High-throughput sequencing of IgH gene in minor salivary glands from Sjögren’s syndrome patients reveals dynamic B cell recirculation between ectopic lymphoid structures

E. Carlotti¹, W. Murray-Brown¹, K. Blighe¹, M. Caliste¹, E. Astorri¹, N. Sutcliffe¹, A.R. Tappuni², C. Pitzalis¹, E. Corsiero¹, M. Bombardieri¹

¹Centre for Experimental Medicine and Rheumatology, William Harvey Research Institute, Queen Mary University of London; ²Department of Oral Medicine, Queen Mary University of London, UK.

Abstract

B cells play a central role in Sjögren’s syndrome (SS) whereby autoreactive B-cells populate ectopic germinal centres (GC) in SS salivary glands (SG) and undergo somatic hypermutation (SHM) and class-switch recombination of the immunoglobulin genes. However, the capacity of specific B cell clones to seed ectopic GC in different SG and undergo clonal diversification is unclear. To unravel the dynamics of B cell recirculation among minor SG biopsies, we investigated the immunoglobulin heavy chain (IgH) gene usage and the pattern of SHM using a high-throughput sequencing approach. We generated ~166,000 reads longer than 350bp and detected 1631 clonotypes across eight samples from four different SS patients, all characterised by the presence of functional ectopic GC as demonstrated by the expression of activation-induced cytidine deaminase. A large number of shared clonotypes were observed among paired mSG biopsies from each patient but not across different patients. Lineage tree analysis revealed significant clonal expansion within the mSG with the identification of shared dominant B cell clones suggestive of extensive recirculation across different SG. Several shared clonotypes with high proliferating capacity displayed IgH-VH gene usage common in autoreactive B cells, including VH1-69, which is typical of rheumatoid factor+ B cells representing potential lymphoma precursors.

The complex dynamic recirculation of B cells that we observed within ectopic GC responses linked with their ability to independently proliferate, undergo ongoing SHM and Ig class-switching within individual glands may explain the difficulty in achieving consistent eradication of ectopic GCs following B cell depleting agents reported in different studies.

Key words

Sjögren’s syndrome, ectopic germinal centres, B cells, next-generation sequencing, clonal diversification
Emanuela Carlotti*, PhD
William Murray-Brown*, PhD
Kevin Blighe, PhD
Mattia Caliste, MD
Elisa Astorri, MD, PhD
Nurhan Satcliffe, MD
Anwar R. Tappuni, BDS, PhD
Costantino Pitzalis**, FRCP, MD, PhD
Elisa Corsiero**, PhD
Michele Bombardieri**, FRCP, MD, PhD

"E. Carlotti and W. Murray-Brown share first authorship.
"C. Pitzalis, E. Corsiero and M. Bombardieri share last authorship.

Please address correspondence to:
Prof Michele Bombardieri
Centre for Experimental Medicine and Rheumatology,
William Harvey Research Institute, Barts and The London School of Medicine & Dentistry,
Queen Mary University of London, John Vane Science Centre, Charterhouse Square, London EC1M 6BQ, UK.
E-mail: m.bombardieri@qmul.ac.uk

Dr Elisa Corsiero
E-mail: e.corsiero@qmul.ac.uk

Received on August 1, 2022; accepted in revised form on October 24, 2022.
© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2022.

Introduction
Sjögren’s syndrome (SS) is an organ-specific and systemic autoimmune disease characterised by chronic inflammation of exocrine glands, particularly salivary and lachrymal glands and parotid gland enlargement (1). It has a prevalence of 0.2-1% in the adult population (UK 100,000-500,000) (1). Approximately, 5% of SS patients can develop lymphoid malignancies, the majority (80%) being low-grade B cell mucosa associated lymphoid tissue lymphomas (MALT-L) (2, 3). Primary SS (pSS) presents alone whilst secondary SS (sSS) occurs in conjunction with another autoimmune disease such as systemic lupus erythematosus, scleroderma or rheumatoid arthritis (4, 5).

