



OPEN

## Gene expression alterations in salivary gland epithelia of Sjögren's syndrome patients are associated with clinical and histopathological manifestations

Ariana Dela Cruz<sup>1,7</sup>, Vinay Kartha<sup>2,7</sup>, Andrew Tilston-Lunel<sup>3,7</sup>, Rongjuan Mi<sup>1,3</sup>, Taylor L. Reynolds<sup>4</sup>, Michael Mingueneau<sup>4</sup>, Stefano Monti<sup>2</sup>, Janicke L. Jensen<sup>5</sup>, Kathrine Skarstein<sup>6</sup>, Xaralabos Varelas<sup>3</sup>✉ & Maria A. Kukuruzinska<sup>1</sup>✉

Sjögren's syndrome (SS) is a complex autoimmune disease associated with lymphocytic infiltration and secretory dysfunction of salivary and lacrimal glands. Although the etiology of SS remains unclear, evidence suggests that epithelial damage of the glands elicits immune and fibrotic responses in SS. To define molecular changes underlying epithelial tissue damage in SS, we laser capture microdissected (LCM) labial salivary gland epithelia from 8 SS and 8 non-SS controls for analysis by RNA sequencing (RNAseq). Computational interrogation of gene expression signatures revealed that, in addition to a division of SS and non-SS samples, there was a potential intermediate state overlapping clustering of SS and non-SS samples. Differential expression analysis uncovered signaling events likely associated with distinct SS pathogenesis. Notable signals included the enrichment of IFN- $\gamma$  and JAK/STAT-regulated genes, and the induction of genes encoding secreted factors, such as LTF, BMP3, and MMP7, implicated in immune responses, matrix remodeling and tissue destruction. Identification of gene expression signatures of salivary epithelia associated with mixed clinical and histopathological characteristics suggests that SS pathology may be defined by distinct molecular subtypes. We conclude that gene expression changes arising in the damaged salivary epithelia may offer novel insights into the signals contributing to SS development and progression.

Sjögren's syndrome (SS) is a debilitating complex autoimmune disease presenting as exocrinopathy of salivary and lacrimal glands with frequent systemic extraglandular manifestations and an increased risk of non-Hodgkin's lymphoma<sup>1–6</sup>. Damage to salivary and lacrimal glands from SS impairs their ability to produce saliva and tears, resulting in xerostomia and xerophthalmia. SS is a multifactorial disease that comprises various combinations of dry eyes, dry mouth, reduced tear and saliva secretion, presence of autoantibodies, focal lymphocytic infiltration of salivary glands, fatigue, vasculitis, joint and muscle pain, as well as peripheral nervous system dysfunction<sup>4,7</sup>. Although to date, the origin of SS remains unknown, the prevailing model for SS has been that loss of secretory function is secondary to lymphocytic infiltrates.

Since defects in the immune system have long been considered to be the cause of SS, the majority of studies and therapeutic approaches have focused on aberrant immune responses<sup>8–10</sup>. Analysis of genomic and epigenomic changes in SS has provided new insight into the understanding of disease pathogenesis<sup>9,11–13</sup>. Gene

<sup>1</sup>Department of Translational Dental Medicine, Boston University School of Dental Medicine, Boston, USA. <sup>2</sup>Department of Medicine, Boston University School of Medicine, Boston, USA. <sup>3</sup>Department of Biochemistry, Boston University School of Medicine, Boston, USA. <sup>4</sup>Immunology Research, Biogen Idec, Cambridge, MA, USA. <sup>5</sup>Faculty of Dentistry, University of Oslo, Oslo, Norway. <sup>6</sup>Department of Clinical Medicine, University of Bergen, Bergen, Norway. <sup>7</sup>These authors contributed equally: Ariana Dela Cruz, Vinay Kartha and Andrew Tilston-Lunel. ✉email: xvarelas@bu.edu; mkukuruz@bu.edu

expression microarray studies on labial salivary glands and peripheral blood detected dysregulation of type 1 interferon (IFN)-inducible genes<sup>14</sup>. Additional upregulated genes involved in antigen presentation, lymphocyte development and activation, as well as interferon-induced chemokines were also identified. Interestingly, IFN- $\gamma$ -inducible genes, such as Signal transducer and activator of transcription (STAT) family members STAT1 and STAT3, were also shown to be highly expressed in primary SS (pSS)<sup>15</sup>. Furthermore, IFN-mediated innate immune mechanisms have been implicated in the pathogenesis of pSS, where specifically 23 genes known to play a role in IFN signaling were identified<sup>16</sup>. Two GWAS studies have been performed in pSS and among the SS-associated non-HLA genes they identified STAT4 and IRF5 encoding transcription factors, BLK coding for B cell kinase as well as genes encoding the IL-12A cytokine. Genes involved in NF- $\kappa$ B signaling and CXCR5 chemokine production were also shown to be upregulated<sup>9,17</sup>. Collectively, these studies provided a strong support for the role of innate immunity, as well as adaptive immune mechanisms, in the pathogenesis of pSS. Despite these advances, treatment strategies for SS disease are almost inexistent, and molecular mechanisms that drive disease onset are poorly understood.

One consistent feature of human autoimmunity is the finding that there is an initial phase in which patients have many of the serological abnormalities of autoimmune disease before they develop clinical pathology<sup>18</sup>. In this respect, the immunobiology of the glandular epithelial cells within the target tissue likely plays a critical role in initiating events in SS. Indeed, accumulating evidence suggests that impaired glandular secretion develops independent of lymphocytic infiltration, and defective secretion precedes the immune response during SS onset, suggesting a role for the epithelium in the etiology of SS<sup>2</sup>. For example, aberrantly high expression of BMP growth factors, such as BMP6, has been implicated in the induction of salivary and lacrimal gland dysfunction independent of the autoantibodies and immune infiltration associated with SS<sup>19</sup>. Recent studies, including our own, have also shown that the structural integrity of the epithelium of SS patients is lost, and that these defects are associated with aberrant apical-basal polarity<sup>20–25</sup>. Aberrant polarity can arise in various ways, including environmental toxins, viral infection, physical damage, and unbalanced hormone levels, and therefore it may be an early event that triggers disease states, including SS. Since structural changes of the epithelium of salivary glands may drive inflammation, the aim of this study was to elucidate molecular details of how epithelial structural dysfunction contributes to SS. We used laser capture microdissection of glandular epithelial regions (acini and ducts) from resected labial salivary glands of patients fulfilling or not fulfilling SS criteria (8 SS and 8 non-SS controls) and processed them for global transcriptomes using bulk RNAseq analysis. Our studies provide evidence that damage to salivary gland epithelia is associated with changes in gene expression signatures that provide insights into deregulated pathways and cellular processes which may underlie predisposition to, and early onset of, SS. Collectively, these findings provide support for the theory that alterations in gene expression arising in the salivary epithelia contribute to the etiology of SS. We postulate that the molecular sub-groups reflect varying stages of disease predisposition and that they may offer novel insights into the signals contributing to the progression of SS.

## Results

**Clinical characteristics of patients.** Sixteen samples of labial salivary glands were obtained from female patients who presented with sicca symptoms to the Department of Otolaryngology/Head and Neck Surgery at Haukeland University Hospital, Bergen. Patients were assessed for primary Sjögren's Syndrome (pSS) based on the American-European Classification Criteria<sup>26</sup>, where diagnosis of pSS required focal lymphocytic sialadenitis score (FS) of  $\geq 1$  or presence of anti-SSA/Ro or anti-SSB/La autoantibodies, as well as ocular symptoms of dry eyes, oral symptoms of dry mouth, positive Schirmer's test, and reduced unstimulated salivary flow  $\leq 1.5$  mL in 15 min (at least four positive tests including symptoms, or three objective tests). We categorized samples into two clinically distinct groups, patients diagnosed with SS (named SS,  $n = 8$ ) or without SS (named non-SS,  $n = 8$ ). Clinical characteristics of patients are shown in Table 1. The patient mean age was  $53 \pm 12.4$  years, with three patients in the SS group presenting with anti-SSA/Ro and ANA antibodies, of which also two were positive for anti-SSB/La antibodies. All patients reported dry eyes and the majority reported dry mouth. Three SS patients had FS of 1, three had a FS of 2, with one patient presenting with FS of 3 and one with FS of 4. Interestingly, FS of either 1 or 4 were associated with positivity for all anti-SSA/Ro, anti-SSB/La and ANA, suggesting that these may be independent variables in our sample group. Likewise, there was no association between germinal centers (GCs) and autoantibody positivity. Furthermore, two of the non-SS patients had a negligible salivary flow rate, and there was little correlation between salivary flow rate and auto-antibody positivity in SS patients.

