Discovery and validation of novel microRNAs in Sjögren's syndrome salivary glands

Sirs,

Advances in next-generation sequencing (NGS) technology enable a thorough characterisation of DNA and RNA sequences found in saliva, allowing for the discovery of previously unknown and uncharacterised sequences. Small RNAs, such as microRNAs (miRNAs) are of particular interest for researchers for their potential role in progression and pathogenesis of many diseases. In Sjögren's Syndrome (SS), an autoimmune excrinopathy that predominantly affects the lachrymal and salivary glands, changes in miRNAs are not only useful as a potential diagnostic tool, but may also provide insight into the mechanisms underlying salivary dysfunction. We have already successfully applied NGS to both quantitate known and discover novel miRNAs in minor salivary gland biopsies of SS patients. In this report, we present five more validated, novel miRNAs.

Sjögren's Syndrome (SS) is an autoimmune exocrinopathy characterised by lymphocytic infiltration of the exocrine glands including the salivary glands (SG) (1). The infiltrating lymphocytes, predominantly $CD4^+$ T-cells (2), form foci within the major and minor SGs and are used as a diagnostic criterion for SS (3, 4). The persistent and temporally increasing presence of these activated lymphocytes leads to chronic inflammation of the SG and impaired saliva secretion. Patients can exhibit variable symptoms and SS can also present secondary to other autoimmune disorders (5). Therefore, there is a need for non-invasive diagnostic biomarkers and saliva is a potential source of biological material. Moreover, microRNAs are increasingly being investigated both as biomarkers and as potential mediators of pathogenesis of SS (6, 7).

MicroRNAs are short RNA sequences (20 – 24 nt) that are involved in the regulation of mRNA translation (8). Thousands of miRNA genes have been identified in mammalian genomes and many are highly conserved across species. They are transcribed as primary transcripts, several hundred nucleotides long, sometimes containing several distinct miRNAs, and are processed into precursor transcripts of about 80 nt before transport to the cytoplasm, where they are finally processed into mature miRNAs by Dicer. Mature miRNAs are then loaded into the RNA-induced silencing complex (RISC), which can sequester or degrade mRNAs based on their complementarity to the miRNA. Due to their short length and an imperfect binding scheme, a single miRNAs can regulate multiple mRNAs in a competitive manner (9).

While our main focus has been on the detection and characterisation of miRNAs involved in pathogenesis of SS, we have also been searching for novel microRNAs. We have previously reported five novel miRNAs from minor SG biopsies of SS patients (10). Here, we report several more previously unidentified miRNAs found through next generation sequencing data that we have since characterised by qPCR. We used custom Taqman qPCR cards to measure the expression levels of these novel miRNAs across a panel of human tissue types. Since those microRNAs have never previously been described, they have not been profiled or examined for potential roles in Sjögren's syndrome development and pathogenesis.

Total RNA was isolated from minor salivary glands excised from the lower lip and immediately snap-frozen. The tissues were homogenised and microRNA was isolated with the Qiagen miRNeasy mini kit (Qiagen, Valencia, CA). The quality of RNA was assessed with the Agilent Small RNA Kit on the 2100 Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA) and NanoDrop 8000 (NanoDrop Technologies, Wilmington, DE). Patients with SS fulfilled European– American criteria for primary SS (4). The Institutional Review Board of the National Institute of Dental and Craniofacial Research approved the study, and all subjects signed informed consent.

The RNA library preparation and the sequencing were performed on the SOLiD 4 platform from Applied Biosystems (Foster City, CA, USA) by EdgeBio (Gaithersburg, MD, USA). Library preparation was according to manufacturer suggestions, using the small RNA library protocol provided with the SOLiD Total RNA-Seq kit (Applied Biosystems). Briefly, the RNA sample is enriched for miRNAs using the Invitrogen PureLink miRNA isolation kit (Invitrogen, Carlsbad, CA, USA). The amount of small RNA is estimated based on Bioanalyzer data, and equal amounts of small RNA are loaded to be reverse transcribed into cDNA and amplified by PCR for 18 cycles. The amplified cDNA is then loaded onto templated beads, and nucleotides are read as the transcript elongates during emulsion PCR.

The SOLiD platform performs primary reconstruction from the individual reads and compiles the information as 35-bp-long sequences in colorspace with corresponding quality scores for each read. These were successively mapped to three sequence databases, where reads matched to one database were removed before moving on to the next step.

First, a filter database was used to remove known contaminant and highly common sequences, like rRNA and tRNA. Next, the sequences that passed through the filter were aligned to all known miRNA sequences from miRBase version 16. Sequences not identified as known miRNAs were then aligned to the human genome (GrCH37/hg18). Alignment and counting of unique reads were carried out using the Small RNA Analysis Pipeline Tool (RNA2MAP) from Applied Biosystems and custom Perl scripting.

