (2022). Mapping the developing human immune system across organs. Science *376*, eabo0510.

48. de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput. Biol. *11*, e1004219.

49. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15.

50. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS *16*, 284–287.

51. Yu, G., and He, Q.-Y. (2016). ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol. Biosyst. *12*, 477–479.

52. Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Syst. *8*, 281–291.e9.

53. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296.

54. Menden, K., Marouf, M., Oller, S., Dalmia, A., Magruder, D.S., Kloiber, K., Heutink, P., and Bonn, S. (2020). Deep learning-based cell composition analysis from tissue expression profiles. Sci. Adv. *6*, eaba2619.

55. Colaprico, A., Silva, T.C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T.S., Malta, T.M., Pagnotta, S.M., Castiglioni, I., et al. (2016). TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. *44*, e71.

56. Seabold, S., and Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. In Proceedings of the 9th Python in Science Conference (SciPy). https://doi.org/10.25080/majora-92bf1922-011.

57. Troulé, K., Petryszak, R., Prete, M., Cranley, J., Harasty, A., Tuong, Z.K., Teichmann, S.A., Garcia-Alonso, L., and Vento-Tormo, R. (2023). CellPhoneDB v5: inferring cell-cell communication from single-cell multiomics data. Preprint at arXiv. https://doi.org/10.48550/arXiv.2311.04567.

58. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

59. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation *2*, 100141.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-CCR7 antibody [EPR23192-57]	Abcam	Cat# ab253187, RRID: AB_2922673
Critical commercial assays		
Visium Spatial for FFPE Gene Expression Kit, Human Transcriptome, 16 rxns	10x Genomics	Cat# 1000336
DNBSEQ-G400RS High-throughput Sequencing Set (App-A FCL PE100)	MGI Tech	Cat# 1000016994
MGIEasy Universal Library Conversion Kit (App-A)	MGI Tech	Cat# 1000004155
EnVision FLEX TARGET RETRIEVAL SOLUTION HIGH pH(50x)	DAKO	Cat# DM828/K8023
EnVision FLEX PEROXIDASE-BLOCKING REAGENT	DAKO	Cat# SM801/K8023
EnVision FLEX/HRP	DAKO	Cat# SM802/K8023
EnVision FLEX SUBSTRATE BUFFER + EnVision FLEX DAB+	DAKO	Cat# DM803+DM827/K8023
Deposited data		
Spatial transcriptome data	This paper	Raw data: JGAS000672, Processed data: E-GEAD-747 and https://doi.org/10.6084/m9.figshare.25052546
Spatial transcriptome data (Healthy thymus)	Heimli et al.[44] and Suo[45]	GSE207205 and https://developmental.cellatlas.io/fetal-immune
scRNA-seq data from MG thymoma	Yasumizu et al.[6]	JGAS000482
scRNA-seq data from thymoma	Xin et al.[14]	HRA002334
GWAS for MG	Chia et al.[1]	GCST90093061
Gene locations (MAGMA)	De Leeuw et al.[46]	https://ctg.cncr.nl/software/MAGMA/aux_files/NCBI37.3.zip
Reference data (MAGMA)	De Leeuw et al.[46]	https://ctg.cncr.nl/software/MAGMA/ref_data/g1000_eur.zip
Software and algorithms		
Code and algorithms for analysis	This paper	https://github.com/yyoshiaki/MG_Visium_2023 (https://doi.org/10.5281/zenodo.13161044)
Space Ranger	10x Genomics	V2.0.1
Scanpy	Wolf et al.[47]	https://scanpy.readthedocs.io/en/stable/
scvi-tool	Lopez et al.[48]	https://scvi-tools.org/
Squidpy	Palla et al.[49]	https://squidpy.readthedocs.io/en/stable/
scCODA	Büttner et al.[16]	https://sccoda.readthedocs.io/en/latest/
clusterProfiler	Yu et al.[50]	https://github.com/YuLab-SMU/clusterProfiler
ReactomePA	Yu et al.[51]	https://github.com/YuLab-SMU/ReactomePA
scDRS	Zhang et al.[13]	https://martinjzhang.github.io/scDRS/
MAGMA	De Leeuw et al.[46]	https://cncr.nl/research/magma/
Scrublet	Wolock et al.[52]	https://github.com/swolock/scrublet
Harmony	Korsunsky et al.[53]	https://github.com/immunogenomics/harmony
Cell2location	Kleshchevnikov et al.[12]	https://github.com/BayraktarLab/cell2location
Scaden	Menden et al.[54]	https://scaden.readthedocs.io/en/latest/
TCGAbiolinks	Colaprico et al.[55]	https://bioconductor.org/packages/release/bioc/html/TCGAbiolinks.html
statsmodels	Seabold et al.[56]	https://github.com/statsmodels/statsmodels