Histopathological evaluation of H&E sections of biopsies from non-SS and SS patients revealed that SS samples were characterized by acinar atrophy, ductal dilation, detectable fat infiltration, and lymphocytic aggregates with FS  $\geq 1$ , whereas non-SS tissues were generally normal in appearance with some acinar atrophy, ductal dilation, and scattered inflammation (Fig. 1). From analysis of clinical presentation, we observed that there was no correlation between SS and non-SS groups for the parameters of positive Schirmer's test, saliva secretion, dry mouth or dry eyes. However, the presence of anti-SSA/Ro, anti-SSB/La, ANA autoantibodies, FS  $\geq 1$ , and the presence of CD21+ GCs were only observed in the SS group. When assessed for age, the SS group was significantly older than the non-SS group ( $60.2 \pm 10.7$  years vs  $45.6 \pm 9.6$  years respectively;  $p = 0.012$ ).

**Subclustering of epithelial-derived RNA-seq data reveals potential SS subgroups.** To gain insight into the molecular characteristics of two groups with sicca symptoms, eight non-SS and eight SS, all samples were subjected to laser capture microdissection (LCM) of epithelial regions followed by RNAseq analysis to identify biomarkers restricted to the epithelial compartment of the labial salivary glands. We evaluated the entire set by principal component analysis (PCA) selected by the expression level of all differentially expressed (DE) genes. It was evident that the SS and non-SS samples did not entirely cluster together and rather three distinct transcriptomic subclusters were observed: a group containing non-SS samples (named Group A,  $n = 3$ ), a group

Clinical classification	Patient	Age	Gender	Anti-SSA/Ro	Anti-SSB/La	ANA	Schirmer test	Saliva secretion (ml/15 min)	Dry mouth	Dry eyes	FS	Germinal Centers (CD21 <sup>+</sup> )
Non-SS	Non-SS1	45	F	-	-	-	+	0.5	+	+	-	-
	Non-SS2	53	F	-	-	-	-	0.0	+	+	-	-
	Non-SS3	51	F	-	-	-	-	0.0	+	+	-	-
	Non-SS4	51	F	-	-	-	+	3.2	+	+	-	-
	Non-SS5	34	F	-	-	-	+	1.0	+	+	-	-
	Non-SS6	59	F	-	-	-	+	2.6	+	+	-	-
	Non-SS7	32	F	-	-	-	+	4.0	-	+	-	-
	Non-SS8	40	F	-	-	-	+	2.1	+	+	-	-
SS	SS1	62	F	+	+	+	-	1.7	+	+	1	-
	SS2	74	F	-	-	-	+	0.9	+	+	2	+
	SS3	51	F	-	-	-	+	1.7	+	+	1	-
	SS4	56	F	-	-	-	+	2.8	+	+	1	-
	SS5	53	F	-	-	-	+	1.8	+	+	2	+
	SS6	79	F	+	+	+	NT	1.0	+	+	4	+
	SS7	55	F	-	-	-	+	4.4	+	+	2	+
	SS8	52	F	+	-	+	+	6.6	-	+	3	+

**Table 1.** Clinical characteristics of patients whose salivary gland biopsies were analyzed by RNAseq. ANA antinuclear antibodies; Saliva secretion (unstimulated flow); normal flow > 1.5 ml/15 min; Schirmer test positive when reduced tear secretion (objective test); Dry mouth and dry eyes, subjective symptoms from the patient, FS Focus Score, NT not determined. Germinal center-like structures may result from the deregulated proliferation and infiltration of lymphocytes in glandular tissue.

containing SS (named Group B,  $n = 3$ ), and a group containing samples from both groups (named Group AB,  $n = 10$ ) (Fig. 2a,b). The position of these groupings demonstrates different features of RNA expression in the epithelial tissues, suggesting that transcriptomic profiling of patient samples may characterize important molecular events in SS pathogenesis compared to grouping by clinical characteristics alone.

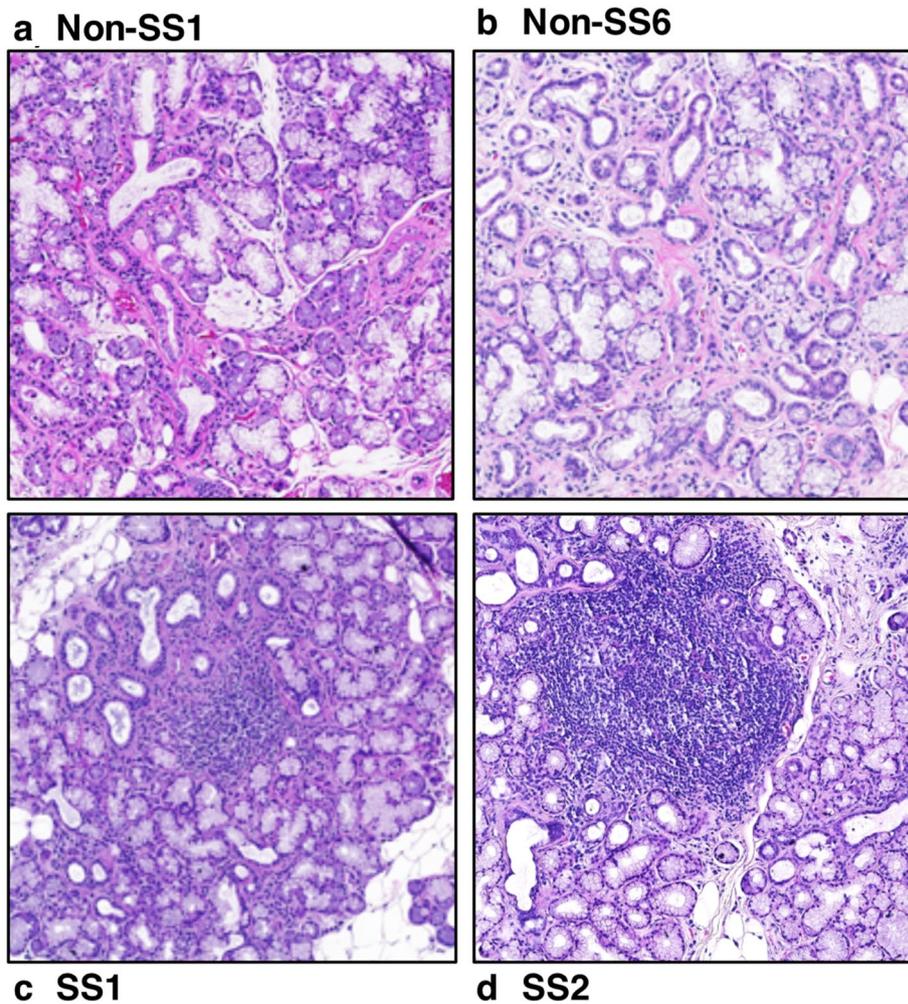
When assessed for age, Group B ( $69.3 \pm 12.7$  years) was significantly older than both the Group AB ( $52.2 \pm 7.4$  years;  $p = 0.020$ ) and Group A ( $39.0 \pm 6.6$  years respectively;  $p = 0.002$ ), whereas there was no significant age difference between Group A and Group AB. Clinical findings for the RNA-Seq groupings revealed that all samples in Group B had GCs with FS of  $\geq 2$ . One of the patients in this group developed lymphoma during the course of this study. In the mixed Group AB, two of the five SS patients also had GCs and lower mean FS. There was no correlation observed for RNAseq grouping for the presence of anti-SSA/Ro, anti-SSB/La and ANA autoantibodies, salivary flow and Schirmer test. Likewise, no specific pattern regarding Group A patients was detected, except that all presented with positive Schirmer tests and a subjective feeling of dry eyes.

**Differential expression testing of SS subgroups.** Genome-wide differential expression testing was performed using DESeq2 for multiple group comparisons. First, we tested the clinical grouping of SS against non-SS. Second, we compared the RNAseq grouping of Group B against Group A samples to determine if there were differences in gene expression between clinical diagnosis of Sjögren's and principal component analysis (PCA) for clustering. Third, the RNAseq Group A and Group B groups were compared to the Group AB.

