Whole Exome Sequencing in Early-onset Systemic Lupus Erythematosus

Ezgi Deniz Batu, Can Koşukçu, Ekim Taşkıran, Sezgin Sahin, Sema Akman, Betül Sözeri, Erbil Ünsal, Yelda Bilginer, Ozgur Kasapcopur, Mehmet Alicaşifoğlu, and Seza Ozen

ABSTRACT. Objective. Systemic lupus erythematosus (SLE) is a multisystem autoimmune disorder. Early-onset, familial, and/or syndromic SLE may reveal monogenic pathologies. The aim of this study was to examine genetic associations in patients with early-onset or familial SLE.

Methods. We enrolled 7 SLE cases (from different families) with disease onset ≤ 5 years of age and family history consistent with an autosomal recessive inheritance. Whole exome sequencing (WES) was performed in 6 index cases. Suspected variants were confirmed by Sanger sequencing. We did not perform WES in 1 patient who had features similar to the first 3 cases; only the exons of C1QA, C1QB, and C1QC were screened with Sanger sequencing.

Results. We demonstrated 2 novel and 3 previously reported variants in genes associated with SLE:

1. A homozygous non-sense alteration (c.622C>T/p.Gln208Ter) in C1QA in 2 patients; homozygous non-sense alteration (c.79C>T/p.Gln27Ter) in C1QC in 1 (novel variant); homozygous missense alteration (c.100G>A/p.Gly34Arg) in C1QC in 1 baby; homozygous missense alteration (c.1945G>C/p.Thr97Ilefs*4) in DNASeIL3 in 1 patient. Further, in 1 patient, we determined a strong candidate variant in HDAC7 (histone decetylase 7).

Conclusion. Five patients had homozygous alterations in genes coding early complement components. This may lead to decreased clearance of apoptotic bodies. One patient had DNASeIL3 variant, which functions in the clearance of self-antigens. In 1 patient, we determined a novel gene that may be important in SLE pathogenesis. We suggest that monogenic causes/associations should be sought in early-onset and/or familial SLE. (J Rheumatol First Release July 15 2018; doi:10.3899/jrheum.171358)

Key Indexing Terms: EARLY-ONSET SYSTEMIC LUPUS ERYTHEMATOSUS MONOGENIC COMPLEMENT SYSTEM WHOLE EXOME SEQUENCING
Examples include genes important for apoptosis (e.g., *FASLG*), DNA damage repair (e.g., *TREX1*), clearance of self-antigens (e.g., *DNASE1L3*), and nucleic acid sensing and type I interferon (IFN) overproduction (e.g., *TREX1, TMEM173*)\(^6\). In addition, deficiencies in complement components may predispose to SLE through the inability to clear apoptotic cells (C1q deficiency) and negative selection of self-reactive B cells (C4 deficiency)\(^6\).

Our aim was to investigate genetic causes in cases of early-onset SLE whose family history suggested an autosomal recessive inheritance.

**MATERIALS AND METHODS**

**Patients.** We enrolled 7 pediatric SLE cases who had early disease onset (before or at 5 yrs of age) and a family history consistent with an autosomal recessive inheritance (affected siblings or parenteral consanguinity). The pedigrees of the patients are shown in Figure 1 and Figure 2. All patients fulfilled the Systemic Lupus International Collaborating Clinics (SLICC) criteria\(^7\). Demographic data, clinical manifestations, laboratory and histopathological findings, treatment, and outcome (the most recent SLICC/American College of Rheumatology Damages Index\(^8\)) were documented. The subjects’ written consent was obtained according to the Declaration of Helsinki (1964) and the study was approved by the ethics committee of Hacettepe University (GO 14/368-09).

Whole exome sequencing (WES) was performed on all patients except Patient 4. Sanger sequencing of *C1Q* genes was performed in Patient 4 because he had similar characteristics (normal C3, C4 levels, and negative anti-dsDNA) with Patients 1–3.

