Published in final edited form as:

J Autoimmun. 2020 February ; 107: 102354. doi:10.1016/j.jaut.2019.102354.

Salivary Dysbiosis and the Clinical Spectrum in Anti-Ro Positive Mothers of Children with Neonatal Lupus

R.M. Clancy^{1,*}, M.C. Marion², H.C. Ainsworth², M.J. Blaser³, M. Chang¹, T.D. Howard², P.M. Izmirly¹, C. Lacher³, M. Masson¹, K. Robins¹, J.P. Buyon¹, C.D. Langefeld²

¹NYU Langone Health, Department of Medicine, Division of Rheumatology, New York, NY;

²Department of Biostatistics and Data Science, Wake Forest School of Medicine, Winston-Salem, NC;

³Rutgers University, Center for Advanced Biotechnology and Medicine, Piscataway, NJ

Abstract

Mothers giving birth to children with manifestations of neonatal lupus (NL) represent a unique population at risk for the development of clinically evident pathologic autoimmunity since many are asymptomatic and only become aware of anti-SSA/Ro positivity (Anti-Ro+) based on heart block in their fetus. Accordingly, we hypothesized that the microbiome in saliva is associated with the development of autoreactivity and in some cases the progression in health status from benign to overt clinical disease including Sjögren's Syndrome (SS) and Systemic Lupus Erythematosus (SLE). The study comprised a clinical spectrum of anti-Ro+ mothers, all of whom gave birth to a child with NL: 9 were asymptomatic or had an undifferentiated autoimmune disease (Asym/UAS) and 16 fulfilled criteria for SS and/or SLE. Microbial diversity was reduced across all levels from kingdom to species for the anti-Ro+ mothers vs healthy controls (HC); however, there were no significant differences between Asym/UAS and SS/SLE mothers. Relative abundance of *Proteobacteria* and more specifically class *Betaproteobacteria* decreased with clinical severity (HC<Asym/UAS<SS/SLE). These ordered differences were maintained through the taxonomic hierarchy to three genera (*Lautropia*, *Comamonas*, and *Neisseria*) and species within these genera (*L. mirabilis*, *N. flavescens* and *N. oralis*). Biometric analysis comparing von Willebrand Factor domains present in human Ro60 with *L. mirabilis* proteins support the hypothesis of molecular mimicry. These data position the microbiome in the development of anti-Ro reactivity and subsequent clinical spectrum of disease.

Keywords

human; autoimmunity; autoantibodies; microbiome

*Corresponding Author: Robert M. Clancy, Phone: (212) 263-0751; Fax: (646) 501-0577, robert.clancy@nyulangone.org, Address: 550 1st Avenue, MSB 611, New York, NY 10016.

Declarations of interest: none.

1. Introduction

Mounting evidence suggests that autoantibodies can antecede overt clinical illnesses, such as Sjögren's syndrome (SS) and systemic lupus erythematosus (SLE), with anti-SSA/Ro positivity (anti-Ro+) among the earliest reactivities [1, 2] but developing predictive models is needed. The mothers in the Research Registry for Neonatal Lupus cohort represent a unique population at risk for overt clinical autoimmunity, including SLE. In 2009, with a focus on anti-Ro+ mothers, we examined maternal disease progression to SLE in 229 anti-Ro+ mothers with neonatal lupus (NL) children followed for at least six months [2]. Of 51 who were asymptomatic at the NL child's birth, 26 progressed and of 37 with minimal symptoms of a rheumatic disease, 16 progressed. Overall, the probability of an asymptomatic mother developing SLE by 10 years was 18.6%, and developing probable/definite SS was 27.9%. Only four mothers developed lupus nephritis. NL manifestations did not predict disease progression in an asymptomatic mother. Within each group, progression fulfills acquisition of one or more rheumatological manifestations. The median time to develop any symptom was 3.15 years. Importantly however, many mothers remained asymptomatic for over 25 years, leading us to hypothesize that there are additional genetic and environmental factors influencing progression to disease.

Though autoimmune diseases have been demonstrated to arise in part due to complex genetics, susceptibility polymorphisms are necessary but insufficient for full expression of disease [3]. Interactions of genes and environmental factors are often invoked to explain the development of disease. For example, a clinical feature of SLE is high susceptibility to Epstein-Barr virus (EBV) infection compared to the general population [4]. A relatively high proportion of SLE patients carry antibodies against EBV nuclear antigen 1 (EBNA-1) that may cross-react with Ro60 and drive lymphoid tissue B cells towards the formation of anti-Ro autoantibodies [5]. Responses to the microbiota also may have an impact on autoimmunity and autoimmune disease development [6]. Direct effects of commensals include the well-established molecular mimicry of specific epitopes as potential ligands to autoimmunity through the existence of class II MHC molecules with cross-reactivity to self-antigens [7, 8]. These studies have largely been carried out in mice or human subjects with SLE [7, 8], and subjects with pre-clinical disease have not been studied.

This study focused on anti-Ro+ mothers with the goal of identifying bacterial taxa as biomarkers in the oral microbiome which might associate with specific autoimmune reactivity and in some cases with pathological autoimmunity as supported by clinical progression. We placed the strongest associations in the context of host genetics, by ranking bacterial peptides and their affinity to DRB1*03:01 with positional scanning of peptides and biometrical analysis. To provide insight into progression, cross-reactivity of bacterial protein(s) with human Ro60 was explored to identify potential bacterial target autoantigens in anti-Ro+ mothers with differing clinical health status.

2. Materials and Methods

2.1 Clinical Characterization of Subjects

Subjects were healthy controls or mothers enrolled in the Research Registry for Neonatal Lupus (RRNL) [9]. All subjects signed informed consent for participation in the study approved by the NYU internal review board. All participating mothers in the RRNL included in this study had a child with congenital heart block (CHB) and/or characteristic rash and have antibodies to at least one component of the SSA/Ro-SSB/La ribonuclear complex including 52kD SSA/Ro, 60kD SSA/Ro or 48kD SSB/La as confirmed by ELISA in the laboratory of RC and JB [10]. No mothers were pregnant at the time of saliva donation to avoid any influence of pregnancy on the microbiota. Likewise no mother had been on antibiotics within 3 months of sample collection.