RNAseq analysis of the epithelial regions for the groupings clinically diagnosed with SS showed no significantly altered differentially expressed (DE) genes compared to the non-SS group out of 25,444 total genes analyzed (Fig. 2c). However, when analyzed according to PCA clustering, there were 1247 DE genes between Group B and Group A at a FDR cut-off of 0.05 (Fig. 2c), further demonstrating a benefit of clustering the samples according to RNAseq data. Table 2 summarizes the top 10 DE upregulated and downregulated genes for both clinical Sjögren's (SS vs non-SS) and RNAseq (Group B vs Group A vs Group AB) groupings.

Inter-group analysis with the Group AB subcluster revealed the following comparisons: 767 DE genes compared to the Group A (FDR cut-off = 0.05) and 1480 DE genes compared to Group B (FDR cut-off = 0.05). While strictly not statistically meaningful, the FDR-value ranking was used to define signatures to be annotated by pathway enrichment analysis.

**Pathway analysis identifies epithelial signals associated with SS.** Gene set enrichment analysis (GSEA) was conducted for each of the groups, and hallmark gene sets were evaluated (Fig. 3a). Comparison of SS vs non-SS groupings revealed 8 signaling pathways that were statistically upregulated and 27 pathways downregulated in the SS group (Supplementary Table S1). Notably, immunological and autoimmune disease pathways were enriched in the SS group while cholesterol and protein-related metabolic pathways were downregulated (Fig. 3b,c). Interferon response genes were upregulated in both the clinical SS and RNAseq (Group B) groupings compared to non-SS and Group A samples, respectively, suggesting conservation of relevant processes reflected in RNA profiling. Further hallmark GSEA of comparisons between the subcluster groupings revealed 9 upregu-



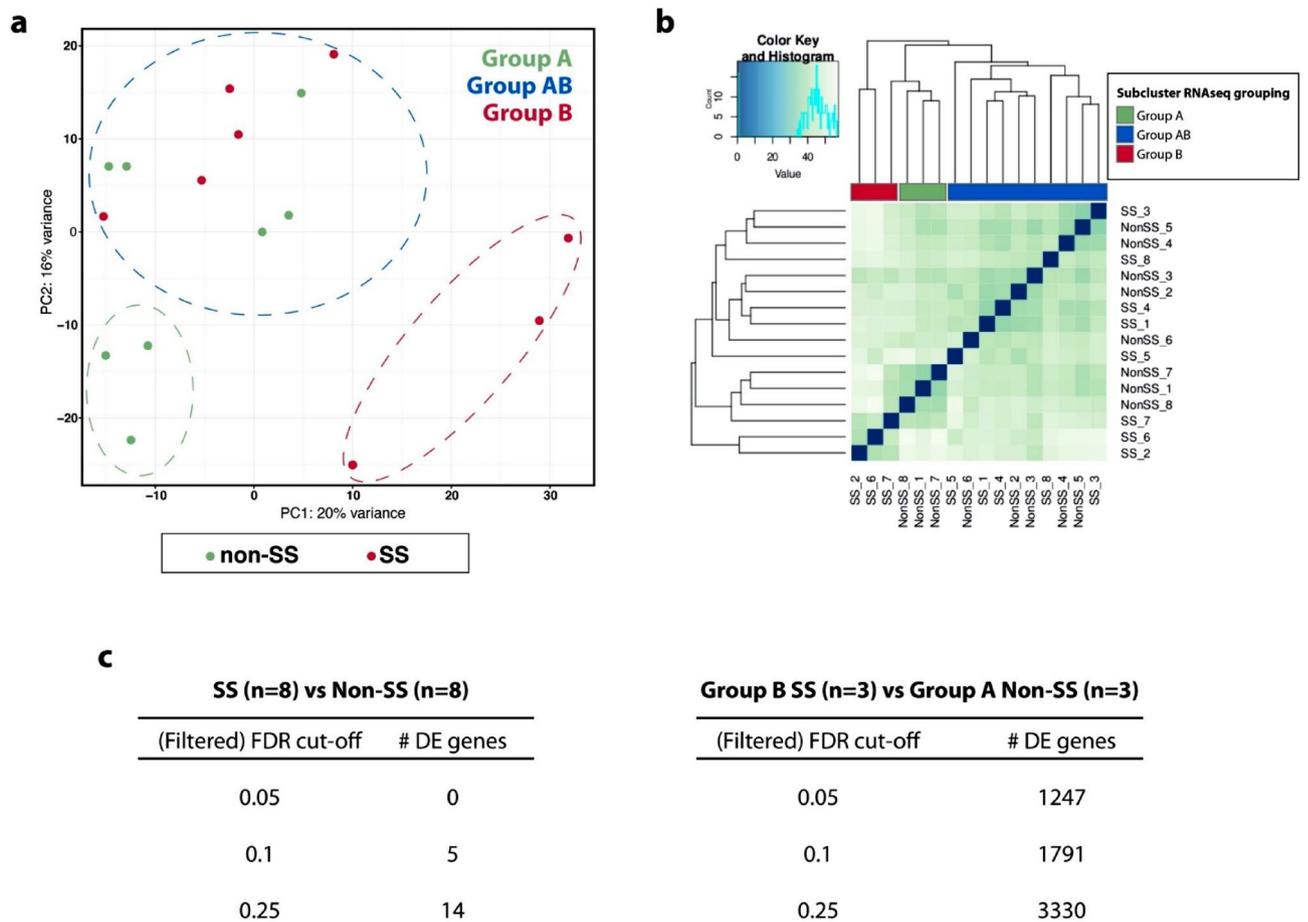
**Figure 1.** Minor salivary glands biopsies. H&E staining of minor salivary gland biopsies from: (a) Non-SS control showing normal appearing salivary gland tissue with few lymphocytic cells, (b) Non-SS control with scattered lymphocytes and some interstitial fibrosis, (c) SS patient with FS=1 showing a focal lymphocytic infiltrate and some acinar atrophy and (d) SS patient with FS=2 with a germinal center within a large lymphocytic focal infiltrate.

lated and 23 downregulated pathways for Group B vs Group A, 21 upregulated and 13 downregulated pathways for Group B vs Group AB, and 30 upregulated and 1 downregulated pathway for Group A vs Group AB (Fig. 3a).

Compared to Group AB, both Group A and Group B showed upregulated pathways in genes encoding components of the apical junction complex and apical surface of epithelial cells, epithelial–mesenchymal transition and coagulation, all indicative of changes in epithelial cell polarity. Significantly, the Group AB showed intermediary expression of genes involved in protein secretion and those activated through the mTORC1 complex as enrichment of genes in these pathways was low in Group B and high in Group A.

While the RNAseq data generated from the epithelia of non-SS and SS salivary glands showed complexity in gene signatures among patient samples, a closer comparison revealed genes that encode secreted proteins that may contribute to the disease. These included genes encoding lactotransferrin (LTF), matrix metalloproteinase 7 (MMP7), bone morphogenetic protein 3 (BMP3), interferon-inducible proteoglycan testican-2 (SPOCK2) as well as cytokines IL-15, IL-19 and BAFF (TNFSF13B). Both BAFF and IL-15 have been shown to be associated with B cell development, physiology and malignancy<sup>13,27</sup>. Furthermore, SPOCK2 was also identified as a downstream target of lncRNA-dependent miRNAs involved in the modulation of ECM genes in SS<sup>13</sup>.

**RNAscope analysis reveals increased LTF, MMP7, and BMP3 expression in Sjögren’s labial salivary gland epithelia.** Among the top genes whose transcriptional expression was increased in the RNAseq analysis of the SS grouping were LTF and in Group B were MMP7 and BMP3. LTF is a glycoprotein found in exocrine fluids, such as milk, saliva, tears, and nasal secretions, as well as in secondary granules of polymorphonuclear leukocytes (PMNs) and was previously shown to be upregulated on the protein level in saliva of SS patients<sup>28,29</sup>. MMP7 belongs to a family of ubiquitous proteolytic enzymes, the MMPs, and has been shown to be



**Figure 2.** RNAseq of LCM isolated epithelia from 8 non-SS and 8 SS patients. **(a)** PCA of total gene expression data reveals segregation between non-SS (Group A) and SS (Group B) samples and similarities between non-SS and SS (Group AB), identifying a new sub-group with mixed clinical features but the same transcriptional profile. Groups were defined as: Group A (non-SS, n = 3), Group AB (mixed non-SS/SS, n = 10), and Group B (SS, n = 3). **(b)** Sample-sample distance heat map further clusters the sample set. **(c)** Comparison of outputs of the number differentially expressed (DE) genes at different filtered FDR cut-off levels. When the biopsies were analyzed for DE genes by RNAseq cluster, a greater number of DE genes were observed.

elevated in corneal specimens of pSS patients<sup>30</sup>. BMP3 is a member of the TGF- $\beta$  superfamily of proteins, associated with inflammation in rheumatoid arthritis patients with concomitant inhibition of bone repair<sup>31</sup>.