**DNA extraction, library preparation, and Ion Proton sequencing.** Genomic DNA was extracted by standard phenol-chloroform protocol from peripheral blood of all patients and family members after the receipt of the informed consent. Ion AmpliSeq Exome RDY Kit and Ion Proton sequencing platform allowed for the capturing, amplification, and sequencing. Emulsion PCR was performed on Ion OneTouch 2 instrument using the Ion PI Hi-Q OT2 200 Kit. Enrichment of ion sphere particles (ISP) was performed using the One Touch ES module and ISP were loaded onto Ion PI chips, then sequenced according to the manufacturer’s instructions (Thermo Fisher Scientific).

**Bioinformatic analysis.** Exome sequencing was performed on 6 pediatric patients with SLE. Base coverage depth was 62x and 94% of the mapped reads were aligned over a target region in average. Torrent Mapping Alignment Program was used for alignment to hg19/GRCh37 human reference genome. Raw sequence file was processed using the Torrent Suite software v5.0.5 for generating mapped reads, analyzing coverage data, and variant calling. Variants with a coverage <5x for single-nucleotide polymorphisms (SNP) and 10x for INDEL were excluded from further analysis. Phred Quality Scores were specified as Q15 for SNP variants and Q20 for INDEL. All variants that passed the quality score filters were annotated with the Ion Reporter 5.2 software. We applied variant filtration steps to select and filter the variants detected with exome sequencing.

For each sample, 48,492 variants were identified on average, including single-nucleotide variants and small INDEL. After discarding common polymorphisms (minor allele frequency > 0.01), a mean of 3542 variants remained. After that, exonic and nonsynonymous substitutions were selected, which we used for in-house filtering. The alterations that were in the boundaries of runs of homozygosity (ROH) regions ≥ 1.5 Mb within all chromosomes were extracted. We used AgileVCFMapper, a software that performs mapping of disease loci by SNP genotyping using variant caller format file obtained from exome sequence data, to identify the ROH in index cases\(^9\). As a final step, variants that were observed in ExAC Browser (in homozygous state) were ruled out. All Human Gene Mutation Database–linked mutations were collected with Ingenuity Variant Analysis software (IVA/QIAGEN Bioinformatics).

**RESULTS**

The clinical and laboratory characteristics of the patients are summarized in Table 1 and Table 2. The characteristics of Patient 1 and 4 have been presented recently\(^10\). We identified 2 novel alterations as a homozygous non-sense variant in *C1QC* (Q27X/c.79C>T) in 1 patient (Patient 2) and a missense variant in *C1S* (A649P/c.1945G>C) in another patient (Patient 5). These 2 alterations were not present in ExAC Browser or other public databases such as dbSNP (Build 151) for default global population and Exome Variant Server (ESP6500SI-V2) for European and African American populations neither in homozygous state nor as heterozygous carrier. In addition, these changes were not present in 245 clinically unrelated Turkish individuals who constitute the in-house database established within the Hacettepe Exome Project. Open access variant functional prediction software tools were used to estimate the severity of these sequence variants in the genome. A649P alteration on *C1S* was predicted as “polymorphism” with MutationTaster and “damaging/probably damaging” with both PolyPhen2 and SIFT. Q27X non-sense variant on *C1QC* was predicted as “disease causing” with MutationTaster. Our study also demonstrated a homozygous non-sense alteration in *C1QA* (Q208X/c.622C>T) in 2 patients (Patients 1 and 4), a homozygous missense alteration in *C1QC* (G34R/c.100G>A; Patient 3) and a homozygous frameshift alteration in *DNASE1L3* (T971fs*2/c.289_290delAC) (Patient 6) in 1 case each, which are previously reported by different studies\(^11,12,13\). All indicated variants were confirmed with Sanger sequencing for index cases as well as their affected siblings. Heterozygosity for parents and genotypes of healthy siblings were also shown (Figure 1 and Figure 2). Following WES, within 15 different chromosomes, about 126 Mb ROH regions, which are larger than 1.5 Mb, were identified for Patient 7 in total. Variant filtering revealed 6 homozygous missense variants (SNV) in different protein coding genes, which met the filtering criteria (Supplementary Table 2, available from the authors on request): a c.1472C>T in the ubiquitin specific peptidase 17-like family member 2 (*USP17L2*), a c.887A>C in the olfactory receptor family 5 subfamily M member 10 (*OR5M10*), a c.163C>T in the macrophase-expressed gene 1 (*MPEG1*), a c.28T>C in the Calpain 1 (*CAPN1*), a c.238C>T in the histone decetylase 7 (*HDAC7*), and a c.488T>C in the NK2 Homeobox 4 (*NKX2-4*). Sanger sequencing of the index case and additional family members suggested that the most probable
candidate variant segregating with the affected phenotype was the HDAC7 p.R80C variant (chr12:48192588G>A; Figure 3A and 3B). We should stress that both Patient 7 and her 3.5-year-old sister were homozygous for OR5M10, MPEG1, and CAPN1 variants. Her sister did not have any signs or symptoms of SLE. However, it is too early to decide whether her sister will be affected because our index case had symptoms when she was 5 years old.