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CellphoneDB	Garcia-Alonso et al.[57]	https://github.com/ventolab/CellphoneDB
ktplotspy	Troulé et al.[56]	https://ktplotspy.readthedocs.io/en/latest/index.html
Fiji	Schindelin et al.[58]	https://imagej.net/software/fiji/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human samples

This study was reviewed and approved by the Research Ethics Committee of Osaka University and was conducted in accordance with the guidelines and regulations. Human samples were collected with the approval of Osaka University's review board (protocol: ID 10038-13. Detailed information on the participants is provided in Table S1.

METHOD DETAILS

Spatial transcriptomics (CytAssist Visium)

Formalin-fixed, paraffin-embedded (FFPE) thymoma samples were used. The samples were sliced into 8-μm-thick sections using a microtome. RNA quality was examined using DV200, and samples with DV200 > 25% were used for all subsequent analyses. Libraries were then constructed using the Visium workflow with CytAssist, according to the manufacturer's guidelines (CG000518, 10× Genomics, Pleasanton, CA, USA). Sequencing was performed at the Research Institute for Microbial Diseases, Osaka University. Libraries were sequenced using an MGI DNBSEQ-G400RS (MGI Tech Co., Shenzhen, China) system. The generated data were processed using Space Ranger v2.0.1 software, using GRCh38-2020-A as a reference.

Visium data analysis

For the assessment of normal thymus tissues, data from eight pediatric thymuses[44] and three fetal thymuses[45] were downloaded (GSE207205 and https://developmental.cellatlas.io/fetal-immune) and processed using Scanpy (1.9.5).[47] Briefly, the data were loaded as anndata objects and concatenated. Spots classified as "in tissue" were retained. Thereafter, we performed normalization (sc.pp.normalize_total), log transformation (sc.pp.log1p) and extraction of HVGs (sc.pp.highly_variable_genes with the options, n_top_genes = 3000, flavor = 'seurat_v3', batch_key = 'sample_id'). We then applied the variational inference model implemented in the scvi-tool (1.0.4).[48] Sample IDs and Projects were specified as categorical covariates and total counts per cell were used as continuous covariates. The model (n_layers = 2, n_latent = 30) was trained using the default parameters and latent space for the UMAP embeddings and Leiden clustering using Scanpy. Marker genes were extracted using the scvi.model.differential_expression function. Gene scores were calculated using the sc.tl.score_genes function implemented in Scanpy with default parameters. Spatial neighborhood enrichment analysis was performed using the sq.gr.spatial_neighbors function implemented in Squidpy (1.3.1).[49] Cell proportions were compared using the Bayesian framework implemented in scCODA.[56] The mixed effect model was implemented using the Python package, statsmodels (v0.14.0). The gene expression comparison between seropositive and seronegative samples was conducted in the medulla and cortex regions where there were sufficient spot numbers. Differences in gene expression were calculated using the sc.tl.rank_genes_groups function (with method = 't test_overestim_var'). Pathway enrichment analysis was performed using the gsePathway function from the R package ReactomePA,[59] targeting genes with a mean corrected expression of 0.1 and log2foldchange. The results were visualized using the dotplot function from the clusterProfiler package.[58]