In order to validate the increased LTF, MMP7, and BMP3 expression identified in our RNAseq data, we performed independent RNAscope analyses for expression of these genes. This in situ hybridization approach is based on the amplification of target-specific signals but not non-specific background hybridization. Samples were counterstained with pan-cytokeratin in order to restrict analysis of RNA to epithelial regions and analyzed using the CellProfiler image analysis software<sup>32–34</sup> (<http://www.cellprofiler.org>). To increase our sample size, we included an additional group of labial salivary gland samples obtained from the University of Oslo. Figure 4 illustrates combined RNAscope analysis of LTF, MMP7 and BMP3 expression from the Bergen samples described in Table 1 and Oslo samples described in Table 3. We note that the age of SS patients in the Oslo cohort was not significantly different than the non-SS group ( $51.0 \pm 19.1$  years vs  $44.5 \pm 13.1$  years respectively) although the new Oslo cohort included SS samples from two male patients and the SS groups were mostly positive for autoantibodies (Table 3). Quantitation of the RNA showed significant increases in the epithelial percentage of LTF, MMP3, and BMP3 expression in the SS group compared to the non-SS group (Fig. 4a,b). The observed increases in LTF, MMP3 and BMP3 were not driven by age as there was no significant correlation between patient age and expression of either gene (Fig. 4, “Methods”, “Statistical analysis”). These results correlate with our RNAseq findings of increased epithelial expression of LTF, MMP7, and BMP3 in SS.

## Discussion

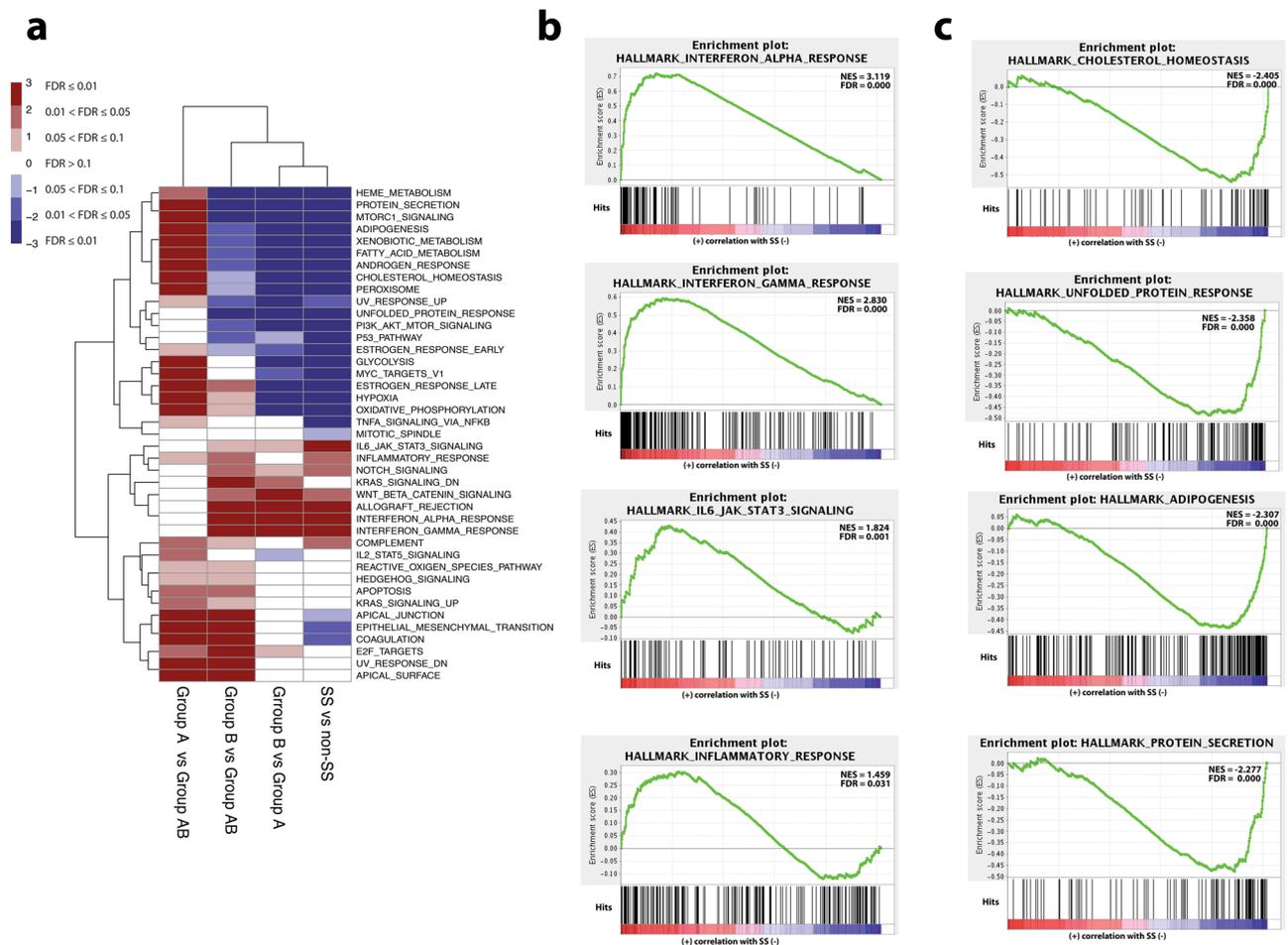
Sjögren's syndrome (SS) is a debilitating autoimmune disease with poorly understood etiology and limited biomarkers available for early detection and prognosis. In this study, we generated and annotated gene expression signatures of labial salivary gland epithelia from human SS and non-SS patients using RNAseq and computational analyses with a goal of obtaining new insights into molecular changes associated with this disease. Our studies