The patients with C1Q variants (Patients 1–4) have some common features such as skin involvement, negative anti-dsDNA, and normal C3 and C4 levels (Table 1 and Table 2). CH50 levels were low in all these patients. The patients with C1Q and C1S variants (Patients 1–5) improved with the addition of fresh frozen plasma therapy to conventional SLE treatment. There was no history of recurrent infections in these patients. Of note, currently, Patient 3 (with C1Q variant) experienced a flare of alopecia, malar rash, and oral aphthosis. Patient 5 (with C1S variant) had a flare of rash, oral ulcers, and neurologic involvement.

Regarding family histories, Patients 1 (C1Q variant), 2 (C1Q variant), and 5 (C1S variant) had affected siblings. The sister of Patient 1 had disease onset at 3 years of age, skin
involvement (malar and discoid rash, photosensitivity), oral ulcers, nephritis, arthritis, normal C3 and C4 levels, positive antinuclear antibody (ANA), anti-SSA, and negative anti-dsDNA antibody. The brother of Patient 2 had disease onset at 4.5 years of age, malar and discoid rash, oral ulcers, arthritis, positive ANA, and hypocomplementemia. Patient 5 had an affected sister and brother. Her sister had disease onset at 5 years of age, malar and discoid rash, normal C3 and C4 levels, positive ANA, and anti-SSA antibodies. Her brother had disease onset at 5.5 years of age, oral ulcers, arthritis, neurologic involvement, normal C3 and C4 levels, positive ANA, anti-SM, and U1-RNP antibodies.

The patient with \textit{DNASE1L3} alteration (Patient 6) had a severe disease course. She received cyclophosphamide, methotrexate, and corticosteroid treatments. She is stable now on corticosteroid and hydroxychloroquine treatment with rare flares of skin and eye manifestations. It should be noted that the mother of Patient 6 had asthma and postpartum cardiomyopathy.

The patient with \textit{HDAC7} variant also had a severe disease course. She has been stable for 7 months while taking oral prednisolone, mycophenolate mofetil, and hydroxychloroquine. We performed Sanger sequencing for segregation of the indicated variant in \textit{HDAC7} gene within family members. Her parents were heterozygous while 2 siblings were carriers or negative for \textit{HDAC7} variant, which was located in a highly evolutionary conserved region of \textit{HDAC7} (Figure 3C and 3D).

None of the other family members except the ones mentioned above described features associated with autoimmunity.

**DISCUSSION**

In our study, Patients 1–5 had homozygous alterations in the genes coding for early complement proteins (\textit{C1Q} and \textit{C1S}), Patient 6 had a \textit{DNASE1L3} alteration, and we found a novel candidate variant in \textit{HDAC7} in Patient 7.

Primary complement defects have long been known to lead to an increased susceptibility to SLE\textsuperscript{14}. Complement deficiencies (especially early components of the classical pathway) causing SLE have been one of the classic examples of single-gene alterations leading to the phenotype of a multifactorial disease\textsuperscript{4}. Such monogenic disorders tend to present early in life or have familial causes because they are transmitted in an often-autosomal recessive fashion. Although C1q, C1s, and C1r complement deficiencies are rare, these are associated with a high risk of developing pediatric SLE, estimated at around 90% for C1q and 65% for C1r/s\textsuperscript{4}.