All mothers in the RRNL completed a questionnaire at the time of enrollment and during serial follow-up visits [2]. The questionnaire included: 1) general demographic information, 2) dates of birth of all the mother's affected and unaffected children, and 3) mother's symptoms of SLE (as per the revised ACR criteria [11] and SLICC criteria[12]) and Sjögren's syndrome (SS) (adapted from the American-European Consensus Group criteria for the diagnosis of SS [13]). To the extent possible, information obtained from the questionnaires was verified from the medical records of the mother's internist/rheumatologist and obstetrician and/or telephone or in person interviews with the patients by JB and PI.

Mothers were classified into one of several categories by two investigators (JB and PI) separately, and discordant assignments were adjudicated. A mother was considered asymptomatic if she had no clinical symptoms of a rheumatic disease. Given that the clinical diagnosis of SS in clinical practice does not always rely on invasive tests, which are rarely performed in routine care [14, 15], we applied two definitions for SS: 1) probable SS if the mother has at least two of the following: dry eyes, dry mouth or parotid enlargement but objective evidence of keratoconjunctivitis sicca, xerostomia, or lymphocytic foci on salivary gland biopsy was not available [13], or 2) definite SS based on the revised American-European Consensus Group criteria for SS [13] and the more recent ACR/EULAR criteria [16]. SLE was assigned based on fulfillment of one or more of the following sets of criteria: 1) ACR [11], 2) SLICC [12], or 3) the most recent ACR/EULAR [17]. When mothers had manifestations in which attribution was considered to be equally likely to be SLE or SS (for example lymphopenia) the clinical assignment was both SS/SLE as long as they fulfilled the criteria for both as described above [14, 18]. Mothers not classified into one of these categories were assigned as undifferentiated autoimmune syndrome (UAS) with Raynaud's phenomenon, leukopenia/lymphopenia, and/or photosensitivity being the symptoms most often separating them from classification as asymptomatic. Clinical diagnoses as described were assigned to the anti-Ro+ mothers with neonatal lupus children at the time of specimen collection which coincided with the most recent time of follow up. At the time of sampling, none of the SS/SLE patients providing saliva had renal involvement.

2.2 Collection and Preparation of Saliva Samples

Oral sampling was performed following a standardized and validated collection protocol. Microbial DNA from oral samples was isolated following the IHMS DNA extraction protocol Q [19], with extraction taking place directly or after storage at -80°C . The 16S ribosomal RNA (16S rRNA) gene libraries were prepared for each oral sample by generating a PCR derivative containing a sample-specific 12 nt barcode at the forward primer (515F) and a universal reverse (806R) primer. For each sample, there were three replicate libraries, which were generated with the same bar-coded oligonucleotide primer pair. Then, the material was pooled, purified, and stored until sequencing. Sequencing was performed on an Illumina MiSeq platform using paired end (PE) $2 \times 150\text{bp}$ reads, resulting in an average of 25,000 sequencing reads/sample.

2.3 Sequence Data Processing

Fastq files were processed using the Divisive Amplicon Denoising Algorithm 2 (DADA2) pipeline [20]. Primer sequences and amplicon reads with low quality scores were trimmed using Trim Galore and filtered in DADA2 based on the number of ambiguous bases, minimum quality score, and expected errors per read. Chimeras were identified and removed per DADA2 protocol. Taxonomic assignments were made using the SILVA (release 128) database, generating a table of amplicon sequence variants (ASVs) for further analysis. The R package phyloseq was used to import and manipulate the ASV table for analysis and graphical display of the data. In total, 3,556 taxa were identified. After removing ASVs that summed to zero within each specimen group, the final dataset included 864 taxa.

2.4 Statistical Analysis

Alpha diversity measures, including Shannon's Index (H) and the corresponding effective number of elements ($\exp(H)$), were computed for healthy controls (HC), anti-Ro+ mothers of NL children, together and separately by maternal disease subgroups (Asym/UAS and SS/SLE). Median differences in Shannon's diversity between healthy controls and Ro+ mothers were tested using a Wilcoxon rank sum test. In addition, differences in alpha diversity were tested assuming an ordering in clinical severity across healthy controls and the two subgroups of mothers ($\text{HC} < \text{Asym/UAS} < \text{SS/SLE}$) using the ordinal contrast $(-1, 0, 1)$ under a generalized linear model (GLM).

To circumvent the bias problem associated with the compositional nature of microbiome data, a centered log-ratio (clr) transformation was applied to the abundance counts using the R package ALDEx2 [20]. Principal component analysis (PCA) was applied to the clr-transformed abundances and characterized in Euclidean space. A logistic regression model was computed (outcome anti-Ro+ mothers vs. healthy controls) with PC1 and PC2 as independent variables. Using the parameter estimates from this model and a classification probability of 0.50, each individual's predicted class (i.e., anti-Ro+ mother or healthy control) and the corresponding discriminant line were computed.

To test for global and taxa-specific differences in clr-transformed relative abundances between HC and anti-Ro+ mothers and among the three disease groups (HC, Asym/UAS, SS/SLE), a series of tests were performed. First, global differences in microbial

communities between disease groups were tested using permutational multivariate analysis of variance (PERMANOVA). Second, generalized linear models (GLM) were used to test for differences in clr-transformed abundances between healthy controls and Ro+ mothers. An ordinal, three-group contrast (−1, 0, 1) within a GLM was used to test for differences across healthy controls, Asym/UAS, and SS/SLE, assuming an ordering in clinical severity (HC<Asym/UAS<SS/SLE).