SS salivary glands (SG) are characterised by infiltration of mononuclear cells such as T (predominantly CD4) and B cells, macrophages, plasma cells and dendritic cells. Together with inflammatory infiltrates, autoantibodies are considered a hallmark of the disease. Autoreactivity against ribonucleoproteins Ro/SSA and La/SSB are observed in 50-80% of SS patients (6, 7). Moreover, between 30-40% of SS patients develop ectopic lymphoid structures (ELS) (8-10) displaying germinal centre (GC) features, such as the expression of the enzyme activation-induced cytidine deaminase (AID) required for the somatic hypermutation (SHM) and class switch recombination of expressed Ig genes (11). Together with the risk of MALT lymphoma the recent observation that mutations in lymphoma driver genes in conjunction with SHM of the immunoglobulin genes can lead to pathogenic antibodies have suggested that B cells might play a primary role in the immune-pathogenesis of SS (12, 13).

Overall, we know that the representation of B cell subpopulations in pSS is skewed with increased IgM+ memory B cells (14, 15) and CD5+ B cells. In addition, perturbation in the follicular helper T cells population (Tfh), highly specialised T cells, present in the GC lymphoid structures, that play a key role in B cells maturation, seem to be altered in SS patients (16) Finally, clinical studies which have investigated the use of the anti-CD20 monoclonal antibody Rituximab for targeting B cells in SS patients have yielded conflicting results in terms of clinical response with only partial effectiveness in SG B cell depletion and modulation of ELS (17) highlighting our incomplete understanding of the mechanisms regulating B cell recirculation in ectopic GC responses.

Previous analysis of the Ig VH gene B cell repertoire in SS have shown that the repertoire is different depending on the anatomical site investigated, with an enrichment of memory B cells in the SGs as compared to the peripheral blood (14, 18). The analysis of IgM antibodies from lymph-nodes, blood and SGs in a mouse model of pSS (Id3−/−) showed that the SG repertoire displays unique characteristics that are likely to contribute to the disease in a very specific manner compared to draining lymph-nodes and the peripheral compartment (19). The introduction of high-throughput sequencing technologies in the analysis of the immunoglobulin gene has unveiled the B cell diversity at an unprecedented depth, allowing enlightening the dynamics of circulation of B cells in the setting of haematological malignancies and in other autoimmune diseases (20, 21). Therefore, in this work we performed next-generation sequencing (NGS) and analysis of the IgH-VH genes in different ELS+ SG biopsies taking from the same patients in order to define the capacity of B cells to recirculate and enter independent ectopic GC responses across different SG.

Material and methods

Patient samples
All labial salivary glands biopsies were obtained for diagnostic purposes, as approved by the National Research Ethics Service Committee London-Westminster (LREC 05/O0702/1), from patients attending the Rheumatology/Oral Medicine Sjögren’s Clinic at Barts Health NHS Trust. In total four pairs of minor salivary gland (mSG) biopsies, collected at the same time, from four different patients (3 females and 1 male, average age 64) were included in the study. All the patients had

Funding. This work was supported by MRC project grants MR/N003063/1 and MR/T016736/1 to M. Bombardieri, Versus Arthritis Grant 21268 to M. Bombardieri and E. Corsiero.

Competing interests: none declared.
a diagnosis of primary SS at the time of the biopsy, but two of them (pt1 and pt3) were subsequently diagnosed with secondary SS. Additional clinical details are provided in Supplementary Table S1. Patients were selected on the basis of the availability of RNA extracted from matched but separated mSG biopsies collected at the same time, the presence in both the mSG of ELS and of a high B cell infiltration, as detected by immunohistochemistry using antibodies targeting CD20 (B cells), CD138 (plasma cells), CD3 (T cells) and CD21 (FDCs) (9, 22) and the expression of AICDA gene required for germinal centre reaction (Suppl. Fig. S1a). The IHC analysis was performed on a third matched mSG, different from the pair used for the high-throughput sequencing analysis and the RQ-PCR study.