Gene symbol	Description	Fold change (log2)	Adjusted p value
<b>(a) List of top 10 genes upregulated in SS (n=8) than non-SS (n=8)</b>			
LTF	Lactotransferrin	0.980	0.099
REC8	REC8 meiotic recombination protein	0.947	0.099
HCP5	HLA complex P5 (non-protein coding)	0.935	0.099
SMPX	Small muscle protein, X-linked	0.908	NA
IL15	Interleukin 15	0.882	NA
CA13	Carbonic anhydrase XIII	0.816	0.099
SNORD116-6	Small nucleolar RNA, C/D box 116-6	0.806	0.272
APOL1	Apolipoprotein L, 1	0.802	0.197
HLA-F	Major histocompatibility complex, class I, F	0.793	0.256
CD8A	CD8a molecule	0.792	NA
<b>(b) List of top 10 genes downregulated in SS (n=8) than non-SS (n=8)</b>			
CA8	Carbonic anhydrase VIII	-0.762	0.275
CABS1	Calcium-binding protein, spermatid-specific 1	-0.757	0.275
FDPSP5	Farnesyl diphosphate synthase pseudogene 5	-0.730	NA
SCD	Stearoyl-CoA desaturase (delta-9-desaturase)	-0.729	0.275
SCDP1	Stearoyl-CoA desaturase (delta-9-desaturase) pseudogene 1	-0.725	NA
LINC01207	Long intergenic non-protein coding RNA 1207	-0.718	0.326
RSL24D1P1	Ribosomal L24 domain containing 1 pseudogene 1	-0.716	NA
VSIG10L	V-set and immunoglobulin domain containing 10 like	-0.669	0.348
SEMA3D	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3D	-0.668	NA
UNC5A	Unc-5 homolog A	-0.659	0.408
<b>(c) List of top 10 genes after RNAseq subclustering that are upregulated in Group B (n=3) than Group A (n=3)</b>			
MMP7	Matrix metalloproteinase 7	2.522	$7.36 \times 10^{-8}$
LINC00639	Long intergenic non-protein coding RNA 639	2.297	$2.96 \times 10^{-6}$
KCNB1	Potassium channel, voltage gated Shab related subfamily B, member 1	2.274	$3.30 \times 10^{-7}$
LINC00284	Long intergenic non-protein coding RNA 284	2.273	$3.61 \times 10^{-6}$
CLU	Clusterin	2.223	$2.73 \times 10^{-9}$
RARRES1	Retinoic acid receptor responder (tazarotene induced) 1	2.125	$2.70 \times 10^{-6}$
C19orf68	Chromosome 19 open reading frame 68	2.119	$4.17 \times 10^{-6}$
SH3D21	SH3 domain containing 21	2.062	$5.81 \times 10^{-6}$
BMP3	Bone morphogenetic protein 3	2.057	$6.22 \times 10^{-6}$
SLC38A3	Solute carrier family 38, member 3	2.022	NA
<b>(d) List of top 10 genes after RNAseq subclustering that are downregulated in Group B (n=3) than Group A (n=3)</b>			
CALML5	Calmodulin-like 5	-2.407	$4.46 \times 10^{-17}$
VSIG10L	V-set and immunoglobulin domain containing 10 like	-2.375	$9.37 \times 10^{-15}$
HPSE	Heparanase	-2.152	$1.52 \times 10^{-6}$
CABS1	Calcium-binding protein, spermatid-specific 1	-2.081	$7.43 \times 10^{-6}$
MUC7	Mucin 7, secreted	-2.008	$7.88 \times 10^{-8}$
FGF12	Fibroblast growth factor 12	-1.946	$1.28 \times 10^{-7}$
TESC	Tescalcin	-1.911	$1.41 \times 10^{-6}$
DNASE2B	Deoxyribonuclease II beta	-1.898	0.000155
SCD	Stearoyl-CoA desaturase (delta-9-desaturase)	-1.863	$3.99 \times 10^{-6}$
FADS2	Fatty acid desaturase 2	-1.853	$1.54 \times 10^{-5}$

**Table 2.** Lists of differentially expressed genes based on RNAseq.

provide evidence that damage to salivary gland epithelia is associated with changes in gene expression signatures that provide insights into deregulated pathways and cellular processes which may underlie predisposition to and early onset of SS. We postulate that the molecular sub-groups identified in our studies reflect varying stages of disease predisposition and that they may offer novel insights into the signals contributing to the development of SS.

PCA analysis showed a clear segregation of gene signatures between some non-SS and SS patients (Fig. 2). In addition, it revealed a novel subgroup of gene signatures derived from both non-SS and SS tissues that may reflect a mixed patient population. Interestingly, however, analyses of clinicopathological features of salivary epithelia from SS and non-SS patient cohorts did not reveal any consistent association with GCs and focus scores; indeed, there were no substantial differences among these patient groups in autoantibody positivity, saliva secretion, tear secretion, and subjective feeling of dry eyes or dry mouth. Our identification of a mixed gene expression

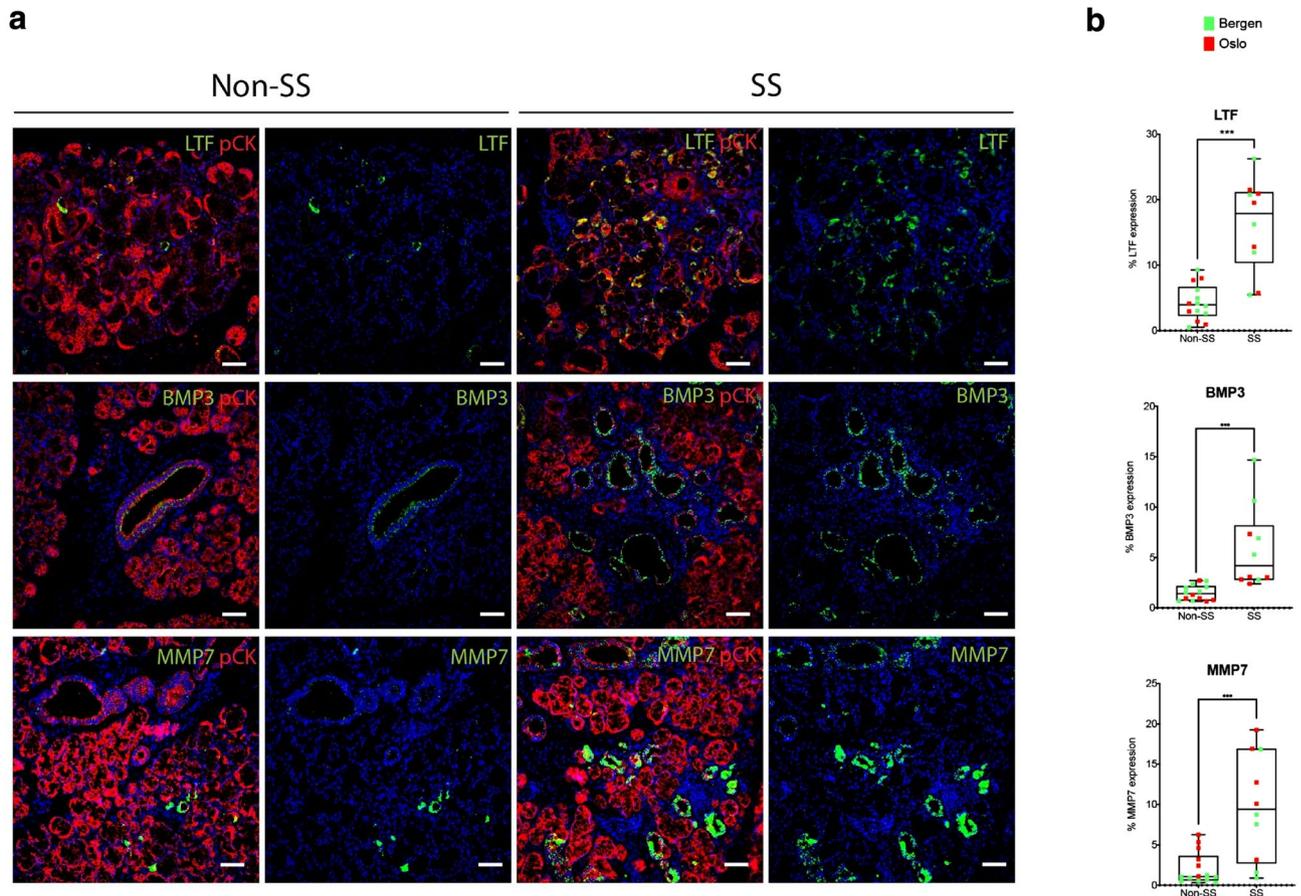


**Figure 3.** Hallmark gene set analyses from the entire RNAseq dataset. **(a)** Heatmap of hallmark gene sets that are significantly up (red) or downregulated (blue) in the respective indicated pathology. **(b)** Gene set enrichment analysis (GSEA) of hallmark gene sets that are significantly enriched in SS ( $n = 8$ ) vs non-SS ( $n = 8$ ). **(c)** GSEA of hallmark gene sets that are significantly enriched in non-SS ( $n = 8$ ) vs SS ( $n = 8$ ). NES nominal enrichment score, FDR false discovery rate. A gene set shows significant enrichment at a FDR < 0.25.

signature subgroup (Group AB) sheds novel insights into molecular changes in labial salivary glands that may reflect early disease in individuals whose symptoms do not meet the criteria for SS diagnosis.

Better understanding of the molecular features that underlie different subsets of patients is essential for disease diagnostics, prognosis, and effective treatment interventions. The focal infiltration of the salivary gland provides the basis for the focus score used to describe the pathological changes of the salivary glands in the currently used classification criteria for pSS. While the European-American Consensus Group criteria<sup>26</sup> include symptoms of dry mouth and dry eyes, the ACR-criteria include only objective criteria of reduced secretion of saliva and tears, ocular staining, positive focus score, and presence of anti-SSA/Ro in serum<sup>35</sup>. Both classification criteria for pSS require that the patients have anti-SSA/Ro antibodies and/or have a positive focus score, but do not include germinal centers. The epithelial gene expression profiles we describe show some association with clinical/pathological features of SS, but such molecular similarities are also evident in non-SS patients. This may be a consequence of the stochastic nature of clinical/pathological characteristics currently used to stratify patients into non-SS and SS groups.