Three different genes (\textit{C1QA}, \textit{C1QB}, \textit{C1QC}) on the short arm of chromosome 1 encode for the C1q protein\textsuperscript{15}. C1q is
### Table 1. Characteristics and treatment of pediatric patients with early-onset systemic lupus erythematosus (SLE).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Pt 1</th>
<th>Pt 2</th>
<th>Pt 3</th>
<th>Pt 4</th>
<th>Pt 5</th>
<th>Pt 6</th>
<th>Pt 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consanguinity</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>SLE in sibling</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Gene</td>
<td>C1QA</td>
<td>C1QC</td>
<td>C1QA</td>
<td>C1QC</td>
<td>C1QA</td>
<td>C1QA</td>
<td>C1QA</td>
</tr>
<tr>
<td>Mutation detection method</td>
<td>WES</td>
<td>WES</td>
<td>WES</td>
<td>WES</td>
<td>WES</td>
<td>WES</td>
<td>WES</td>
</tr>
<tr>
<td>Reported (ref. no.)</td>
<td>Known (10)</td>
<td>Novel</td>
<td>Known (10)</td>
<td>Novel</td>
<td>Known (10)</td>
<td>Novel</td>
<td>Known (10)</td>
</tr>
<tr>
<td>Age at symptom onset, mos</td>
<td>6</td>
<td>36</td>
<td>30</td>
<td>54</td>
<td>15</td>
<td>24</td>
<td>60</td>
</tr>
<tr>
<td>Arthritis</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Malar rash</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pericarditis</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Neurologic involvement</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Other features</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Treatment</td>
<td>CS, HQ, MTX, FFP</td>
<td>CS, HQ, AZA, FFP</td>
<td>CS, HQ, MTX, FFP</td>
<td>CS, AZA, MTX, FFP</td>
<td>CS, HQ, CYC, AZA, FFP</td>
<td>CS, HQ, CYC, AZA, FFP</td>
<td>CS, HQ, CYC, AZA, IVIG, aspirin</td>
</tr>
<tr>
<td>Most recent SLICC/ACR damage index</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>


### Table 2. Laboratory features of pediatric patients with early-onset systemic lupus erythematosus during active disease.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Pt 1</th>
<th>Pt 2</th>
<th>Pt 3</th>
<th>Pt 4</th>
<th>Pt 5</th>
<th>Pt 6</th>
<th>Pt 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anemia</td>
<td>+</td>
<td>—</td>
<td>+</td>
<td>—</td>
<td>+</td>
<td>—</td>
<td>+</td>
</tr>
<tr>
<td>Leukopenia</td>
<td>—</td>
<td>+</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>—</td>
<td>+</td>
</tr>
<tr>
<td>Thrombocytopenia</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Hypocomplementemia</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ana</td>
<td>1/160</td>
<td>1/640</td>
<td>1/160</td>
<td>1/320</td>
<td>1/100</td>
<td>1/160</td>
<td>1/160</td>
</tr>
<tr>
<td>Anti-dsDNA</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>ENA Anti-SM</td>
<td>—</td>
<td>NA</td>
<td>—</td>
<td>+</td>
<td>—</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>LAC</td>
<td>—</td>
<td>NA</td>
<td>+</td>
<td>NA</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>aCL IgG</td>
<td>—</td>
<td>NA</td>
<td>—</td>
<td>NA</td>
<td>—</td>
<td>—</td>
<td>+</td>
</tr>
<tr>
<td>aCL IgM</td>
<td>—</td>
<td>NA</td>
<td>+</td>
<td>NA</td>
<td>—</td>
<td>—</td>
<td>+</td>
</tr>
<tr>
<td>Other autoantibodies</td>
<td>NA</td>
<td>RF+</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

ANA: antinuclear antibody; ENA: extractable nuclear antigens; aCL: anticardiolipin antibodies; LAC: lupus anticoagulant; RF: rheumatoid factor; NA: not available.