Preparation of the libraries and pyrosequencing
Total RNA was extracted from mSG tissue using the RNeasy Mini Kit (Qiagen); the concentration/purity of RNA samples were measured using the Na- nodrop 20C (Lab Tech, UK) and the quality was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, UK). To enrich for the IgV-VH genes, cDNA was synthesised using specific primers designed on the constant Ig-μ, - γ and - α domains (23) (Supplementary Table S2) using 100 ng of total RNA and the Superscript III Reverse Transcriptase (Invitrogen) fol-
following the manufacturer’s instructions. Thirty-six amplicon libraries were generated per sample, using the whole retro-transcribed RNA, 12 VH consensus forward primers, 3 Iγ-μ-, γ and α reverse primers (23) modified with a unique multiplex identifier (MID) tag (Suppl. Table S2) and the Q5 High Fidelity DNA polymerase (NEB) according to the following conditions: a denaturation at 98°C for 1’ followed by 33 cycles (for γ-μ isotypes) or 36 cycles (for α isotype) of 98°C for 30” and 72°C for 45” followed by a final extension of 72°C for 2’. Amplicons were purified using the Agentcourt AMPure beads kit (Beckam Coulter), quantified using the Quanti-iT PicoGreen Assay kit (Thermo Fischer scientific), pooled and then sequenced using the Roche GS-FLX titanium platform, following the manufacturer’s instructions for the Titanium series (454 Life Science, Roche) (Supplementary Fig. S2a).

Data analysis
The sequential steps involved in the analysis of 454 reads are shown in Supplementary Figure S2b. Briefly, sequences were exported using 2 different pipelines: shotgun (designed for libraries prepared using a shotgun strategy) (dataset A) and Long Amplicon 2 (LAMP2) (designed for long-range PCR experiments) (Dataset B). We then merged the reads belonging to unique clonotypes from both the data sets and generated the Dataset A1, used for our downstream analysis. After a quality control filtering, only reads ≥ 350 bases were included in the analysis (Suppl. Fig. S2c). The IgH VDJ reads were identified using IMGT/HighV-QUEST (http://www.imgt.org/) and the different VDJ combinations were determined using the IgAT software (Rogosh, Front Imm 2012). IgH VH-JH rearrangements across the whole reads were plotted using Circos software (24). Sequences were aligned with ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and lineage trees generated using IgTree (25) (kindly provided by Dr Ramit Mehr, Bar-Ilan University, Ramat-Gan, Israel) and Cytoscape (http://www.cytoscape.org) (25). Clonotypes were defined as reads

Fig. 2. Circos plots of the VH-JH gene recombination present in the 4 paired mSG libraries. Only VH and JH genes belonging to clonotypes with more than 10 reads were included in this analysis. For each plot the width of each ribbon is proportional to the frequency of each VH-JH rearrangement detected. JH genes are all shown in black whilst the different VH families are represented with different colours: yellow VH1, green VH3, blue VH4, red VH5 and purple VH6.
with the same VH-DH-JH gene usage and equal CDR3 length. Clones instead were defined as reads, identified by at least one forward and one reverse primer, belonging to the same clonotype, carrying the same CDR3 but displaying a unique SHM pattern (Suppl. Table S3). The analysis was performed across the whole VDJ region, using the sequences registered in the IMGT database as reference; the mutation profile in N1 and N2 regions was arbitrarily defined according to the level of conservation of the nucleotides: the less conserved base identified after aligning different clones belonging to the same clonotype was considered a mutation. Reads with stop codons or frame shift mutations were excluded from the analysis.