To account for the number of statistical tests computed, the taxonomic stepdown (TSD) method was developed and applied to adjust for multiple comparisons. The TSD method integrates the taxonomic hierarchy (i.e., relatedness, similarity) into the multiple comparisons scheme to improve power by informatively narrowing the hypothesis space compared to taxonomy-agnostic multiple comparison methods. The taxonomic stepdown method tests groups of hypotheses defined by taxonomic level with appropriate group-wise multiple comparison adjustment, here the Benjamini-Hochberg false discovery rate (FDR) in a sequential manner, and each taxonomic level serves as a gate for the lower taxonomic levels within. If a specific phylum was statistically significant after applying an FDR adjustment for all of the phylum level tests, then all classes within that phylum were examined, applying an FDR adjustment just for the classes within that phylum. Testing continued down the subsequent taxonomic hierarchy in this manner until the species level was reached or until the FDR-adjusted test was no longer statistically significant. The software SPIEC-EASI [21] used the clr-transformed relative abundances to calculate the inverse covariance matrix and construct a graphical network of the bacteria.

2.5 Immunodominant Peptide of Candidate Taxa

We examined taxa showing differential relative abundance between anti-Ro+ mothers and controls for sequence homology of Ro60 and von Willebrand factor type A domain protein (vWFA) from the oral microbe. Ro60, which is also known as TROVE2, has two distinct conserved domains, TROVE and the vWFA. Within the latter, the core protein contains an immunodominant peptide at amino acids 371–381, residing at 15-amino acid which is now recognized as a T cell epitope. Many gram negative taxa, which are facultative anaerobes of the oral cavity, express a bacterial vWFA. Using a search tool at <https://www.ncbi.nlm.nih.gov/protein/advanced>, we queried “vwfa” and the specific taxon of interest in the two search fields of the builder feature, which yielded as output the vWFA primary sequence of each candidate taxon. Next, we leveraged the MHC class II binding predictions using the IEDB analysis resource “Consensus tool” with the goal to obtain a value of percentile-rank, where a lower number indicates higher binding affinity (<http://tools.immuneepitope.org/mhcii/>).

3. Results

3.1 Sample Demographics, Clinical Characteristics, and Serological Evaluation

The oral microbiome of 25 anti-Ro+ mothers of NL children and seven healthy controls were assayed using the 16S ribosomal RNA methods as described above (see 2.2). The mean ($\pm$ standard deviation) age at the time of saliva collection was 46 ± 12 years (Table 1). Additional demographic and clinical characteristics of the anti-Ro+ mothers at the time of

sample collection, including medication use, are described in Table 1. Evaluation of anti-Ro titers in blood showed that 100% and 96% of the patients had titers >1000 u/mL for anti-SSA Ro60 and anti-SSA Ro52 antibodies, respectively (Supplemental Figure 1). The seven healthy controls were 71% (n=5) of European ancestry, 85% (n=6) female, had an average age of 32±16, and were negative for anti-SSA Ro60 and anti-SSA Ro52. On review of the clinical status, the 25 anti-Ro+ mothers were further classified based on review of clinical symptoms as described in the methods into Asym/UAS, N=9, or SS/SLE, N=16 (see 2.1). There were no association of anti-Ro titers and the maternal disease status (Supplemental Figure 1).

3.2 Saliva Alpha Diversity is Reduced in Anti-Ro+ Mothers of NL Children

Overall, the sequencing of the 16S rDNA identified microorganisms from 2 kingdoms, 16 phyla, 25 classes, 41 orders, 70 families, 164 genera, and 166 species, documenting a rich oral community of microorganisms, as expected. We computed the Shannon Index (H), the entropy-based diversity index, and the corresponding effective number of elements (e^H) within each taxonomic level to estimate the alpha diversity of the oral microbiota (Table 2). For each taxonomic level, except species, there were significant reductions in H in the anti-Ro+ mothers, relative to healthy controls ($P < 0.05$). The effective number of elements, a more direct measure of diversity, shows a difference of more than 0.7 phyla relative to healthy controls (4.5 in healthy controls, 3.8 in Asym/UAS, and 3.6 SS/SLE). The corresponding difference at the genus level was over six genera (21.1 in healthy controls and 15.0 in both Asym/UAS and SS/SLE groups). In contrast, there was no evidence of a difference in the oral microbiota diversity between anti-Ro+ mothers who were diagnosed as Asym/UAS versus those diagnosed as SS/SLE ($P > 0.50$). Thus, the overall biodiversity at each taxonomic level appears to be reduced in the anti-Ro+ mothers, but comparable for the Asym/UAS and SS/SLE classified subjects.

3.3 Principal Component Analysis Separates Healthy Controls from Anti-Ro+ Mothers

We computed a principal component (PC) analysis on the clr-transformed relative abundances using Aitchison distance (see 2.4). The first two PCs explained 13.8% and 9.7% of the variation, respectively (Figure 1), and a simple linear discriminant line classified five of the seven healthy controls together with only one anti-Ro+ mother; conversely, 24 of the anti-Ro+ mothers classified together with only two of the seven healthy controls. This classification resulted in an observed concordance rate between predicted and observed of 0.91 (i.e., 29 correct classifications of 32), a positive predictive value of 0.92, a negative predictive value of 0.17, and sensitivity and specificity values of 0.71 and 0.96, respectively. Although the sample size is too small to develop a predictive algorithm, the microbiota composition provides important discriminating information relative to the presence or absence of anti-Ro+. These principal components do not correlate with age, gender, ethnicity, medications, or Asym/UAS vs. SS/SLE status ($P > 0.05$).

3.4 Microbiota Differences Across Healthy Controls, Anti-Ro+ Mothers and Disease Subgroups

There were global differences in the microbiota of anti-Ro+ mothers relative to healthy controls (PERMANOVA $P < 0.001$) but not between Asym/UAS and SS/SLE anti-Ro+ mothers (PERMANOVA $P = 0.47$).