Changes in SS epithelial and stromal tissue biology have been increasingly shown to play a role in SS pathobiology<sup>13,36</sup>. Previous studies revealed alterations in subpopulations of epithelial cells in NOD/ShiLJ mouse model of SS, revealing a decrease in the tissue area of secretory cells in favor of ductal epithelial cells and expansion of keratin 5-positive basal cells<sup>37</sup>. Additional studies highlighted the key roles of the innate and adaptive immune systems, as well as stromal cells, in promoting the pro-inflammatory milieu<sup>38,39</sup>. A more recent report using an LCM approach coupled with RNAseq analyses described identification of cell-type and disease-specific markers in SS patients<sup>40</sup>. However, the latter study did not strictly focus on salivary gland epithelia, and thus precluded identification of gene signatures associated only with the epithelial compartment. Additionally, recent analyses of PBMCs from SS patients provided evidence for the involvement of the perturbed long non-coding RNAs (lncRNAs)—microRNAs axes in the expression of genes involved in the pathogenesis of SS, including those with roles in B cell function and extracellular matrix remodeling, among others<sup>13</sup>. Here, our results identifying IL-15 and BAFF cytokines with important functions in B cell physiology, as well as SPOCK2 proteoglycan



**Figure 4.** RNAscope validation of gene expression changes identified from RNAseq of SS epithelial tissues. **(a)** SS vs non-SS biopsies were probed for lactotransferrin (LTF), bone morphogenetic protein 3 (BMP3), and matrix metalloproteinase 7 (MMP7) RNA (green) by in situ RNAscope and counterstained with pan-cytokeratin to mark epithelium (pCK, red). Scale bars = 20  $\mu$ m. **(b)** LTF, BMP3, and MMP7 RNAs were quantified in the respective pCK-positive tissues from patients in the Bergen (green) and Oslo (red) cohorts, revealing increased epithelial expression of LTF, BMP3, and MMP7 in SS tissues. The average % LTF, BMP3, and MMP7 area per patient sample in epithelial regions is shown. Data points represent each patient from Bergen and Oslo cohorts per condition of non-SS or SS. Each data point denotes the average of a range of 2000–4000 epithelial cells scored across multiple tissue fields from each patient sample (\*\* $p < 0.001$  compared to non-SS group determined with unpaired  $t$  tests).

Clinical classification	Patient	Age	Gender	Anti-SSA/Ro	Anti-SSB/La	ANA	Schirmer test	Saliva secretion	Dry mouth	Dry eyes	FS	Germinal centers (CD21 <sup>+</sup> )
Non-SS	Non-SSa	42	F	–	–	–	+	–	+	+	–	–
	Non-SSb	47	F	–	–	–	+	–	+	+	–	–
	Non-SSc	64	F	–	–	–	–	+	+	+	–	–
	Non-SSd	25	F	–	–	–	+	–	+	+	–	–
	Non-SSe	51	F	–	–	–	+	+	+	+	–	–
	Non-SSf	38	F	–	–	–	+	–	+	+	–	–
SS	SSa	49	F	+	+	+	+	+	+	+	1	–
	SSb	30	M	+	+	+	+	+	+	+	1	–
	SSc	38	F	–	–	–	+	+	+	+	2	+
	SSd	79	M	+	+	+	+	+	–	+	1	–
	SSe	59	F	+	–	+	+	+	+	+	1	–

**Table 3.** Additional patient samples used for RNAscope (Oslo cohort). ANA antinuclear antibodies; Saliva secretion (unstimulated flow); Schirmer test positive when reduced tear secretion (objective test); Dry mouth and dry eyes, subjective symptoms from the patient, FS Focus Score.

contributing to the increasingly recognized roles of the stromal compartment in the pathogenesis of SS<sup>36</sup>, provide important new information into their epithelial origin in SS.

In addition to signals known to be enriched in SS, including IFN- $\gamma$  and JAK/STAT-regulated genes, and genes encoding secreted factors implicated in immune responses, our studies identified LTF, also known as lactoferrin (LF), to be upregulated in expression in SS epithelia compared to non-SS tissues. Increased levels of salivary LTF with some relationship to the labial salivary gland histopathology grade were reported previously<sup>28</sup>, while no significant changes in LTF expression were detected in tear secretions in SS<sup>41</sup>. LTF is an iron-binding multifunctional glycoprotein known to function in the prevention of infections, late onset sepsis, and necrotizing enterocolitis<sup>29</sup>. LTF has been shown to stimulate a balanced T-helper-1/T-helper-2 cytokine immune response and to be a part of the innate defense, mainly at mucosa<sup>42</sup>. In our study, the observed increase in LTF gene expression in SS ductal epithelia suggests that it may be a relevant biomarker predictive of early disease with a potential causal effect on the induction of immune response.

Our data also show that MMP7 and BMP3 transcripts are upregulated in labial salivary glands from SS patients. MMP7 is a member of ubiquitous MMPs with important functions in tissue remodeling during development, but whose inappropriate expression has been linked to tissue destruction in autoimmune rheumatoid arthritis<sup>43</sup> and with corneal melting in patients with primary SS<sup>30</sup>. This suggests that upregulation of MMP7 in labial tissues from SS patients is likely to contribute to salivary gland destruction. In contrast, relatively less is known about the role of BMP3 in pSS. Previous studies documented increased expression of BMP6 in salivary and lacrimal glands from human SS patients independent of autoantibodies<sup>19</sup>. To date, increased expression of BMP3 has been linked to immune infiltration in rheumatoid arthritis and inhibition of bone repair<sup>31</sup>. Thus, our RNAseq-based identification of BMP3 in salivary epithelia from SS patients may provide additional mechanistic insights into how inflammation induces BMP3 expression to drive epithelial dysfunction in pSS.

Accumulating evidence suggests a role for the epithelium in the etiology of SS. Impaired epithelial secretion has been reported to develop independently of lymphocytic infiltration and frequently, defective secretion precedes the immune response during SS onset<sup>2</sup>. Epithelial defects are often accompanied by pro-fibrotic changes in the stroma indicating remodeling of the extracellular matrix and fibroblast activation<sup>36</sup>. Here, our findings align alterations in gene expression signatures in the salivary epithelia with different degrees of compromised salivary and lacrimal glands' function. Collectively, these studies offer a molecular explanation for exocrine tissue dysfunction as it progresses to SS and suggest that transcriptional signatures may serve as predictive biomarkers for early and advanced disease.

## Methods

**Patient selection and sample collection.** The study protocol was approved by the Norwegian Regional Committee for Medical and Health Research Ethics of West (REK 2009/686) and South-East Norway (REK 2010/1292-1). The study was performed in compliance with the tenets of the Declaration of Helsinki. Prior to participation in the study, written informed consent was obtained from all participants. The data was de-identified prior to analysis.

Labial salivary glands were obtained from a total of 27 patients with sicca symptoms. Based on the American-European Consensus Criteria<sup>26</sup>, these patients were sub-divided into 13 pSS (focus score (FS) 1–4) and 14 non-SS sicca controls (FS < 1). Inclusion criteria were: SS-patients fulfilling the AECG classification for pSS; non-SS controls were patients evaluated for SS, but not fulfilling the AECG classification. Importantly, none of the non-SS controls had a positive focus score or positive anti-SSA/Ro or anti-SSB/La autoantibodies. All participants provided informed consent prior to the initiation of this study. Clinical characteristics of patients whose minor salivary gland (SG) biopsies were analyzed for RNAseq are summarized in Table 1 (Bergen cohort) and for RNAscope analyses in Table 3 (Oslo cohort). These included assessment of objective tests of antinuclear antibodies (ANA, anti-SSA/Ro, anti-SSB/La), saliva secretion (unstimulated flow of whole saliva), tear secretion by Schirmer test, FS, and germinal center (GC)-like structures, as well as subjective symptoms noted by the patients (dry mouth and dry eyes).