**Results**

**High-throughput sequencing data and VH-DH-JH analysis**

After exporting and filtering the reads from DatasetA1 for sequences longer than 350 bases we obtained ~166,000 reads (range 14,244 – 29,737, pt2B and pt2A respectively) (Fig. 1a and Supplementary Fig. S2c). IgAT showed similar VH family usage patterns across the paired biopsies, with VH3, VH4 and VH1 gene families being the most detected (VH3 most rearranged in the samples 1B, 2A, 3A, 3B and VH4 in the biopsies 1A, 2B, 4A, 4B) (Suppl. Fig. S3a). Similar consistent results were observed for JH genes, for which we confirmed JH4 family to be the most rearranged across all the samples, followed by JH5 and JH3 (Supplementary Fig. S3c). D3 was the gene DH family most detected across all the samples but mSG 4B, with DH3-10 and DH3-22 being the most represented genes (Supplementary Fig. S3b). With the exception of patient 4, which had a dominant presence of γ isotype reads in both biopsies (68% in sample A and 56% in sample B), the most prevalent constant region detected was μ range detection: 48% in Pt2B and 49% in Pt3B) (Fig. 1b). When we extended our analysis to individual VH genes.
we observed some variability not only across patients but also between paired biopsies, with VH3-23, VH3-30 and VH4-39 being the most represented (Fig. 1c). Interestingly VH1-69 gene, frequently detected in MALT from SG (27) and in RF positive cells (28) was detected (at a prevalence >1%) in both biopsies from Pt1, Pt2 and Pt3 but in none from Pt4, whereas VH4-34 gene, usually associated with features of autoreactivity (29) was observed in both the mSGs of Pt2, Pt3 and Pt4, but in none of the samples from Pt1. By using Circos plots (Fig. 2) we clearly confirmed the preferential rearrangement of VH3/VH4 and JH4 and highlighted preferential VH-JH combinations. Overall, we observed that VH4-39 gene, the most represented in all our libraries, rearranged preferentially with JH4 and JH5, VH3-23 with JH4 and VH3-30 mainly with JH4, but also with JH5, JH6 and JH1. JH1, JH2 and JH3 genes did not seem to rearrange preferentially with any VH family. VH4-34 and VH1-69 rearranged not only with JH4 but also with JH5 and JH6.

Identification and analysis of clonotypes

We then investigated the number of clonotypes (reads with the same VH-DH-JH genes and identical CDR3 length) identified in each library. In total we detected 6331 clonotypes (range 452 - 1354, libraries SSA1A and SS3A respectively). We observed a high level of variation in the number of clonotypes detected across the different mSGs: e.g., pt2 mSG-A 1352 clonotypes and mSG-B 531 observed; interestingly, when we selected the clonotypes with ≥10 reads, only an average of 26.5% (range 19.3% – 32.3%, mSG libraries 4A and 3A respectively) and a total of 1631 clonotypes were carried over for further analysis (Fig. 3a). When we looked at the prevalence of these clonotypes and we observed a difference in number and distribution across the different libraries (Table IV and Supplementary Fig. S4). We found a statistically significant correlation between the number of reads identified and all the clonotypes detected (Pearson, p=0.02), the number of reads and the number of clonotypes

![Fig. 4. BCR clonotype tracking in the 4 paired mSGs.](image-url)

The clonotype composition is represented by stacked bar plots, which are coloured according to IGHV families. Only clonotypes with clonal size ≥3 are plotted, and proportion of BCRs in these clonotypes are compared between paired mSGs. The graphs show the most abundant clonotypes observed between paired mSGs (SSA and SSB) in a) Pt1, b) Pt2, c) Pt3 and d) Pt4.
with ≥10 reads (Pearson, p=0.007) and the number of total clonotypes as compared to the clonotypes with ≥10 reads (Pearson, p=0.0026).

We also examined the clonality of BCRs based on unique VDJ sequences for the paired mSGs in each patient (Fig. 4). In total, 93 clonotypes were common in the repertoires isolated from Pt1, 122 for Pt2, 56 for Pt3 and 31 for Pt4.

Lineage tree analysis and patterns of circulations of B cells across matched mSGs

In order to understand whether B cells clones re-circulate across different mSG we focused our analysis on those clonotypes that were shared between the two matched biopsies. We identified 22 shared clonotypes across the paired biopsies and the majority (16/22) shared the same VH-DH-JH rearrangement but displayed two or more isotypes (Table I).

The number of unique clones belonging to each clone ranged from 2 (detected in pt1, clonotype VH3-23/DH2-2/JH3-02; pt3, clonotype VH4-39/DH4-11/JH5-01; pt4, clonotype VH3-09/DH3-22/JH4-02) to 28 (pt2, clonotype VH4-39/DH5-17/JH4-02) (Table I). Interestingly, three of the shared clonotypes displayed the VH1-69 and one the VH4-34 gene rearrangements; most of this clonotypes (3/4) were observed in the shared libraries from pt2. Based on the pattern of SHM and the isotype detected in the B cells clones, we identified three different patterns of circulation between the two paired mSGs. In 12 clonotypes we observed a unidirectional pattern of circulation with clones moving from the mSG-A to the mSG-B (in both pt1 and pt3) and B cell clones belonging to 1 clonotype (pt1) and 2 clonotypes (pt3) moving in the opposite direction (Table I). We then investigated the dynamics of B cells differentiation and clonal expansion within each patient by analysing the pattern of SHM and switch recombination across all the different clonotypes by constructing phylogenetic trees. As exemplified in Figure 5, we frequently observed highly expanded and profoundly hypermutated B cell clones which were dominant in both mSG biopsies from the same patient. Interestingly, across different SGs we identified B cell clones undergoing independent rounds of SHM either as unswitched IgM+ clones or following class-switching to IgG or IgA (Fig. 5 and Supplementary Fig. S5). In most of the clonotypes, more than one of these profiles were simultaneously present within the same lineage tree.