Three phyla exhibited differential relative abundances between healthy controls and anti-Ro+ mothers (*Actinobacteria*, $P_{FDR} = 0.0232$; *Firmicutes*, $P_{FDR} = 0.0235$; Figure 2; Supplemental Table 1), along with *Bacteroidetes*, $P_{FDR} = 0.0680$ (Supplemental Table 1), and all three of these phyla were more abundant in the anti-Ro+ mothers than healthy controls. Within *Actinobacteria*, the class *Coriobacteriia* and subsequent lower taxonomic levels down to *Atopobium parvulum*, all exhibited significant increases in relative abundance in the anti-Ro+ mothers compared to healthy controls (Figure 2A; Supplemental Table 1). Similarly, all taxonomic levels from class *Actinobacteria* down to *Actinomyces graevenitzi* exhibited higher relative abundance, with significant differences for taxa of anti-Ro+ mothers versus healthy controls.

The phylum *Firmicutes* and two of its classes, *Bacilli* and *Negativicutes*, were all significantly more abundant in anti-Ro+ mothers compared to healthy controls (Figure 2B; Supplemental Table 1). For a comparison of taxa of anti-Ro mothers vs healthy controls, there was a pattern of significantly increased relative abundance, which was observed down to the genus *Streptococcus* within *Bacilli*, and to *Veillonella* within *Negativicutes*. Here, the strongest associations were with the family *Veillonellaceae* ($P_{FDR} = 0.0001$; anti-Ro+ mothers 0.073 ± 0.027 , healthy controls 0.037 ± 0.016).

There were also differences among the three groups (i.e., healthy controls, Asym/UAS, SS/SLE) in relation to clinical severity (Figure 2C; Supplemental Table 2) with a significant reduction in the clr-transformed relative abundance as clinical severity increased within one of the most frequent phyla, *Proteobacteria* ($P_{FDR} = 0.030$; healthy controls 0.24 ± 0.068 ; Asym/UAS 0.19 ± 0.12 ; SS/SLE 0.11 ± 0.082); the difference in the relative abundance between Asym/UAS and SS/SLE also was significant ($P = 0.042$). Within *Proteobacteria*, the common class *Betaproteobacteria* also showed reduced relative abundance with increasing clinical severity ($P_{FDR} = 0.0037$; healthy controls 0.11 ± 0.045 ; Asym/UAS 0.072 ± 0.066 ; SS/SLE 0.031 ± 0.035). These ordered differences were maintained down the taxonomic hierarchy to three genera (Figure 2C; Supplemental Table 2; *Lautropia*, $P_{FDR} = 0.0072$; *Comamonas* $P_{FDR} = 0.0025$; *Neisseria* $P_{FDR} = 0.0098$) and three species within these genera (*L. mirabilis*, $P_{FDR} = 0.012$; *N. flavescens* $P_{FDR} = 0.041$; *N. oralis* $P_{FDR} = 0.041$). In addition to the reduced relative abundance with increased clinical severity, there were differences detected between the Asym/UAS and SS/SLE groups (i.e., reduced relative abundance in SS/SLE subjects) within the order *Neisseriales* ($P_{FDR} = 0.0419$), family *Neisseriaceae* ($P_{FDR} = 0.0420$), and genus *Neisseria* ($P_{FDR} = 0.0431$).

While the relative abundances of *Veillonellaceae* and *Fusobacterium* were among the most significant differences between anti-Ro+ and healthy controls in the study, even without multiple comparison adjustments, taxa could not distinguish Asym/UAS from SS/SLE ($P > 0.05$; Supplemental Table 1).

3.5 Microbial Networks in Saliva

The oral microbiome is an ecological community, complete with interdependencies, succession and competition for resources [22]. To explore the potential of microbial ecological networks while accounting for the compositional nature of these data, we computed a network analysis using SPIEC-EASI. We observed one dominant network of interacting taxa that included members of the phyla *Actinobacteria* and *Firmicutes* (Figure 3). The network contained several genera and families identified as differentially abundant in the two-group comparison (healthy controls vs. anti-Ro+ mothers) reported above (Figure 2, Supplemental Table 1): *Actinomycetaceae*, *Coriobacteriaceae*, *Streptococcaceae*, and *Veillonellaceae*.

3.6 Association of Anti-Ro+ Mothers, *L. mirabilis*, and HLA Binding of Bacterial-derived Peptides at DRB1*03:01

The human leukocyte antigen (HLA) region on chromosome 6 has been strongly associated with both SS and SLE across multiple ethnicities, with extensive overlap in risk alleles between the two diseases [23–25] including DRB1*03:01. Because 67% of anti-Ro+ mothers have at least one copy of DRB1*03:01, we sought evidence of shared peptide structure between oral taxa and Ro60, with structural fitness to contain a DRB1*03:01 binding motif. The MIDAS domain of Ro60 (containing FLLAVDVSASMNQ) was identified as the highest predicted binding affinity to this class II molecule, suggesting DRB1*03:01 restricted T cell recognition of this peptide [7]. Having a set of candidate taxa, we sought to identify those sequences within the microbial MIDAS domain at vWFA with the use of a scale involving percentile-rank, wherein a lower number indicates higher binding affinity. Shown in Supplemental Table 3 is the primary vWFA sequence of ten taxa (selected from Figure 2) along with the corresponding percentile rank. The sequences of three species, *Lautropia mirabilis*, *Veillonella atypica*, and *Prevotella veroralis*, yielded lower values of percentile rank within the MIDAS domain at the vWFA motif (0.13, 0.25, and 0.4, respectively) than the value of human Ro60 (% rank, 1.25). For the former, the primary protein sequence of *L. mirabilis* was identical to human Ro60 at 7 amino acids, consistent with mimicry of Ro60 and vWFA; for this taxon, its abundance was significantly negatively correlated with disease severity (Figures 2C, 4). In contrast, the vWFA sequence of *Atopobium parvulum* yielded a peptide with a higher value of percentile rank, reflecting a dissimilarity to human Ro60. Here abundance was positively correlated with disease severity (Figure 4).