**Histopathology.** Paraffin-embedded, formalin-fixed human salivary gland tissue sections, 5- $\mu$ m-thick, were stained with hematoxylin and eosin (H&E) and focal lymphocytic infiltration areas in H&E slides were captured and assessed using Aperio digital technology (Leica Biosystem). Minor SG FS of SS patients was assessed based on the currently accepted criteria<sup>4</sup>. The sections were evaluated by one oral pathologist (KS) in order to determine their FS. The tissue sections were further assessed for the presence of GC-like structures both on H&E sections and with an additional staining with CD21 (a marker of follicular dendritic cells) to improve the reliability of GC identification<sup>44</sup>. Fresh frozen biopsies from the Bergen cohort characterized in Table 1, were used for RNAseq analyses. Additional minor SG biopsies from another patient cohort (Oslo cohort, Table 3) were used for immunofluorescence labeling and RNAscope analyses.

**Laser capture microdissection (LCM) analysis.** Fresh surgical specimens were preserved in the Tissue-TEK OCT compound and frozen in isopentane, followed by storage at  $-80^{\circ}\text{C}$ . Tissue sections (5- $\mu$ m-thick) were stained with an ethanol-based cresyl violet and eosin stain, followed by examination under the microscope and annotation of epithelium and inflammatory infiltrates on digital images. Acinar and ductal epithelium was subsequently captured by LCM using a Leica LMD6 microdissection system with regions containing inflammatory cells being excluded.

**RNAseq analyses.** *RNA extraction and RNA sequencing.* RNA was isolated from laser capture microdissected labial SG epithelia. Approximately 30–100 ng of RNA was isolated from each patient sample in 4  $\mu$ l with

concentration of 7.5–25 ng/μl. Libraries for RNAseq were prepared using the Clontech Mammalian SMARTer Stranded Total RNAseq Kit Pico Input and then sequenced on an Illumina HiSeq 2500 instrument.

**Processing RNA-sequencing data.** Paired-end RNA-sequencing reads for each sample were first trimmed for Illumina TruSeq adapters (AGATCGGAAGAGCACACGTCTGAACT CCAGTCAC and AGATCGGAAGAG CGTCGTGTAGGGAAAGAGTGTAGATCGGAAGA GCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTG GTCGCCGTATCATT) using Cutadapt tool<sup>45</sup>, setting the minimum read quality threshold to 20 for the start and end of the reads, and 15 for the minimum read length. Reads were then aligned to a reference hg19 genome using Tophat<sup>46</sup>, with the number of allowed read mismatches, gap length and edit distance set to 1. Reads were then mapped to gene features using Ensembl hg19 (human genome release 19) genome annotations, and quantified as raw counts per feature using featureCount<sup>47</sup> using default parameters.

**RNA differential expression testing.** Gene feature raw RNA expression counts were normalized and tested for differential expression between different sample groupings using DESeq2<sup>48</sup>. In our analyses, we did not include age as a variable to be adjusted for when performing differential expression testing, although we are aware that age can be a confounding factor. Pairwise comparisons for the non-SS and SS RNAseq data provided sufficient coverage to estimate the differential expression of moderately to highly expressed genes.

**Gene set enrichment analysis.** For testing enrichment of pathways, genes were first ranked (genome-wide) based on their signed average log-fold change between group comparisons (e.g. Sjogren's vs control), using the pre-ranked GSEA tool v2.1.1 given a specific differential expression ranked list (based on the 't-statistic' of differential expression). Data were then tested for enrichment of annotated hallmark gene sets from the molecular signature database (MSigDB)<sup>49</sup> using the pre-ranked Gene Set Enrichment Analysis (GSEA) tool<sup>50</sup>. Pathway-based enrichment testing allowed the identification of coordinated sets of genes, and GSEA computationally determined if a defined set of genes was of statistical significance. We note that we did not use any fold-change cut-off here in addition to an FDR threshold.

**RNAscope analyses.** Formalin-fixed, paraffin-embedded non-SS and SS patient minor SG biopsies from the Bergen and Oslo patient cohorts were cut into 4-μm-thick sections and mounted onto Superfrost Plus slides. Samples were deparaffinized and in situ hybridization was conducted using the RNAscope 2.5 HD Assay Reagent Kit (ACDBio), probing for lactotransferrin Hs-LTF (LTF) RNA, matrix metalloproteinase 7 Hs-MMP7 (MMP7) RNA, and bone morphogenetic protein 3 (BMP3) RNA, Positive Control Probe Hs-PPIB, and Negative Control probe Dap-B according to the manufacturer's instructions (ACDBio). After the final amplification, fast-red chromogenic detection was performed. The epithelial localizations of LTF, MMP7, and BMP3 were determined by counterstaining the tissue sections with pan-cytokeratin protein. Sections were blocked with 5% donkey serum-TBST (DS-TBST) and incubated in primary antibody 1:100 in 5% DS-TBST for 60 min, followed by incubation in 1:500 secondary goat anti-mouse conjugated with Alexa 647, counterstained with DAPI, and mounted in ProLong Gold Antifade and coverslipped. Image processing of the LTF/BMP3/MMP7 and pan-cytokeratin (pCK) signals was carried out on the free, open-source software CellProlifer v4.0.6 across all 2048X2048px images<sup>32–34</sup> (<http://www.cellprofiler.org>). We utilized the ability to mask nuclei and epithelial cells, followed by merging the two masks to define a full epithelial cell. Nuclei outside the epithelial mask was discarded from further analysis. Following the estimation of epithelial cells, we measured the area of pixels of RNAscope probes within the of area covered by the epithelial mask to derive the percentage of area covered by a probe signal over all the epithelial areas (Supplementary Figure 1). This analysis was carried out on all patients on a minimum of 3 images per patient.

**Statistical analysis.** All analyses were performed with GraphPad Prism statistical software version 8.4.3 (GraphPad Software Inc., La Jolla, CA, USA). Unpaired *t* tests were performed to compare values among groups for RNAscope of LTF, BMP3, and MMP7 expression. Each data point represents a patient. In order to determine the impact of age on gene expression, simple linear regression was calculated for each condition (Non-SS and SS), where age was the independent variable and expression of genes (LTF, BMP3, MMP7) quantified in the RNAscope experiment was the dependent variable.  $p < 0.05$  considered to be statistically significant.

Received: 20 June 2020; Accepted: 11 May 2021

Published online: 27 May 2021

## References

1. Delaleu, N., Jonsson, M. V., Appel, S. & Jonsson, R. New concepts in the pathogenesis of Sjogren's syndrome. *Rheum. Dis. Clin. N. Am.* **34**, 833–845, vii (2008).
2. Jonsson, M. V., Delaleu, N., Brokstad, K. A., Berggreen, E. & Skarstein, K. Impaired salivary gland function in NOD mice: Association with changes in cytokine profile but not with histopathologic changes in the salivary gland. *Arthritis Rheumatol.* **54**, 2300–2305 (2006).
3. Margaix-Munoz, M., Bagan, J. V., Poveda, R., Jimenez, Y. & Sarrion, G. Sjogren's syndrome of the oral cavity. Review and update. *Med. Oral Patol. Oral Cir. Bucal* **14**, E325–330 (2009).
4. Vivino, F. B. *et al.* Sjogren's syndrome: An update on disease pathogenesis, clinical manifestations and treatment. *Clin. Immunol.* **203**, 81–121 (2019).
5. Rischmueller, M., Tieu, J. & Lester, S. Primary Sjogren's syndrome. *Best Pract. Res. Clin. Rheumatol.* **30**, 189–220 (2016).