Each RNA sample was reversely transcribed to cDNA using the TaqMan® MicroRNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA). Single-stranded cDNA was synthesised from 500 ng of total RNA using RT Primer Pool Custom (TaqMan MicroRNA Assay, Applied Biosystems, Foster City, CA). Five µl of cDNA was used as a template in a 450 µl PCR reaction. PCR products were amplified using the TapMan Array MicroRNA Custom Card (PN 4371129, Applied Biosystem, Foster City, CA) and the TaqMan Universal PCR Master Mix (PN 4324018, Applied Biosystems, Foster City, CA), and detected using Applied Biosystems 7900 HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA).

A total of 17,035,312 transcripts were read by the SOLiD sequencer and were sequentially mapped to a known contaminants (18.45% of total reads), miRNA sequences from miRBase 16 (0.56% of total reads), and the human genome (30.49% of total reads). Note that less than 1% of all reads mapped to the human miRNome and about 50% were not matched to the human genome at all. The most abundant miRNA in the saliva was hsa-miR-1273d, accounting for about 18% of all detected miRNAs.

The analysis done by the miRanalyzer tool predicted several hundred sequences found in the NGS dataset to be potential miRNAs, 15 of which were tested – and 5 validated – in our previous study (10). Here we used a Taqman qPCR card containing assays for 26 new candidate miRNAs to measure their expression in 13 tissue types (liver, kidney, placenta, lung, testes, adipose tissue, ovary, bladder, colon, thymus, spleen and salivary glands) shown in Table I. Five of the 26 custom Taqman assays were able to amplify the target novel miRNA in most of the tissue types tested. Although the C_T values in many samples were close to or over 30 cycles, indicating low copy number of the mature miRNA, the assays were reproducible and consistent. Genomic information about these new miRNA sequences is given in Table II.

The rest of the candidates shown in Table I showed expression in only some samples, although many are expressed at higher levels.

This brief report is an extension of our previous study aimed at discovering novel miRNAs using NGS. We were able to identify 5 more previously unidentified sequences and validate their expression in human tissues. Further investigation and more samples are required to test whether these sequences do exhibit a more tissue-specific expression or if they are artifacts from other RNA sources. As a part of the post-transcriptional regulation machinery in the cell, miRNAs offer a compelling explanation of conflicting expression measurements. In SS, for example, one proteomics study (11) identified over 50 proteins that could be detected only in either SS patients or controls. A comparable microarray dataset of SS patients and healthy volunteers (12) [GEO Dataset GDS3940] reveals, however, that only 8 of the genes found in the proteomics study show a significant change between groups in the microarray, and two of those show a significant change in the opposite direction. For example, at the protein level, the gene *MIF* was only detected in SS samples, even though the microarray data shows a large and significant decrease in the transcript level. Although this discrepancy may at least partially be explained by the

Table I. qPCR results for each of the miRNA candidates tried. Values represent the Ct difference from the housekeeper average, where positive values denote higher expression than the housekeeper and negative values indicate lower expression.

Target	Salivary Gland			Adipose	Bladder	Brain	Colon	Kidney	Liver	Lung	Ovary	Placenta	Spleen	Testes	Thymus	Num Detect
	pSS	sicca	HV													
Candidate-20	-10.80	-9.05	-7.50	-10.63	-10.26	-10.42	-11.92	-10.99	-11.84	-10.23	-13.04	-10.57	-12.80	-11.15	-13.89	15
Candidate-33	-9.52	-9.61	-10.36	-12.37	-12.40	-9.08	-11.04	-10.16	-11.01	-10.16	-11.69	-8.91	-14.00	-10.85	-15.13	15
Candidate-38	-7.74	-7.89	-8.84	-11.88	-8.74	-8.15	-9.68	-9.13	-9.20	-8.94	-8.87	-8.53	-10.68	-10.17	-14.06	15
Candidate-40	-8.39	-9.59	-7.69	-10.60	-11.18	-13.54	-12.38	-12.51	-12.79	-10.57	-11.14	-12.20	-15.75	-13.30	-16.25	15
Candidate-44	-12.38	-9.38	-11.58	-8.18	-11.62	-11.27	-13.80	-11.65	-14.03	-11.95	-11.92	-11.03	--	-12.55	--	13
Candidate-13	--	-2.88	--	--	--	--	--	-15.35	-15.13	--	--	--	--	-15.64	--	4
Candidate-19	--	--	1.38	--	0.37	--	-6.50	--	-12.67	--	--	--	--	--	--	4
Candidate-25	--	--	--	-7.21	--	-7.82	--	--	2.09	--	--	--	--	-12.29	--	4
Candidate-34	--	--	--	--	--	--	--	-16.53	-14.29	--	--	-16.37	--	-17.12	--	4
Candidate-43	-2.96	-7.70	--	--	--	-13.49	--	--	--	--	--	--	--	--	--	3
Candidate-21	--	--	--	-11.86	--	--	--	--	--	--	--	--	--	--	--	1
Candidate-22	--	--	--	--	--	--	--	--	3.69	--	--	--	--	--	--	1
Candidate-27	--	--	--	--	--	--	--	--	--	-15.63	--	--	--	--	--	1
Candidate-37	--	--	--	--	--	--	--	--	--	--	--	--	--	--	-8.15	1
Candidate-39	--	--	--	--	-6.62	--	--	--	--	--	--	--	--	--	--	1
Housekeeper Avg Ct	24.22	24.35	23.75	23.11	22.88	22.12	22.81	21.44	21.89	21.70	21.80	21.72	21.49	20.87	21.55	

Table II. Sequence and location information for the five reproducible candidate miRNAs.