**Discussion**

The pathogenic role of B cells in SS is a hallmark of the disease and goes beyond the mere presence of circulating autoantibodies such as anti-Ro/SSA and anti-La/SSB and rheumatoid

Clinical and Experimental Rheumatology 2022
Fig. 5. Lineage trees of one representative clonotype shared between paired mSGs. The clonal relationship across the different clones detected in each clonotypes was determined on the CDR1-CDR3 region and using IgTree; lineage tree was then draw using Cytoscape. Each node represents a clone. The size of each node is proportional to the frequency of the number of reads identified for that particular clonotype. Black circles = germline sequence (as for IMGT). White circles are progenitor cells, that based on the pattern of SHM are inferred to exist but that have not been detected in our NGS experiment. The isotypes μ, α, and γ are represented as partitions of the circles, whose size is proportional to the frequency of each individual isotype. One representative clone (VH4-39/D5-12/JH5-02) shared between paired mSGs (Biopsy A and B) is shown. In this example, both SHM and switch recombination occur. B cell clones, with same SHM but with μ (blue/purple) and α isotypes (green/light blue), migrated from the mSG biopsy A to mSG biopsy B, where they continue the SHM process but also starts to switch into γ (orange).

factor (RF) (30). SS patients are characterised by profound alterations in peripheral B-cell sub-populations with increase in the proportion of naïve B cells and a decrease in the circulating memory compartment, which is partly explained by the preferential accumulation of CD27+ memory B cells in the SG (14, 18). In advanced SG lesions B cells predominate and in around 30-40% of SS patients are organised in ELS, characterised by aggregates of segregated B and T cells with differentiation of FDC networks sustaining an ectopic GC response with perifollicular plasma cells accumulation (31). While the mechanisms of B cell recruitment and positioning within the SG of SS patients with ELS are largely related to the ectopic expression of lymphoid chemokines CXCL13 and CXCL12 which bind to their specific receptors CXCR5 and CXCR4, respectively, on B cells (32), little is known about the capacity of B cells once activated and diversified in the SG to recirculate and further seed other SG. A recent NGS single cells analysis of selected T cells has shown the existence of expanded CD4+ T clones in the SGs of SS patients that seem to recognise gland specific antigens (33). In the present work, we demonstrate that B cells in SS have the capacity to extensively recirculate to different glands. Specifically, by clonotype tracking and lineage tree analysis of shared clonotypes present in different SG of the same SS patients, we identified dominant B cell clones shared between different mSG and undergoing clonal expansion and diversification through different cycles of SHM and proliferation. These results suggest extensive recirculation of expanded B cell clones across different mSG through the lymphatic system with re-entering via the blood circulation. However, the difficulty in identifying a common precursor and the lack of consensual biopsy of a draining lymph node cannot exclude the possibility that progenitor clones migrated to different mSG gave rise to independent clonal expansion. This notion is in line with the evidence of extensive lymphatic neo-vascularization in the SG of SS patients with a focus score >1 (34). Another novel observation of this study is the evidence through lineage tree analysis that B cells not only can extensively recirculate between ectopic GCs in separate SG from the same patient but can also undergo independent rounds of Ig VH genes somatic hypermutation giving rise to highly proliferative clones which can represent up to 10% of the overall diversity in a given SG. This evidence suggests continuous shuffling of B cell clones in/out of ectopic GCs whereby for some B cells the whole process of affinity maturation took place within the SG tissue. This observation is in line with the known capacity of ELS to support a functional GC response trough the induction of AID, the enzyme responsible for SHM and class switch recombination of the Ig genes and extend previous data reporting highly somatically hypermutated B cells in SG with ectopic GCs (35, 36).