scDRS-spatial

The GWAS summary statistics deposited at GCST90093061 [1] were used for analysis. These summary statistics describe the meta-analysis results for MG. The cohort included 1,873 cases and 36,370 controls from the US and Italy, respectively. Gene scores were computed using MAGMA[46] (v1.10) software as described by Zhang et al.[13] First, we performed single nucleotide polymorphism (SNP) annotation with gene locations (NCBI37.3, https://ctg.cncr.nl/software/MAGMA/aux_files/NCBI37.3.zip) and the reference data created from 1000 genomics Phase3 (g1000_eur, https://ctg.cncr.nl/software/MAGMA/ref_data/g1000_eur.zip) using magma –annotate (with the option, window = 10,10). Next, we calculated the gene scores from the *p*-values using MAGMA. To include a variety of cell types in the dataset, we downloaded public Visium data (Table S3) and created a Visium control dataset. We then combined these with the thymus datasets. We pre-processed the dataset by normalizing the total counts to the median of the total counts (scanpy.pp.normalize_total), log transformation (scanpy.pp.log1p), and imputing gene expression using MAGIC[17] (scanpy.external.pp.magic). Thereafter, the polygenic enrichment for each cell was evaluated using scdrs compute-score (v1.0.3, options: –flag-filter-data True –flag-raw-count False –*n*-ctrl 1000); the number of genes for each cell was used as the covariate. Group-level statistics were calculated using scdrs perform-downstream and visualized using scdrs.util.plot_group_stats.

A null simulation was performed as described by Zhang et al.[13] We randomly selected 1000 genes 100 times, and the enrichment for the Visium control dataset was evaluated using scdrs compute-score (–flag-filter-data True –flag-raw-count False –*n*-ctrl 1000 for imputed data, –flag-filter-data True –flag-raw-count True –*n*-ctrl 1000 for raw data).

Single-cell RNA-seq analysis

We pre-processed the scRNA-seq data of thymomas generated by Xin et al.[14] First, doublets were removed using Scrublet[52] with default parameters, and cells with >200 and <8000 genes and <20% mitochondrial RNA were retained. The data were then merged with the thymoma and PBMC datasets generated by Yasumizu et al.[6] To remove the effect of immune receptors on highly variable genes, genes related to T cell receptors and B cell receptors were removed. The retained expression was normalized (sc.pp.normalized_total with the option target_sum = 1e4) and transformed (sc.pp.log1p), and highly variable genes were assessed (sc.pp.highly_variable_genes with the options flavor = 'seurat_v3', batch_key = 'project'). Cell cycle was inferred using the sc.tl.score_genes_cell_cycle function following a tutorial (https://nbviewer.jupyter.org/github/theislab/scanpy_usage/blob/master/180209_cell_cycle/cell_cycle.ipynb). The total UMI counts, percentage of mitochondrial genes, S score, and G2M score were regressed using sc.tl.regress_out and scaled using sc.tl.scale. The principal components were then computed using sc.tl.pca. The batch effect of the samples was eliminated using the Harmony algorithm.[53] Neighbors were calculated using sc.pp.neighbors with the options n_neighbors = 30 n_pcs = 50. Cells were embedded in UMAP using sc.tl.umap (spread = 2) and clustered using sc.tl.leiden. The initial layer clusters (cluster L1) were manually defined based on Leiden clusters. For Layer 2 clustering, we recursively extracted cells from a population and performed the same procedures with manually optimized parameters (number of highly variable genes: 1000–3000, number of neighbors: 15–30, n_pcs: 10–50, spread of UMAP: 1). Doublets assigned in subcluster analysis were removed, and the final embedding was generated following the same procedures. For marker gene detections, sc.tl.rank_genes_groups(method = 'wilcoxon') were used.