Novel ID	Location (hg19)	Strand	Mature sequence
Candidate-19	chr16:89874564-89874670	+	ATTCTGGGCTTTGAAATATAA
Candidate-20	chr1:227628927-227629035	+	CTGAGCTGGCCAAGGAGCAGGC
Candidate-33	chr11:59976516-59976656	+	GGGAGGTCTGCATCCTAGTCCCC
Candidate-38	chr10:97034956-97035084	+	AGGCAACTGGAGAGCCACCCCA
Candidate-40	chr16:27513163-27513283	+	GAGAGCCCAGGGACCCAC
Candidate-44	chr11:68909067-68909195	+	CCTGGAGGAGCCTGGGCTCTTT

technical and biological differences of the detection methods, it could also hint at more complex mechanisms.

Additionally, salivary glands being target organs for Sjögren's syndrome, have not been fully characterised in terms of the presence and expression of microRNAs and other non-coding RNAs, both in health and disease. Based on an increasing number of studies published and also based on work in our laboratory, we believe that non-coding RNAs and especially microRNAs contribute significantly to the homeostasis of the salivary glands as well as Sjögren's syndrome progression. Therefore, comprehensive and detailed characterisations of the non-coding RNAs of the salivary glands can facilitate further studies.

This research was supported by the Intramural Research Program of the NIH, NIDCR.

A. Gallo
M. Tandon
G. Illei
I. Alevizos

Sjögren's Syndrome and Salivary Gland Dysfunction Unit, Molecular Physiology and Therapeutics Branch, National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, MD, USA.

Address correspondence to:
Dr Ilias Alevizos, Sjögren's Syndrome and Salivary Gland Dysfunction Unit, Molecular Physiology and Therapeutics, National Institute of Dental and Craniofacial Research, 10 Center Drive Bld 10, Rm 1N110, Bethesda, MD, USA.
E-mail: alevizosi@mail.nih.gov

Competing interests: none declared.

References

1. TAPINOS NI, POLIHRONIS M, TZIOUFAS AG, SKOPOULI FN: Immunopathology of Sjögren's syndrome. *Ann Med Interne* (Paris) 1998; 149: 17-24.
2. CHRISTODOULOU MI, KAPSOGEORGOU EK, MOUTSOPOULOS HM: Characteristics of the minor salivary gland infiltrates in Sjögren's syndrome. *J Autoimmun* 2010; 34: 400-7.
3. SHIBOSKI SC, SHIBOSKI CH, CRISWELL L *et al.*: American College of Rheumatology classification criteria for Sjögren's syndrome: a data-driven, expert consensus approach in the Sjögren's International Collaborative Clinical Alliance cohort. *Arthritis Care Res* (Hoboken) 2012; 64: 475-87.
4. VITALI C, BOMBARDIERI S, MOUTSOPOULOS HM *et al.*: Assessment of the European classification criteria for Sjögren's syndrome in a series of clinically defined cases: results of a prospective multicentre study. The European Study Group on Diagnostic Criteria for Sjögren's Syndrome. *Ann Rheum Dis* 1996; 55: 116-21.
5. BALDINI C, TALARICO R, TZIOUFAS AG, BOMBARDIERI S: Classification criteria for Sjögren's syndrome: a critical review. *J Autoimmun* 2012; 39: 9-14.
6. ALEVIZOS I, ALEXANDER S, TURNER RJ, ILLEI GG: MicroRNA expression profiles as biomarkers of minor salivary gland inflammation and dysfunction in Sjögren's syndrome. *Arthritis Rheum* 2011; 63: 535-44.
7. KAPSOGEORGOU EK, GOURZI VC, MANOUSSAKIS MN, MOUTSOPOULOS HM, TZIOUFAS AG: Cellular microRNAs (miRNAs) and Sjögren's syndrome: candidate regulators of autoimmune response and autoantigen expression. *J Autoimmun* 2011; 37: 129-35.
8. AMBROS V, BARTEL B, BARTEL DP: A uniform system for microRNA annotation. *RNA* 2003; 9: 277-9.
9. YATES LA, NORBURY CJ, GILBERT RJ: The long and short of microRNA. *Cell* 2013; 153: 516-9.
10. TANDON M *et al.*: Deep sequencing of short RNAs reveals novel microRNAs in minor salivary glands of patients with Sjögren's syndrome. *Oral Dis* 2012; 18: 127-31.
11. HJELMERVIK TO, JONSSON R, BOLSTAD AI: The minor salivary gland proteome in Sjögren's syndrome. *Oral Dis* 2009; 15: 342-53.
12. GREENWELL-WILD T, MOUTSOPOULOS NM, GLIOZZI M *et al.*: Chitinases in the salivary glands and circulation of patients with Sjögren's syndrome: macrophage harbingers of disease severity. *Arthritis Rheum* 2011; 63: 3103-15.