4. Discussion

In this cross-sectional study, we evaluated anti-Ro positivity and microbiome associations in the context of host genetics and disease status. Anti-Ro+ mothers had distinct differences in their oral microbiome compared to healthy controls. Specifically, anti-Ro+ mothers exhibited reduced saliva microbial diversity relative to healthy controls although significant differences between the Asym/UAS and SS/SLE individuals were not detected. Using the taxonomic stepdown method and an ordering of clinical severity, we identified taxa associated with clinical severity (healthy controls<Asym/UAS<SS/SLE). Given that nearly all anti-Ro+ mothers have at least one copy of DRB1*03:01, we evaluated candidate taxa

and their predicted binding affinity to this class II molecule by peptides within the MIDAS motif of vWFA domain. Cross-reactivity was identified between T cell epitopes and bacterial peptides, specifically the primary protein sequence of *L. mirabilis*.

The high diversity of the oral microbiome is a result of numerous distinct microenvironments (e.g., hard surfaces of the teeth, subgingival regions, epithelial surfaces of the mucosal membrane) subjected to frequent and dramatic disturbances (e.g., mastication, drinking, temperature, pH) [22]. Saliva helps buffer these microenvironments while providing nutrition and antimicrobial factors antagonistic to exogenous species [22] and helps to maintain homeostasis. When healthy resident microbiota outcompete exogenous pathogens, they contribute to normal tissue and immune development [22]. However, when the normal oral microbiota is displaced by a more pathological mix of organisms, multiple microbiota have been shown to be associated with increased risk of cardiovascular disease, autoimmune diseases such as rheumatoid arthritis, and pre-term deliveries of low-birthweight neonates [22]. Since overlap between the genetic risk loci for SS and SLE is very high, differing environmental triggers in the context of genetic susceptibility may drive an imbalance of the microbiome secondary to mimicry of bacterial and host proteins. Thus, mimicry positions the microbiome as a plausible component in the development of autoimmunity, supporting the hypothesis that deviations from the normal oral microbiome correlate with anti-Ro positivity as well as progression of disease to SS and SLE. While it has been shown that SS and SLE patients have reduced oral microbial diversity relative to controls [26, 27], we now report reduced diversity in subjects with pre-clinical disease. Several of the identified taxa have been previously associated with differential relative abundances in patients with autoimmune diseases compared to healthy controls. Our study is consistent with the results of Zhou with the phylum *Firmicutes*, down to the taxonomic level of the genus *Veillonella*, increased in a cohort of SS patients, albeit Zhou does not report significance at the phylum level as he initiates their findings at *Negativicutes* [27]. Thus, we extend their result to the phylum level. Given that *Veillonella* and *Streptococcus* are genera within *Firmicutes*, taxonomically separating at the class level, it is unclear what are the biologically relevant taxa. *Veillonella* has been associated with dental caries [28] and is considered potentially cariogenic [29]; however, these biological features may not be important for disease progression. To this point, although there are similarities between the oral microbiome of clinical and pre-clinical disease, in our study *Veillonella* did not distinguish Asym/UAS from SS/SLE.

The phylum *Actinobacteria* and each taxonomic level down to the two genera, *Actinomyces* and *Atopobium*, which diverge at the class level, show greater relative abundances in anti-Ro + mothers compared to healthy controls. Prior studies show associations of *Actinomyces* with several disease processes but in both directions [30–33]. Consistent with our results, the saliva of individuals diagnosed with RA had the genus *Atopobium* at increased relative abundance compared with healthy controls [34]. Comparison of our three clinical groups (healthy controls, Asym/UAS, SS/SLE) identified numerous differences in relative abundances down the taxonomic hierarchy. Three genera, all in different families, showed decreasing abundance with increasing clinical severity. In fact, the relative frequency of the genus *Neisseria* was significantly less in SS/SLE patients compared to Asym/UAS subjects,

results consistent with a prior report [27]. *Neisseria* species are common commensals across the globe [35, 36].

Our study supports a path of inflammation related to molecular mimicry [37], with a possible microbiome-anti-Ro+ connection leading to a deregulation of immune responses. Here, the immune system may consider particular microbiome constituents as pathogenic. We evaluated specific microbial taxa in the context of cross-reactivity between T cell epitopes and bacterial peptides. Prior HLA typing of anti-Ro+ mothers shows a high reactivity against SSA/Ro60 in DRB3* haplotypes, with the demonstration that the Ro60 MIDAS motif plays a critical role in the immunogenicity [7, 38]. Using a transgenic DRB3* murine model, Szymula et al found that a microbial protein such as vWFA could lead to autoantibody production secondary to T cell activation, in the context of HLA presentation of a peptide with the bacterial MIDAS motif. Here, we report a peptide of the MIDAS motif of the vWFA domain of *L. mirabilis* showing the highest predicted binding affinity to this class II molecule, a result implicating DRB1*03:01 restricted T cell recognition of this peptide. An alternative possibility is that the Ro60 orthologs in bacteria initiate the immune response against self Ro60 [8]; however, the Ro60 orthologs are intracellular, while bacterial proteins with vWFA domains are extracellular [39].