Our clonotypes analysis showed that VH3-23, VH4-39 and VH1-69 were the most commonly observed family usages among the top 10 dominant clones shared by different SG from the same patient. Interestingly, VH1-69 usage is typically expressed by RF-producing B cells which have been reported to frequently bear lymphoma driver mutations in the circulation of SS patients with cryoglobulinaemic vasculitis (13), and are known to harbour within premalignant lymphoepithelial lesions (36) and to give rise to malignant B cell clones in almost 50% of SG MALT-lymphomas (37). Generally, the highly proliferative VH1-69 B cell clones that we identified in the SG displayed profoundly somatically hypermutated Ig genes and preferential _isotype usage which is typically associated with SG MALT-L originating from IgM+ memory B cells (38). The above evidence provides novel insights, albeit indirect, in support of the previously reported association between ectopic GCs and the evolution to MALT-L in SS patients. While we were
the first to report a higher prevalence of ELS in labial SG biopsies of SS patients who later developed parotid MALT-L (11) subsequent work estimated that the presence of ectopic GCs confers between 8- to 16-fold higher risk of developing MALT-L (10, 39) although controversial results have also been reported (40-42).

In addition to VH1-69, among highly proliferating and hypermutated shared clonotypes recirculating in different SG we frequently observed other VDJ rearrangements, such as VH4-34, used by inherently autoreactive B cells (43), supporting the evidence that ELS in SS patients are unable to exert the control mechanisms of follicular exclusion of autoreactive B cells allowing their entrance in ectopic GC and their local differentiation in autoantibody producing cells (44) resulting in the accumulation of anti-Ro/SSA and anti-La/SSB reactive perifollicular plasma cells (8, 45). The well described association between ELS and circulating autoantibodies reported in several studies in SS (46, 47) further suggest that local autoantibody production in the SG is reflected, at least partially, in the peripheral compartment.

The combined analysis of IgH-VH gene with Ig isotype usage that we performed also allowed complete reconstruction of the extent by which infiltrating B cell clones are driven to switch isotype class within the SG microenvironment. In agreement with previous work, we demonstrated that mSG B cells are highly enriched in hypermutated IgM+ B cells which have been previously described to bear a human transitional/marginal zone-like phenotype and be primarily responsible for the evolution towards MALT-lymphoma (48). The extensive clonal expansion and intra-clonal diversification of both IgM+ and isotype switched hypermutated B cells that we observed in the SG is in keeping with the notion that both subsets of memory B cells can re-enter the GC response and acquire further rounds of hypermutation and affinity maturation, with switched memory B cells considered more likely to differentiate into plasma cells (49, 50). This could be of significant relevance as a reduction in IgA+ plasma cells associated with a relative increase in IgG+ plasma cells has been described in the SG of SS compared to controls with a threshold of <70% IgA indicated as superior to the FS in diagnostic accuracy.

In summary in this work, by using a NGS strategy to capture the complexity of the B cells infiltrating different labial SG tissues within the same patients we demonstrated a highly dynamic recirculation of B cells within ectopic GC responses. The capacity of these B cells to infiltrate, hypermutate and class switch their Ig genes and extensively proliferate provide novel insights on the intricacy of local B cell responses within SG with ELS. These processes appear driven by locally delivered stimuli and antigens though BCR engagement providing strong survival signals which may partially explain the observed resistance in eradicating ectopic GCs responses upon treatment with B cell depleting agents.

Acknowledgements

We thank the patients who donated tissue samples to this project and Ramit Mehr at Bar-Ilan University, Ramat-Gan, Israel for providing IgTree.

References


B-cells recirculation in Sjögren’s syndrome / E. Carlotti et al.


42. PAPE KA, TAYLOR JJ, MAUL RW, GEARHART PJ, JENNINGS MK: Different B cell populations mediate early and late memory during an agenogenic immune response. Science 2011; 331(6021): 1203-7. https://doi.org/10.1126/science.1201730


Clinical and Experimental Rheumatology 2022