the first subcomponent of the classical complement pathway. The genes for C1s and C1r are located on the short arm of chromosome 12. C1r/s interacts with the hinge region of C1q to form the activated C1 complex. Then, C1s in this complex activates C4 and further activation of the classical complement pathway occurs. When C1 complex is not functional, the classical complement pathway cannot be activated and consumption of C4, C2, and C3 is reduced. This may result in normal to high levels of these complement proteins in serum, which was the case in Patients 1–4. Early complement proteins are important in decreasing the number of autoantibodies by their role in the clearance of...
Figure 3. A. Pedigree of index case with the segregation of p.Arg80Cys variant in HDAC7 included. B. Variant filtering. 50,992 variants passed Q15 and over 5x coverage for SNP, passed Q20 and over 10x for INDEL were detected totally in exome sequencing. After filtering steps, 6 candidate variants were selected for further analysis. Only one of them segregated with disease in the family. C. Sanger sequencing of p.Arg80Cys (c.238C>T) alteration in HDAC7. D. Multiple sequence alignment by Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo). The asterisk indicates positions that have a single, fully conserved residue, whereas colons indicate conserved substitutions and periods indicate semiconserved substitutions. Black frame indicates evolutionary conservation of the arginine at position 80. ANA: antinuclear antibody; asx: symptomatic; NA: not available; SLE: systemic lupus erythematosus; MAF: minor allele frequency; NM: nonmutated.
developed SLE at age 5.5 years\(^1\). The authors have demonstrated severely reduced levels of DNASE1L3 in homozygotes as expected because of non-sense–mediated mRNA decay triggered by the premature termination codon\(^1\). Al-Mayouf, \textit{et al} also reported a deletion in \textit{DNASE1L3}, which is associated with SLE in the study of 7 SLE families with a probable autosomal recessive pattern of inheritance\(^3\). They also showed that protein encoded by the mutant \textit{DNASE1L3} completely lacked DNase activity\(^3\). In fact, a GWAS peak was observed on 3p14.3 to be significantly associated with the risk of developing SLE\(^4\). This peak was attributed to \textit{PXK} gene, which encodes the Phox homology domain-containing serine-threonine kinase. However, subsequent studies did not replicate the \textit{PXK} association in SLE\(^4,4^4\). Because \textit{DNASE1L3} is only 140 kb upstream of \textit{PXK}, the variants in this gene might have also generated the GWAS peak, as speculated by Al-Mayouf, \textit{et al}\(^3\). When we look at the clinical characteristics of the reported SLE patients with \textit{DNASE1L3} variants (\(n = 18\))\(^11,3^8\); all of them have hypocomplementemia, positive ANA, and anti-dsDNA, while most had early disease onset, nephritis, and positive antineutrophil cytoplasmic antibodies. Our patient (Patient 6) also had early disease onset, hypocomplementemia, positive ANA, anti-dsDNA, and nephritis.

HDAC are chromatin-modifying enzymes that are involved in regulation of many aspects of cell biology including tissue differentiation, autophagy, apoptosis, migration, mitosis, and angiogenesis\(^4^5\). There are 11 different \textit{HDAC} genes encoding HDAC proteins grouped by their homology to yeast orthologues: class I (\textit{HDAC} 1–3, 8), class IIa (\textit{HDAC} 4, 5, 7, 9), class IIb (\textit{HDAC} 6 and 10), and class V (\textit{HDAC} 11)\(^4^5\). The role of individual HDAC proteins has not been fully identified. HDAC inhibitors have previously been considered for the treatment of SLE based on \textit{in vitro} cell culture and \textit{in vivo} studies in immune models of SLE; however, clinical efficacy remains to be established\(^4^6\). HDAC inhibition increases the production and suppressive function of regulatory T cells and decreases dendritic cell function. In addition, it is effective in the reduction of Th1- and Th17- inducing cytokines, which are important in SLE pathogenesis. In the case of identified alteration affecting the function of HDAC, the immune balance may favor autoimmunity, which may cause SLE. We will proceed with the functional study for this new candidate variant that we identified in 1 patient.

We have found alterations in early complement proteins, DNase enzyme, and a new candidate gene, \textit{HDAC7}, in pediatric patients with early-onset SLE and family history suggesting autosomal recessive inheritance. WES is an effective method for identifying clinically significant exonic variants. However, there are some limitations for evolutionarily conserved regulator DNA elements in untranslated, intronic, and intergenic regions that may be associated with the disease. Whole genome sequencing can also identify small copy number variations and mitochondrial DNA mutations, as well as exonic SNV missed by WES\(^4^7\). The definition of these single-gene disorders will shed further light on disease mechanisms.
REFERENCES