This study identified a number of differentially abundant taxa across a spectrum of clinical autoimmunity and links oral dysbiosis with subclinical disease. Nevertheless, the study has several limitations. Sequencing based on 16S does not fully delineate all organisms (e.g., viruses, eukaryotes), nor does it take into account strain differences but is less prone to host contamination. As with any cross-sectional observational study, we were not able to discern whether the observed differences in the oral microbiome contribute to the development of disease or result from disease. This can be further addressed in future studies as we prospectively evaluate our recently referred anti-Ro+ asymptomatic mothers over time. Although the study is powered to detect moderate to large effects, it cannot efficiently adjust for a large number of confounders or complete extensive multivariate modeling. We note that there are disease comorbidities outside of SS/SLE, including other autoimmune diseases (e.g., autoimmune liver patients), that were not part of this study. Similarly, there were environmental factors such as smoking that were not evaluated. Some results could be influenced by the use of specific medication; however, over half of the patients (15/25) were not receiving any medications, and the majority of the remaining participants were prescribed only hydroxychloroquine. Furthermore, no mother had been on antibiotics within 3 months of sample collection. It is also an acknowledged limitation that we did not have sufficient numbers of subjects to separate SLE from SS, which may have yielded additional insights; however, these two diseases share many clinical features in common such as photosensitivity, arthritis, lymphopenia, leukopenia, and low complement levels. In fact, data exist that patients diagnosed with SS often fulfill criteria for SLE [14]. Since the SLE mothers in this evaluation were on the milder spectrum of disease, did not have renal involvement, and were on little to no medications, identification of even greater differences in microbiome composition along disease severity may have been muted. Finally, the inferences from this study are based on the oral microbiome and the spectrum of autoimmunity studied. Studies that corroborate and expand upon these results are needed,

particularly the molecular mimicry result, to shape future directions and help elucidate the microbiome-autoimmune connection.

To improve our power to detect associations, we introduced a novel, biologically informed multiple comparisons method (taxonomic stepdown method) designed to leverage our 16S-based knowledge of taxonomic relatedness and similarity. However, microbial taxonomy is an active field of research, and some of the lower taxonomic relationships may be less precise. Taking advantage of taxonomic structure results in a much smaller burden of multiple comparisons and highlights the taxonomic hierarchy and insight inherent in many of these associations (e.g., multiple closely related genera within the same family).

4.1 Conclusions

In summary, we observed microbiota differences between healthy controls, anti-Ro+ mothers and disease subgroups. We placed a strong association of a candidate taxon, *L. mirabilis*, in the context of host genetics by ranking bacterial peptides and their affinity to DRB1*03:01. These results support the potential for cross-recognition of Ro60 and homologous peptides from the oral microbiota, suggesting a possible role of bacterial antigens as inducers or potentiators of a pathogenic autoimmune response in anti-Ro+ mothers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

The authors are grateful to Ben Wainwright for excellent editorial assistance.

Funding: This work was supported by the National Institutes of Health, Bethesda, MD (5P50 AR070591 (J.P.B.)), 5R21 AR071106, (J.P.B. and C.D.L.)). For the genomics facility, DNA sequencing relied on a shared resource, which is partially supported by the National Institutes of Health via a Cancer Center Support Grant (P30 CA016087).

References

- [1]. Arbuckle MR, McClain MT, Rubertone MV, Scofield RH, Dennis GJ, James JA et al. Development of autoantibodies before the clinical onset of systemic lupus erythematosus. *N Engl J Med*, 2003;349:1526–33. [PubMed: 14561795]
- [2]. Rivera TL, Izmirly PM, Birnbaum BK, Byrne P, Brauth JB, Katholi M et al. Disease progression in mothers of children enrolled in the Research Registry for Neonatal Lupus. *Ann Rheum Dis*, 2009;68:828–35. [PubMed: 18625627]
- [3]. Arango MT, Perricone C, Kivity S, Cipriano E, Ceccarelli F, Valesini G et al. HLA-DRB1 the notorious gene in the mosaic of autoimmunity. *Immunol Res*, 2017;65:82–98. [PubMed: 27435705]
- [4]. Kang I, Quan T, Nolasco H, Park SH, Hong MS, Crouch J et al. Defective control of latent Epstein-Barr virus infection in systemic lupus erythematosus. *J Immunol*, 2004;172:1287–94. [PubMed: 14707107]
- [5]. Poole BD, Templeton AK, Guthridge JM, Brown EJ, Harley JB, James JA. Aberrant Epstein-Barr viral infection in systemic lupus erythematosus. *Autoimmun Rev*, 2009;8:337–42. [PubMed: 19167523]

- [6]. Rosser EC, Mauri C. A clinical update on the significance of the gut microbiota in systemic autoimmunity. *J Autoimmun*, 2016;74:85–93. [PubMed: 27481556]
- [7]. Szymula A, Rosenthal J, Szczerba BM, Bagavant H, Fu SM, Deshmukh US. T cell epitope mimicry between Sjogren's syndrome Antigen A (SSA)/Ro60 and oral, gut, skin and vaginal bacteria. *Clin Immunol*, 2014;152:1–9. [PubMed: 24576620]
- [8]. Greiling TM, Dehner C, Chen X, Hughes K, Iniguez AJ, Boccitto M et al. Commensal orthologs of the human autoantigen Ro60 as triggers of autoimmunity in lupus. *Sci Transl Med*, 2018;10.
- [9]. Buyon JP, Hiebert R, Copel J, Craft J, Friedman D, Katholi M et al. Autoimmune-associated congenital heart block: demographics, mortality, morbidity and recurrence rates obtained from a national neonatal lupus registry. *J Am Coll Cardiol*, 1998;31:1658–66. [PubMed: 9626848]
- [10]. Reed JH, Clancy RM, Lee KH, Saxena A, Izmirly PM, Buyon JP. Umbilical cord blood levels of maternal antibodies reactive with p200 and full-length Ro 52 in the assessment of risk for cardiac manifestations of neonatal lupus. *Arthritis Care Res (Hoboken)*, 2012;64:1373–81. [PubMed: 22511615]
- [11]. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum*, 1982;25:1271–7. [PubMed: 7138600]
- [12]. Petri M, Orbai AM, Alarcon GS, Gordon C, Merrill JT, Fortin PR et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum*, 2012;64:2677–86. [PubMed: 22553077]
- [13]. Vitali C, Bombardieri S, Jonsson R, Moutsopoulos HM, Alexander EL, Carsons SE 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*, 2002;61:554–8. [PubMed: 12006334]
- [14]. Izmirly PM, Buyon JP, Wan I, Belmont HM, Sahl S, Salmon JE et al. The Incidence and Prevalence of Adult Primary Sjogren's Syndrome in New York County. *Arthritis Care Res (Hoboken)*, 2019;71:949–60. [PubMed: 30044541]
- [15]. Maciel G, Crowson CS, Matteson EL, Cornec D. Prevalence of Primary Sjogren's Syndrome in a US Population-Based Cohort. *Arthritis Care Res (Hoboken)*, 2017;69:1612–6. [PubMed: 27998024]
- [16]. Shiboski CH, Shiboski SC, Seror R, Criswell LA, Labetoulle M, Lietman TM 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*, 2017;76:9–16. [PubMed: 27789466]
- [17]. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R et al. 2019 European League Against Rheumatism/American College of Rheumatology Classification Criteria for Systemic Lupus Erythematosus. *Arthritis Rheumatol*, 2019.
- [18]. Tedeschi SK, Johnson SR, Boumpas DT, Daikh D, Dorner T, Diamond B et al. Multicriteria decision analysis process to develop new classification criteria for systemic lupus erythematosus. *Ann Rheum Dis*, 2019;78:634–40. [PubMed: 30692164]
- [19]. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F et al. Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol*, 2017;35:1069–76. [PubMed: 28967887]
- [20]. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*, 2016;13:581–3. [PubMed: 27214047]
- [21]. Kurtz ZD, Muller CL, Miraldi ER, Littman DR, Blaser MJ, Bonneau RA. Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput Biol*, 2015;11:e1004226. [PubMed: 25950956]
- [22]. Lamont RJ, Koo H, Hajishengallis G. The oral microbiota: dynamic communities and host interactions. *Nat Rev Microbiol*, 2018;16:745–59. [PubMed: 30301974]
- [23]. Langefeld CD, Ainsworth HC, Cunninghame Graham DS, Kelly JA, Comeau ME, Marion MC et al. Transancestral mapping and genetic load in systemic lupus erythematosus. *Nat Commun*, 2017;8:16021. [PubMed: 28714469]