Cell deconvolution of Visium samples using Cell2location

Cell deconvolution of the Visium samples using Cell2location[12] was performed according to the tutorial guidelines (https://cell2location.readthedocs.io/en/latest/notebooks/cell2location_tutorial.html). The combined scRNA-seq reference without doublets (described below) was filtered (cell2location.util.filtering.filter_genes with the options cell_count_cutoff = 5, cell_percentage_cutoff2 = 0.03, nonz_mean_cutoff = 1.12) and prepared (cell2location.models.RegressionModel.setup_anndata with the options batch_key = 'sample', labels_key = 'clusterL2'). A regression model was created using cell2location.models.RegressionModel and trained (model training with max_epochs = 250). Cell proportions were inferred for each Visium sample at each time point. In the inference step, a model for the Visium sample was created using cell2location.models.Cell2location(N_cells_per_location = 30, detection_alpha = 20) and trained (max_epochs = 30000). Co-localization analysis was performed using cell2location.run_colocation(model_name = 'CoLocatedGroupsSklearnNMF'), and the optimal number of factors was manually selected.

Cell deconvolution of TCGA bulk RNA-seq samples using scaden

Cell deconvolution of TCGA samples was performed using a neural-net-based algorithm, Scaden (v1.1.1), as described by Yasumizu et al.[6] We created 30,000 simulation datasets using a scaden simulate. The count matrices of our single-cell dataset and the TCGA thymoma dataset quantified by HTseq and downloaded from TCGAbiolinks were pre-processed using the scaden process command. Thereafter, the network was trained using the command, scaden train with the option, –steps 5000. Finally, the bulk RNA-seq matrix was deconvoluted using scaden predict. The deconvoluted cell proportion was tested using a multiple linear regression provided as the formula.api.ols function using the Python package statsmodels (0.12.0) with a model, cells ~ MG + WHO + days_to_birth + Gender + 1.

Cell-cell interaction analysis by CellphoneDB

CCI inference was performed using the CellphoneDB[22] framework. Cells with a loading of 0.1 or higher in the NMF-based cell co-localization analysis of Cell2location were used as the microenvironments. CCI inference was performed using the cellphonedb.src.core.methods.cpdb_statistical_analysis_method.call (score_interactions = True, threshold = 0.1) function. The results were visualized using ktplotspy and Scanpy software.

Immunohistochemistry staining and quantification

We randomly selected 10 cases for each group, MG seropositive thymoma and non-MG seronegative thymoma (comprising 5 B1 and 5 B2 cases for each group). Immunohistochemical staining was performed using the DAKO Autostainer Link 48 system (Agilent Technologies, Inc. Santa Clara, CA, USA). Antigen retrieval was performed with EnVision FLEX TARGET RETRIEVAL SOLUTION HIGH pH (Agilent Technologies). Blocking was performed using EnVision FLEX PEROXIDASE-BLOCKING REAGENT (Agilent Technologies). The sections were incubated with an anti-CCR7 antibody (1:800 dilution, clone EPR23192-57, Abcam plc, Trumpington, Cambridge, UK). Then, the sections were incubated with EnVision FLEX/HRP Diaminobenzidine (Agilent Technologies) and treated with EnVision FLEX SUBSTRATE BUFFER + EnVision FLEX DAB+ (Agilent Technologies). SLIDEVIEW VS200 (Evident, Shinjuku, Tokyo, Japan) was utilized to scan the samples at 40× magnification. Subsequently, ImageJ FIJI software was employed to manually remove non-thymoma regions and isolate the tumor areas. The green channel images obtained were subjected to thresholding in the

range of 0–220 to measure the tumor area. Additionally, the red and blue channel images underwent subtraction using Image calculator, followed by thresholding in the range of 10–255 to measure the CCR7+ area.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were performed in R (4.1.2) and Python (3.8.0). FDR was obtained by the Benjamini-Hochberg procedure implemented by a Python package statsmodels (0.12.0). All other statistical analyses are detailed in the respective sections of the article.