6. Nocturne, G. & Mariette, X. Advances in understanding the pathogenesis of primary Sjogren's syndrome. *Nat. Rev. Rheumatol.* **9**, 544–556 (2013).
7. Tzioufas, A. G., Kapsogeorgou, E. K. & Moutsopoulos, H. M. Pathogenesis of Sjogren's syndrome: What we know and what we should learn. *J. Autoimmun.* **39**, 4–8 (2012).
8. Kiripolsky, J., McCabe, L. G. & Kramer, J. M. Innate immunity in Sjogren's syndrome. *Clin. Immunol.* **182**, 4–13 (2017).
9. Lessard, C. J. *et al.* Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjogren's syndrome. *Nat. Genet.* **45**, 1284–1292 (2013).
10. Haskett, S. *et al.* Identification of novel CD4<sup>+</sup> T cell subsets in the target tissue of Sjogren's syndrome and their differential regulation by the lymphotoxin/LIGHT signaling axis. *J. Immunol.* **197**, 3806–3819 (2016).
11. Michael, D. G., Pranzatelli, T. J. F., Warner, B. M., Yin, H. & Chiorini, J. A. Integrated epigenetic mapping of human and mouse salivary gene regulation. *J. Dent. Res.* **98**, 209–217 (2019).
12. Reale, M. *et al.* MicroRNA in Sjogren's syndrome: Their potential roles in pathogenesis and diagnosis. *J. Immunol. Res.* **2018**, 7510174 (2018).
13. Dolcino, M. *et al.* Long non-coding RNAs modulate Sjogren's syndrome associated gene expression and are involved in the pathogenesis of the disease. *J. Clin. Med.* **8**, 1349 (2019).
14. Hjelmerik, T. O., Petersen, K., Jonassen, I., Jonsson, R. & Bolstad, A. I. Gene expression profiling of minor salivary glands clearly distinguishes primary Sjogren's syndrome patients from healthy control subjects. *Arthritis Rheumatol.* **52**, 1534–1544 (2005).
15. Ramos, H. L., Valencia-Pacheco, G. & Alcocer-Varela, J. Constitutive STAT3 activation in peripheral CD3(+) cells from patients with primary Sjogren's syndrome. *Scand. J. Rheumatol.* **37**, 35–39 (2008).
16. Gottenberg, J. E. *et al.* Activation of IFN pathways and plasmacytoid dendritic cell recruitment in target organs of primary Sjogren's syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 2770–2775 (2006).
17. Li, H., Ice, J. A., Lessard, C. J. & Sivils, K. L. Interferons in Sjogren's syndrome: Genes, mechanisms, and effects. *Front. Immunol.* **4**, 290 (2013).
18. Jonsson, R., Theander, E., Sjostrom, B., Brokstad, K. & Henriksson, G. Autoantibodies present before symptom onset in primary Sjogren syndrome. *JAMA* **310**, 1854–1855 (2013).
19. Yin, H. *et al.* Association of bone morphogenetic protein 6 with exocrine gland dysfunction in patients with Sjogren's syndrome and in mice. *Arthritis Rheumatol.* **65**, 3228–3238 (2013).
20. Enger, T. B. *et al.* The Hippo signaling pathway is required for salivary gland development and its dysregulation is associated with Sjogren's syndrome. *Lab. Investig.* **93**, 1203–1218 (2013).
21. Barrera, M. J. *et al.* Sjogren's syndrome and the epithelial target: A comprehensive review. *J. Autoimmun.* **42**, 7–18 (2013).
22. Barrera, M. J. *et al.* Aberrant localization of fusion receptors involved in regulated exocytosis in salivary glands of Sjogren's syndrome patients is linked to ectopic mucin secretion. *J. Autoimmun.* **39**, 83–92 (2012).
23. Perez, P. *et al.* Aberrant localization of ezrin correlates with salivary acini disorganization in Sjogren's Syndrome. *Rheumatology (Oxford)* **49**, 915–923 (2010).
24. Bahamondes, V. *et al.* Changes in Rab3D expression and distribution in the acini of Sjogren's syndrome patients are associated with loss of cell polarity and secretory dysfunction. *Arthritis Rheumatol.* **63**, 3126–3135 (2011).
25. Lai, Z. *et al.* Aquaporin gene therapy corrects Sjogren's syndrome phenotype in mice. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5694–5699 (2016).
26. Vitali, C. *et al.* Classification criteria for Sjogren's syndrome: A revised version of the European criteria proposed by the American-European Consensus Group. *Ann. Rheum. Dis.* **61**, 554–558 (2002).
27. Ding, J. *et al.* BAFF overexpression increases lymphocytic infiltration in Sjogren's target tissue, but only inefficiently promotes ectopic B-cell differentiation. *Clin. Immunol.* **169**, 69–79 (2016).
28. Jezequel, N. *et al.* Salivary lactoferrin in primary Sjogren's syndrome. *Clin. Exp. Rheumatol.* **7**, 123–125 (1989).
29. Gonzalez-Chavez, S. A., Arevalo-Gallegos, S. & Rascon-Cruz, Q. Lactoferrin: Structure, function and applications. *Int. J. Antimicrob. Agents* **33**(301), e301–e308 (2009).
30. Brejchova, K., Liskova, P., Hrdlickova, E., Filipec, M. & Jirsova, K. Matrix metalloproteinases in recurrent corneal melting associated with primary Sjogren's syndrome. *Mol. Vis.* **15**, 2364–2372 (2009).
31. Matzelle, M. M. *et al.* Inflammation in arthritis induces expression of BMP3, an inhibitor of bone formation. *Scand. J. Rheumatol.* **45**, 379–383 (2016).
32. McQuin, C. *et al.* Cell Profiler 3.0: Next-generation image processing for biology. *PLoS Biol.* **16**, e2005970 (2018).
33. Jones, T. R. *et al.* Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1826–1831 (2009).
34. Jones, T. R. *et al.* Cell Profiler Analyst: Data exploration and analysis software for complex image-based screens. *BMC Bioinform.* **9**, 482 (2008).
35. Shiboski, C. H. *et al.* 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjogren's syndrome: A consensus and data-driven methodology involving three international patient cohorts. *Ann. Rheum. Dis.* **76**, 9–16 (2017).
36. Asam, S., Neag, G., Berardicurti, O., Gardner, D. & Barone, F. The role of stroma and epithelial cells in primary Sjogren's syndrome. *Rheumatology (Oxford)* <https://doi.org/10.1093/rheumatology/kez050> (2019).
37. Gervais, E. M. *et al.* Changes in the submandibular salivary gland epithelial cell subpopulations during progression of Sjogren's syndrome-like disease in the NOD/ShiLTJ mouse model. *Anat. Rec. (Hoboken)* **298**, 1622–1634 (2015).
38. Colafrancesco, S. *et al.* CXCL13 as biomarker for histological involvement in Sjogren's syndrome. *Rheumatology (Oxford)* **59**, 165–170 (2020).
39. Barone, F. *et al.* IL-22 regulates lymphoid chemokine production and assembly of tertiary lymphoid organs. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11024–11029 (2015).
40. Tandon, M., Perez, P., Burbelo, P. D., Calkins, C. & Alevizos, I. Laser microdissection coupled with RNA-seq reveal cell-type and disease-specific markers in the salivary gland of Sjogren's syndrome patients. *Clin. Exp. Rheumatol.* **35**, 777–785 (2017).
41. Dogru, M. *et al.* Lactoferrin in Sjogren's syndrome. *Ophthalmology* **114**, 2366–2367 (2007).
42. Donovan, S. M. The role of lactoferrin in gastrointestinal and immune development and function: A preclinical perspective. *J. Pediatr.* **173**(Suppl), S16–S28 (2016).
43. Perez, P. *et al.* Increased acinar damage of salivary glands of patients with Sjogren's syndrome is paralleled by simultaneous imbalance of matrix metalloproteinase 3/tissue inhibitor of metalloproteinases 1 and matrix metalloproteinase 9/tissue inhibitor of metalloproteinases 1 ratios. *Arthritis Rheumatol.* **52**, 2751–2760 (2005).
44. Fisher, B. A. *et al.* Standardisation of labial salivary gland histopathology in clinical trials in primary Sjogren's syndrome. *Ann. Rheum. Dis.* **76**, 1161–1168 (2017).
45. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal [S.I.]* **17**, 10–12 (2011).
46. Kim, D. *et al.* TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
47. Liao, Y., Smith, G. K. & Shi, W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

48. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
49. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
50. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).

## Acknowledgements

This study was performed at Biogen Idec, and in the Departments of Biochemistry, Boston University School of Medicine and Department of Molecular and Cell Biology, Boston University School of Dental Medicine. Patient data and specimens were obtained from the University of Bergen and University of Oslo Biorepositories, Norway.

## Author contributions

A.D.C., J.L.J., K.S., T.R., M.M., X.V., and M.A.K. participated in the conception and design of the experiments. A.D.C., A.T.-L., V.K., and R.J. performed the experiments and data analysis. A.D.C., J.L.J., K.S., X.V. and M.A.K. wrote the manuscript, which was revised by all authors.

## Funding

Studies described in this manuscript were supported in part by a Biogen Research Grant (XV and MAK), a National Institute of Dental and Craniofacial Research, National Institutes of Health R21 DE024954 (MAK) and a National Heart, Lung, and Blood Institute, National Institutes of Health R01 HL124392 (XV). ADC was the recipient of the American Association for Dental Research Student Research Fellowship.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-90569-w>.

**Correspondence** and requests for materials should be addressed to X.V. or M.A.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021