- [24]. Hanscombe KB, Morris DL, Noble JA, Dilthey AT, Tombleson P, Kaufman KM et al. Genetic fine mapping of systemic lupus erythematosus MHC associations in Europeans and African Americans. *Hum Mol Genet*, 2018;27:3813–24. [PubMed: 30085094]
- [25]. Lessard CJ, Li H, Adrianto I, Ice JA, Rasmussen A, Grundahl KM et al. Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjogren's syndrome. *Nat Genet*, 2013;45:1284–92. [PubMed: 24097067]
- [26]. Correa JD, Calderaro DC, Ferreira GA, Mendonca SM, Fernandes GR, Xiao E et al. Subgingival microbiota dysbiosis in systemic lupus erythematosus: association with periodontal status. *Microbiome*, 2017;5:34. [PubMed: 28320468]
- [27]. Zhou Z, Ling G, Ding N, Xun Z, Zhu C, Hua H et al. Molecular analysis of oral microflora in patients with primary Sjogren's syndrome by using high-throughput sequencing. *PeerJ*, 2018;6:e5649. [PubMed: 30280027]
- [28]. Belda-Ferre P, Alcaraz LD, Cabrera-Rubio R, Romero H, Simon-Soro A, Pignatelli M et al. The oral metagenome in health and disease. *ISME J*, 2012;6:46–56. [PubMed: 21716308]
- [29]. Aas JA, Griffen AL, Dardis SR, Lee AM, Olsen I, Dewhirst FE et al. Bacteria of dental caries in primary and permanent teeth in children and young adults. *J Clin Microbiol*, 2008;46:1407–17. [PubMed: 18216213]
- [30]. Casarin RC, Barbagallo A, Meulman T, Santos VR, Sallum EA, Nociti FH et al. Subgingival biodiversity in subjects with uncontrolled type-2 diabetes and chronic periodontitis. *J Periodontal Res*, 2013;48:30–6. [PubMed: 22762355]
- [31]. de Groot PF, Belzer C, Aydin O, Levin E, Levels JH, Aalvink S et al. Distinct fecal and oral microbiota composition in human type 1 diabetes, an observational study. *PLoS One*, 2017;12:e0188475. [PubMed: 29211757]
- [32]. Long J, Cai Q, Steinwandel M, Hargreaves MK, Bordenstein SR, Blot WJ et al. Association of oral microbiome with type 2 diabetes risk. *J Periodontal Res*, 2017;52:636–43. [PubMed: 28177125]
- [33]. Vielkind P, Jentsch H, Eschrich K, Rodloff AC, Stingu CS. Prevalence of *Actinomyces* spp. in patients with chronic periodontitis. *Int J Med Microbiol*, 2015;305:682–8. [PubMed: 26324012]
- [34]. Zhang X, Zhang D, Jia H, Feng Q, Wang D, Liang D et al. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nature medicine*, 2015;21:895–905.
- [35]. Nasidze I, Li J, Quinque D, Tang K, Stoneking M. Global diversity in the human salivary microbiome. *Genome Res*, 2009;19:636–43. [PubMed: 19251737]
- [36]. Liu G, Tang CM, Exley RM. Non-pathogenic *Neisseria*: members of an abundant, multi-habitat, diverse genus. *Microbiology*, 2015;161:1297–312. [PubMed: 25814039]
- [37]. van der Meulen TA, Harmsen HJM, Vila AV, Kurilshikov A, Liefers SC, Zhernakova A et al. Shared gut, but distinct oral microbiota composition in primary Sjogren's syndrome and systemic lupus erythematosus. *J Autoimmun*, 2019;97:77–87. [PubMed: 30416033]
- [38]. Paisansinsup T, Deshmukh US, Chowdhary VR, Luthra HS, Fu SM, David CS. HLA class II influences the immune response and antibody diversification to Ro60/Sjogren's syndrome-A: heightened antibody responses and epitope spreading in mice expressing HLA-DR molecules. *J Immunol*, 2002;168:5876–84. [PubMed: 12023392]
- [39]. Konto-Ghiorgi Y, Mairey E, Mallet A, Dumenil G, Caliot E, Trieu-Cuot P et al. Dual role for pilus in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*. *PLoS Pathog*, 2009;5:e1000422. [PubMed: 19424490]

Highlights

- Anti-Ro+ neonatal lupus mothers are at risk for pathologic autoimmunity.
- Disease progression depends on genetics and microbiotic taxa in Anti-Ro+ mothers.
- Molecular mimicry of salivary microbiome and Ro60 associate with clinical disease.

Figure 1.

Principal component plot with discriminant line distinguishing healthy controls (n=7) from anti-Ro+ mothers of neonatal lupus children (n=25) based on 16S V4 sequencing of oral microbiome. Principal components (PCs) were computed on the clr-transformed relative abundances (see 2.4); PC1 and PC2 explained 13.8% and 9.7% of the variation in the oral microbiome. A discriminant line was computed from a logistic regression model (outcome anti-Ro+ mothers vs. healthy controls), where the estimated parameters are $\text{logit}(p) = 2.1 + 0.076 \cdot \text{PC1} + 0.056 \cdot \text{PC2}$ and the predicted probability was set to $p=0.50$ for classification. This classification resulted in a positive predictive value of 0.92, a negative predictive value of 0.17, and sensitivity and specificity values of 0.71 and 0.96, respectively.

Figure 2.

Taxa meeting statistical significance (FDR P-value < 0.05) using the taxonomic stepdown method (full results available in Supplemental Tables 1 and 2). FDR P-values are shown in the graph on the left with the corresponding taxa (indicated by color) shown in context of taxonomic hierarchy on the right. The significance threshold of $P = 0.05$ is represented as a dotted line. (A) Taxa within phylum *Actinobacteria* showing statistically significant differences in clr-transformed relative abundance between healthy controls and anti-Ro+ mothers of neonatal lupus children. (B) Taxa within phylum *Firmicutes* showing statistically significant differences in clr-transformed relative abundance between healthy controls and anti-Ro+ mothers of neonatal lupus children. (C) Taxa within phylum *Proteobacteria* showing statistically significant ordinal differences in clr-transformed relative abundance when comparing three populations in order of increasing disease severity (healthy controls < Asym/UAS < and SS/SLE).

Figure 3.

Correlation network of oral taxa at the family level showing statistically significant differences (FDR P-value < 0.05) in clr-transformed relative abundance between healthy controls and anti-Ro+ mothers using the taxonomic stepdown method. Correlations were calculated using SPIEC-EASI. The presence of an edge between two nodes suggests the two families interact. Nodes are colored by phylum. Edges are labeled with values from the inverse covariance matrix.

AA Sequence	Protein
FLLAVDVSASMNQ	Ro60
LLLLLDVSGSMAG (WP_005674879.1)	vWFA domain of <i>Lautropia mirabilis</i>
VLVVF ^Y DN ^R SS ^R MTA (WP_012809344.1)	vWFA domain of <i>Atopobium parvulum</i>

Figure 4.

Sequence homology of human Ro60 and peptides found in bacterial vWFA (above) and associations of targeted taxa and anti-Ro+ mothers of neonatal lupus children vs controls (below). For the former, note that Ro60 shares seven of its first 11 aa with *L. mirabilis* vWFA, while it shares only five with *A. parvulum* vWFA (sharing denoted by yellow; and polar and large residue, by red). Note that for different groups arranged by severity, there was the depletion of *L. mirabilis* along with an expansion of *A. parvulum* (lower panel). P-value denotes a significant difference between two groups (SS/SLE and healthy controls).

Table 1.

Demographic, clinical and serological characteristics of anti-Ro+ mothers at time of saliva sample.

	Total (N=25)	Asym/UAS (N=9)	SS/SLE (N=16)
Age, mean±SD	46.4±12.0	37.6±7.4	51.4±11.3
Race/ethnicity			
White	17	5	12
Black	0	0	0
Asian	5	2	3
Multiethnic	3	2	1
Hispanic	2	2	0
Non-Hispanic	23	7	16
Medications			
None	15	7	8
Anti-malarials Only	6	2	4
Anti-malarials + Steroids + Immunosuppressants	2	0	2
Steroids Only	1	0	1
Steroids + Immunosuppressants	1	0	1

Table 2.

Alpha diversity in healthy controls and anti-Ro+ mothers of neonatal lupus children.

T axonomic Level	Healthy Controls (N=7) H (e ^H)	Asym/UAS* (N=9) H (e ^H)	SS/SLE (N=16) H (e ^H)	P-value ¹ Healthy Controls vs. All RRNL	P-value ² Healthy Controls vs. Asym/UAS vs. SS/SLE	P-value ¹ Asym/UAS vs. SS/SLE
Phylum	1.51±0.11 (4.53)	1.34±0.22 (3.82)	1.29±0.23 (3.63)	0.024	0.068	0.562
Class	2.07±0.12 (7.92)	1.91±0.24 (6.75)	1.87±0.22 (6.49)	0.053	0.146	0.709
Order	2.30±0.08 (9.97)	2.08±0.27 (8.00)	2.06±0.20 (7.85)	0.011	0.042	0.863
Family	2.67±0.13 (14.44)	2.36±0.30 (10.59)	2.34±0.20 (10.38)	0.001	0.006	0.796
Genus	3.05±0.19 (21.11)	2.71±0.41 (15.03)	2.71±0.28 (15.03)	0.011	0.043	0.968
Species	3.14±0.28 (23.10)	2.94±0.47 (18.92)	2.87±0.30 (17.64)	0.110	0.257	0.663

Results reported: mean ± standard deviation of the Shannon entropy-based diversity index (H) and the corresponding effective number of elements (exp(H)).

* Asym/UAS represents asymptomatic and undifferentiated autoimmune syndrome (UAS major and minor) anti-Ro⁺ neonatal lupus mothers in the RRNL.

¹ P-value for two group comparison using a t-test.

² P-value based on all three groups using analysis of variance with ordinal alternative reflecting clinical severity.